IBM SPSS Modeler 18.3 — podręcznik użytkownika

Uwaga

Przed skorzystaniem z niniejszych informacji oraz produktu, którego one dotyczą, należy zapoznać się z informacjami zamieszczonymi w sekcji “Uwagi” na stronie 269.

Informacje o produkcie

Niniejsze wydanie publikacji dotyczy wersji 18, wydania 3, modyfikacji 0 produktu IBM® SPSS Modeler oraz wszystkich następnych wydań i modyfikacji do czasu, aż w kolejnym wydaniu publikacji zostanie zawarta informacja o stosownej zmianie.

Spis treści

Rozdział 1. IBM SPSS Modeler – informacje 1
Produkty IBM SPSS Modeler 1
IBM SPSS Modeler 1
IBM SPSS Modeler Server 1
IBM SPSS Modeler Administration Console 2
IBM SPSS Modeler Batch 2
IBM SPSS Modeler Solution Publisher 2
Adaptery serwera IBM SPSS Modeler Server dla IBM SPSS Collaboration and Deployment Services 2
Wydania programu IBM SPSS Modeler 2
Dokumentacja 3
Dokumentacja SPSS Modeler Professional 3
Dokumentacja SPSS Modeler Premium 4
Przykłady aplikacji 4
Folder Demos 4
Monitorowanie wykorzystania licencji 5

Rozdział 2. Nowe funkcje w programie IBM SPSS Modeler 18.3.0 7

Rozdział 3. Przegląd produktu 9
Pierwsze kroki 9
Uruchamianie produktu IBM SPSS Modeler 9
Uruchamianie z wiersza komend 9
Łączenie się z serwerem IBM SPSS Modeler Server 10
Łączenie się z serwerem Analytic Server 12
Zmiana katalogu tymczasowego 13
Uruchamianie wielu sesji programu IBM SPSS Modeler 13
Przegląd interfejsu oprogramowania IBM SPSS Modeler 14
Obszar roboczy strumienia IBM SPSS Modeler 14
Paleta węzłów 15
Panele zarządzania w programie IBM SPSS Modeler 16
Projekty w programie IBM SPSS Modeler 17
Pasek narzędzi programu IBM SPSS Modeler 18
Dostosowywanie paska narzędzi 19
Dostosowywanie okna programu IBM SPSS Modeler 19
Zmiana rozmiaru ikony strumienia 20
Korzystanie z myszy w oprogramowaniu IBM SPSS Modeler 21
Używanie skrótów klawiaturowych 21
Drukowanie 22
Automatyzacja procesów programu IBM SPSS Modeler 23

Rozdział 4. Omówienie eksploracji danych 25
Eksploracja danych — przegląd 25
Uzyskiwanie dostępu do danych 26
Strategia eksploracji danych 27
Model procesu CRISP-DM 28
Typy modeli 29
Przykłady eksploracji danych 34

Rozdział 5. Tworzenie strumieni 35

Tworzenie strumienia — przegląd....35
Tworzenie strumieni danych....35
Praca z węzłami....35
Praca ze strumieniami....41
Opisy strumieni....53
Uruchamianie strumieni....55
Praca z modelami....55
Dodawanie komentarzy i adnotacji do węzłów i strumieni....56
Zapisywanie strumieni danych....61
Ładowanie plików....63
Mapowanie strumieni danych....63
Porady i skróty....65

Rozdział 6. Praca z danymi....67
Budowanie wykresów....67
Układ kreatora wykresów i związane z nim terminy....67
Budowanie wykresu za pomocą galerii typów wykresów....68
Typy wykresów....68
Kokpit....104
Preferencje wizualizacji globalnej....104

Rozdział 7. Praca z wynikami....107
Okno raportu....107
Wyświetlanie i ukrywanie wyników....107
Przenoszenie, usuwanie i kopiowanie wyników....108
Zmiana początkowego wyrównania....108
Zmiana wyrównania wyników....108
Struktura obszaru Okno raportu....108
Edytowanie lub dodawanie elementów w obszarze Okno raportu....109
Znajdowanie i zastępowanie informacji w oknie raportu....110
Kopiowanie wyniku do innych aplikacji....111
Wyniki interaktywne....112
Eksportowanie raportu wynikowego....113
Opcje formatu HTML....114
Opcje raportu WWW....114
Opcje formatu Word....115
Opcje formatu Excel....116
Opcje formatu PowerPoint....116
Opcje PDF....117
Opcje formatu tekstowego....118
Opcja eksportu samych obrazów....118
Opcje formatu graficznego....119
Drukowanie z Edytora raportów....120
Drukowanie raportu i wykresów....120
Podgląd wydruku....120
Atrybuty strony: nagłówki i stopki....120
Atrybuty strony w opcjach wydruku....121
Zapisywanie wyników....122
Zapisywanie dokumentu okna raportu....122
Tabele przestawne....123
Tabele przestawne....123
Manipulowanie tabelą przestawną....123
Praca z warstwami....126
Wyświetlanie i ukrywanie pozycji....127
Szablony TableLook....127
Właściwości tabeli....128
Właściwości komórki....131

Przypisy i nagłówki ... 131
Szerokość komórek danych ... 133
Zmiana szerokości kolumny ... 133
Wyświetlanie ukrytych krawędzi tabeli przestawnej ... 134
Zaznaczanie wierszy, kolumn i komórek w tabeli przestawnej ... 134
Drukowanie tabel przestawnych ... 134
Tworzenie wykresu na podstawie tabeli przestawnej ... 135
Starsze wersje tabel ... 135
Opcje ... 135
Opcje ... 135
Opcje ogólne ... 136
Opcje raportów ... 136
Opcje tabel przestawnych ... 137
Opcje wyników ... 137

Rozdział 8. Obsługa braków danych ... 139
Braki danych — przegląd ... 139
Obsługa braków danych ... 139
Obsługa rekordów z brakami danych ... 140
Obsługa zmiennych z brakami danych ... 140
Obsługa rekordów z systemowymi brakami danych ... 141
Podstawianie lub wypełnianie braków danych ... 143
Funkcje języka CLEM do obsługi braków danych ... 143

Rozdział 9. Tworzenie wyrażeń CLEM ... 145
Informacje o produkcie CLEM ... 145
CLEM — przykłady ... 145
Wartości i typy danych ... 147
Wyrażenia i warunki ... 148
Parametry strumienia, sesji i superwęzła ... 148
Praca z łańcuchami ... 149
Obsługa wartości pustych i braków danych ... 150
Praca z liczbami ... 150
Praca z godzinami i datami ... 150
Podsumowanie wielu zmiennych ... 151
Praca z danymi opisującymi wielokrotne odpowiedzi ... 152
Konstruktor wyrażeń ... 153
Uzyskiwanie dostępu do Konstruktora wyrażeń ... 153
Tworzenie wyrażeń ... 153
Wybieranie funkcji ... 153
Wybór zmiennych, parametrów i zmiennych globalnych ... 157
Wyświetlanie lub wybieranie wartości ... 157
Sprawdzanie wyrażeń CLEM ... 157
Znajdowanie i zastępowanie ... 158

Rozdział 10. CLEM — informacje o języku ... 161
CLEM — przegląd informacji ... 161
Typy danych CLEM ... 161
Liczby całkowite ... 161
Liczby rzeczywiste ... 162
Znaki ... 162
Łańcuchy ... 162
Listy ... 162
Zmienne ... 163
Daty ... 163
Czas ... 164
Operatory CLEM ... 164

Funkcje — informacje....167
Konwencje zastosowane w opisach funkcji....168
Funkcje informacyjne....169
Funkcje przekształcania....169
Funkcje porównawcze....171
Funkcje logiczne....173
Funkcje numeryczne....174
Funkcje trygonometryczne....175
Funkcje prawdopodobieństwa....176
Funkcje przestrzenne....176
Operacje bitowe na liczbach całkowitych....177
Funkcje losowe....179
Funkcje łańcuchowe....179
Funkcje SoundEx....185
Funkcje daty i czasu....186
Funkcje sekwencji....191
Funkcje globalne....197
Funkcje obsługujące wartości puste i null....198
Funkcje specjalne....199

Rozdział 11. Używanie programu IBM SPSS Modeler razem z repozytorium....203
Informacje o programie IBM SPSS Collaboration and Deployment Services Repository....203
Przechowywanie i wdrażanie obiektów repozytorium....204
Łączenie się z repozytorium....204
Wprowadzanie danych uwierzytelniających do repozytorium....204
Przeglądanie danych uwierzytelniających repozytorium....205
Przeglądanie zawartości repozytorium....205
Zapisywanie obiektów w repozytorium....205
Określanie właściwości obiektów....205
Składowanie strumieni....207
Składowanie projektów....208
Składowanie węzłów....208
Składowanie obiektów wynikowych....209
Składowanie modeli i palet modeli....209
Pobieranie obiektów z repozytorium....210
Wybieranie obiektu do pobrania....210
Wybieranie wersji obiektu....211
Wyszukiwanie obiektów w repozytorium....211
Modyfikowanie obiektów w repozytorium....212
Tworzenie, zmiana nazwy i usuwanie folderów....212
Blokowanie i usuwanie blokady obiektów repozytorium....212
Usuwanie obiektów repozytorium....213
Zarządzanie właściwościami obiektów w repozytorium....213
Wyświetlanie właściwości folderu....213
Wyświetlanie i edytowanie właściwości obiektu....214
Zarządzanie etykietami wersji obiektu....215
Wdrażanie strumieni....215
Opcje wdrażania strumieni....216
Gałąź oceniania....217

Rozdział 12. Zapisywanie strumieni na serwerze IBM Cloud Pak for Data....221

Rozdział 13. Eksportowanie do aplikacji zewnętrznych....223
Informacje dotyczące eksportowania do aplikacji zewnętrznych....223
Otwieranie strumienia w programie IBM SPSS Modeler Advantage....223
Importowanie i eksportowanie modeli w formacie PMML....224
Typy modeli obsługujące język PMML....224

Rozdział 14. Projekty i raporty..... 227
Wprowadzenie do projektów..... 227
Widok CRISP-DM..... 227
Widok Klasy..... 228
Tworzenie projektu..... 228
Tworzenie nowego projektu..... 228
Dodawanie do projektu..... 228
Przenoszenie projektów do repozytorium IBM SPSS Collaboration and Deployment Services Repository..... 229
Ustawianie właściwości projektu..... 230
Dodawanie adnotacji do projektu..... 230
Właściwości obiektu..... 231
Zamykanie projektu..... 231
Generowanie raportu..... 231
Zapisywanie i eksportowanie wygenerowanych raportów..... 232

Rozdział 15. Dostosowywanie produktu IBM SPSS Modeler 235
Dostosowywanie opcji produktu IBM SPSS Modeler..... 235
Ustawianie opcji programu IBM SPSS Modeler..... 235
Opcje systemowe..... 235
Ustawianie katalogów domyślnych..... 236
Określanie opcji użytkownika..... 236
Dostosowywanie palety węzłów..... 244
Dostosowywanie ustawień opcji Zarządzanie paletami..... 245
Zmiana widoku karty palety..... 247

Rozdział 16. Czynniki wpływające na wydajność dotyczące strumieni i węzłów... 249
Kolejność węzłów..... 249
Pamięci podręczne węzłów..... 250
Wydajność: węzły procesowe..... 251
Wydajność: węzły modelowania..... 252
Wydajność: wyrażenia CLEM..... 252

Rozdział 17. Ułatwienia dostępu w programie IBM SPSS Modeler 255
Przegląd ułatwień dostępu w programie IBM SPSS Modeler..... 255
Rodzaje funkcji ułatwiających dostęp..... 255
Ułatwienia dostępu dla użytkowników słabo widzących..... 255
Ułatwienia dostępu dla użytkowników niewidomych..... 256
Ułatwienia dostępu z użyciem klawiatury..... 256
Używanie lektora ekranowego..... 264
Porady..... 265
Problemy podczas współpracy z innym oprogramowaniem..... 266
JAWS i Java..... 266
Używanie wykresów w programie IBM SPSS Modeler..... 266

Rozdział 18. obsługa Unicode..... 267
Obsługa kodowania Unicode w IBM SPSS Modeler..... 267

Uwagi..... 269
Znaki towarowe..... 270
Warunki dotyczące dokumentacji produktu..... 271

Indeks..... 273

Rozdział 1. IBM SPSS Modeler – informacje

IBM SPSS Modeler to zestaw narzędzi do eksploracji danych. Produkt umożliwia szybkie opracowywanie modeli predykcyjnych przy wykorzystaniu wiedzy specjalistycznej i stosowanie tych modeli w procesach biznesowych, jako wsparcia przy podejmowaniu decyzji. Rozwiązania zawarte w oprogramowaniu IBM SPSS Modeler zapewniają możliwość wykorzystywania branżowego modelu CRISP-DM i pozwalają na obsługę całego procesu eksploracji danych: od pozyskiwania danych do uzyskiwania lepszych wyników biznesowych.

Oprogramowanie IBM SPSS Modeler umożliwia korzystanie z wielu metod modelowania opartych na sztucznej inteligencji, uczeniu maszynowym i statystykach. Metody dostępne na palecie Modelowanie pozwalają na ekstrahowanie nowych informacji z danych i tworzenie modeli predykcyjnych. Każda metoda ma określone mocne strony i jest dostosowana do rozwiązywania określonych problemów.

Oprogramowanie SPSS Modeler można zakupić jako produkt samodzielny lub jako program kliencki używany wraz z oprogramowaniem SPSS Modeler Server. Dostępnych jest wiele opcji dodatkowych, które przedstawiono w kolejnych rozdziałach. Aby uzyskać więcej informacji, patrz https://www.ibm.com/analytics/us/en/technology/spss/.

Produkty IBM SPSS Modeler

Rodzina produktów IBM SPSS Modeler i towarzyszącego im oprogramowania składa się z elementów przedstawionych poniżej.

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console (jest częścią produktu IBM SPSS Deployment Manager)
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- Adaptery serwera IBM SPSS Modeler Server dla IBM SPSS Collaboration and Deployment Services

IBM SPSS Modeler

Oprogramowanie SPSS Modeler to w pełni funkcjonalna wersja produktu instalowana i uruchamiana na komputerze osobistym. Oprogramowanie SPSS Modeler można uruchomić lokalnie jako produkt samodzielny lub korzystać z niego w trybie rozproszonym wraz z serwerem IBM SPSS Modeler Server. Tego typu rozwiązanie zapewnia zwiększenie wydajności obsługi dużych zbiorów danych.

Dzięki oprogramowaniu SPSS Modeler można szybko tworzyć dokładne modele predykcyjne, stosując intuicyjne metody niewymagające umiejętności programowania. Unikatowy interfejs graficzny pozwala na wizualizowanie procedur eksploracji danych. Zaawansowane metody opracowywania analiz dostępne w programie umożliwiają określanie wcześniej niezauważalnych wzorców i trendów zawartych w danych. Użytkownik może modelować wyniki i poznawać czynniki wpływające na ich wartości. W ten sposób można wykorzystywać nowe szanse biznesowe i obniżać ryzyko.

Dostępne są dwie edycje oprogramowania SPSS Modeler: SPSS Modeler Professional oraz SPSS Modeler Premium. Więcej informacji można znaleźć w temacie “ Wydania programu IBM SPSS Modeler” na stronie 2.

IBM SPSS Modeler Server

Oprogramowanie SPSS Modeler działa w oparciu o architekturę klient-serwer, w której żądania wymagające zaangażowania dużych zasobów kierowane są do zaawansowanego oprogramowania serwerowego. Takie rozwiązanie umożliwia bardziej wydajną obsługę dużych zbiorów danych.

SPSS Modeler Server to produkt wymagający dodatkowej licencji, działający stale na serwerze w trybie analizy rozproszonej. Współpracuje on z co najmniej jedną instalacją oprogramowania IBM SPSS Modeler. W ten sposób oprogramowanie SPSS Modeler Server poprawia wydajność podczas obsługi dużych zbiorów danych, ponieważ operacje wymagające dużej mocy obliczeniowej można wykonywać na serwerze bez potrzeby pobierania danych na komputer kliencki. Oprogramowanie IBM SPSS Modeler Server optymalizuje również obsługę SQL i funkcje modelowania wewnątrz bazy danych, co dodatkowo zwiększa wydajność działania i sprzyja automatyzacji pracy.

IBM SPSS Modeler Administration Console

Oprogramowanie Modeler Administration Console to graficzny interfejs użytkownika służący do obsługi wielu opcji konfiguracji SPSS Modeler Server, które można dostosować również za pomocą pliku opcji. Konsola udostępniona w aplikacji IBM SPSS Deployment Manager pozwala na monitorowanie i konfigurowanie instalacji SPSS Modeler Server. Konsola jest dostępna bez dodatkowych opłat dla aktualnych użytkowników SPSS Modeler Server. Aplikację można zainstalować tylko na komputerach z systemem Windows, jednak administrować można serwerem zainstalowanym na dowolnej obsługiwanej platformie.

IBM SPSS Modeler Batch

Eksploracja danych jest zazwyczaj procesem interaktywnym, jednak oprogramowanie SPSS Modeler można też uruchomić z poziomu wiersza komend i zrezygnować z używania graficznego interfejsu użytkownika. Niekiedy użytkownik wykonuje długotrwałe lub powtarzalne zadania, które mogą być realizowane bez nadzoru. Oprogramowanie SPSS Modeler Batch to specjalna wersja produktu pozwalająca na wykonywanie wszystkich funkcji analitycznych SPSS Modeler bez potrzeby używania standardowego interfejsu użytkownika. Oprogramowanie SPSS Modeler Server jest wymagane do korzystania z aplikacji SPSS Modeler Batch.

IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher umożliwia tworzenie spakowanej wersji strumieni programu SPSS Modeler, które można uruchamiać za pomocą zewnętrznych środowisk wykonawczych lub osadzać w aplikacji zewnętrznej. W ten sposób można publikować i wdrażać pełne strumienie SPSS Modeler w celu używania ich w środowiskach, w których nie zainstalowano programu SPSS Modeler. SPSS Modeler Solution Publisher jest dystrybuowany jako część produktu IBM SPSS Collaboration and Deployment Services - Scoring, który do działania wymaga oddzielnej licencji. Wraz z licencją użytkownik otrzymuje oprogramowanie SPSS Modeler Solution Publisher Runtime umożliwiające uruchamianie opublikowanych strumieni.

Więcej informacji na temat SPSS Modeler Solution Publisher znajduje się w dokumentacji produktu IBM SPSS Collaboration and Deployment Services. W Dokumentacji IBM produktu IBM SPSS Collaboration and Deployment Services dostępne są sekcje „IBM SPSS Modeler Solution Publisher” oraz „IBM SPSS Analytics Toolkit”.

Adaptery serwera IBM SPSS Modeler Server dla IBM SPSS Collaboration and Deployment Services

Dostępnych jest wiele adapterów dla IBM SPSS Collaboration and Deployment Services, które umożliwiają współpracę programów SPSS Modeler i SPSS Modeler Server z repozytorium IBM SPSS Collaboration and Deployment Services. Dzięki temu strumień SPSS Modeler wdrożony w repozytorium można udostępnić wielu użytkownikom lub uzyskać do niego dostęp z poziomu uproszczonej aplikacji klienckiej IBM SPSS Modeler Advantage. Adapter należy zainstalować na systemie hostującym repozytorium.

Wydania programu IBM SPSS Modeler

Dostępne są następujące wydania oprogramowania SPSS Modeler.

SPSS Modeler Professional

Oprogramowanie SPSS Modeler Professional zapewnia wszystkie narzędzia wymagane do obsługi większości typów danych ustrukturyzowanych, takich jak np. zachowania i interakcje śledzone w systemach CRM, dane demograficzne, zachowania zakupowe i dane sprzedażowe.

SPSS Modeler Premium

Oprogramowanie SPSS Modeler Premium wymaga oddzielnej licencji. Dzięki temu rozwiązaniu oprogramowanie SPSS Modeler Professional może obsługiwać wyspecjalizowane dane oraz nieustrukturyzowane dane tekstowe. SPSS Modeler Premium obejmuje IBM SPSS Modeler Text Analytics:

Program **IBM SPSS Modeler Text Analytics** korzysta z zaawansowanych rozwiązań lingwistycznych oraz przetwarzania języka naturalnego w celu szybkiego przetwarzania różnego rodzaju nieustrukturyzowanych danych tekstowych, wyodrębniania i porządkowania kluczowych pojęć oraz grupowania tych pojęć w kategorie. Wyodrębnione pojęcia i kategorie można łączyć z istniejącymi danymi ustrukturyzowanymi, takimi jak dane demograficzne, a następnie stosować w celu modelowania, korzystając z produktu IBM SPSS Modeler i zawartego w nim pełnego pakietu narzędzi do eksploracji danych, aby w rezultacie takiego połączenia podejmować lepsze decyzje przy zmniejszonej ilości zakłóceń.

IBM SPSS Modeler Subscription

IBM SPSS Modeler Subscription oferuje te same funkcje analiz predykcyjnych, co tradycyjny klient IBM SPSS Modeler. Użytkownicy edycji Subscription mogą regularnie pobierać aktualizacje produktu.

Dokumentacja

Dokumentacja jest dostępna w programie SPSS Modeler z poziomu menu Pomoc. Spowoduje to otwarcie internetowej Dokumentacji IBM, która jest zawsze dostępna poza produktem.

Pełna dokumentacja dla każdego produktu (obejmująca instrukcje instalacji) jest dostępna również w formacie PDF, w osobnym, skompresowanym folderze, w ramach materiałów dotyczących produktu do pobrania. Najnowsze dokumenty PDF można także pobrać z Internetu pod adresem https://www.ibm.com/support/pages/spss-modeler-183-documentation.

Dokumentacja SPSS Modeler Professional

Pakiet dokumentacji produktu SPSS Modeler Professional (bez instrukcji instalacyjnych) zawiera następujące publikacje.

- **IBM SPSS Modeler — podręcznik użytkownika.** Ogólne wprowadzenie do obsługi oprogramowania SPSS Modeler, w tym opisy procedur tworzenia strumieni danych, obsługi braków danych, tworzenia wyrażeń CLEM, pracy z projektami i raportami oraz przygotowywania strumieni do wdrożenia w IBM SPSS Collaboration and Deployment Services lub IBM SPSS Modeler Advantage.
- **IBM SPSS Modeler — węzły źródłowe, procesowe i wyników.** Opisy wszystkich węzłów używanych do odczytywania, przetwarzania i tworzenia wynikowych postaci danych w różnych formatach. Czyli wszystkich węzłów poza węzłami modelowania.
- **IBM SPSS Modeler — węzły modelowania.** Opisy wszystkich węzłów używanych do tworzenia modeli eksploracji danych. Oprogramowanie IBM SPSS Modeler umożliwia korzystanie z wielu metod modelowania opartych na sztucznej inteligencji, uczeniu maszynowym i statystykach.
- **IBM SPSS Modeler — podręcznik zastosowań.** Przykłady zawarte w niniejszym przewodniku stanowią skrócone informacje związane z konkretnymi metodami i technikami modelowania. Wersja elektroniczna tego podręcznika jest również dostępna z poziomu menu Pomoc. Więcej informacji można znaleźć w temacie “Przykłady aplikacji” na stronie 4.
- **IBM SPSS Modeler — podręcznik tworzenia skryptów w języku Python i automatyzacji.** Informacje na temat automatyzacji działania systemu za pomocą skryptów Python wraz z właściwościami służącymi do obsługi węzłów i strumieni.

- **IBM SPSS Modeler — podręcznik wdrażania.** Informacje na temat uruchamiania strumieni IBM SPSS Modeler przedstawione w postaci krokowych operacji wykonywanych podczas przetwarzania zadań w oprogramowaniu IBM SPSS Deployment Manager.
- **IBM SPSS Modeler CLEF Developer's Guide.** Z oprogramowaniem CLEF można zintegrować inne programy pozwalające na przetwarzanie danych lub obsługę algorytmów modelujących w postaci węzłów w IBM SPSS Modeler.
- **IBM SPSS Modeler — podręcznik eksploracji w bazie danych.** Informacje na temat wydajnego wykorzystywania bazy danych w celu zwiększenia wydajności i zakresu funkcji analitycznych za pomocą algorytmów innych firm.
- **IBM SPSS Modeler Server — podręcznik administracji i wydajności.** Informacje na temat konfiguracji i funkcji administracyjnych w oprogramowaniu IBM SPSS Modeler Server.
- **IBM SPSS Deployment Manager — Podręcznik użytkownika.** Informacje dotyczące korzystania z interfejsu użytkownika konsoli administracyjnej zawartej w aplikacji Deployment Manager podczas monitorowania i konfigurowania serwera IBM SPSS Modeler Server.
- **IBM SPSS Modeler — podręcznik CRISP-DM.** Szczegółowy podręcznik metodologii CRISP-DM w kontekście eksploracji danych za pomocą oprogramowania SPSS Modeler.
- **IBM SPSS Modeler Batch — podręcznik użytkownika.** Pełny podręcznik obsługi oprogramowania IBM SPSS Modeler w trybie wsadowym obejmujący szczegółowe informacje na temat pracy w trybie wsadowym i korzystania z argumentów z poziomu wiersza komend. Ten podręcznik jest dostępny tylko w formacie PDF.

Dokumentacja SPSS Modeler Premium

Pakiet dokumentacji produktu SPSS Modeler Premium (bez instrukcji instalacyjnych) zawiera następujące publikacje.

- **SPSS Modeler Text Analytics — podręcznik użytkownika.** Informacje na temat używania analiz tekstu za pomocą oprogramowania SPSS Modeler obejmują procedury dotyczące węzłów eksploracji tekstu, interaktywnego pulpitu roboczego, szablonów oraz innych zasobów.

Przykłady aplikacji

Podczas gdy narzędzia do eksploracji danych w programie SPSS Modeler mogą pomóc w rozwiązaniu szeregu problemów biznesowych i organizacyjnych, przykłady aplikacji udostępniają krótkie, ukierunkowane wprowadzenia do konkretnych metod i technik modelowania. Używane tutaj zestawy danych są znacznie mniejsze niż ogromne składnice danych zarządzane przez programy do eksploracji danych, lecz używane koncepcje i metody są skalowalne odpowiednio do potrzeb rzeczywistych aplikacji.

Dostęp do przykładów można uzyskać, klikając opcję **Przykłady aplikacji** w menu Pomoc programu SPSS Modeler.

Pliki danych i przykładowe strumienie są instalowane w folderze Dema, w katalogu instalacyjnym produktu. Aby uzyskać więcej informacji, patrz “Folder Demos” na stronie 4.

Przykłady modelowania w bazach danych. Przykłady zamieszczono w publikacji *IBM SPSS Modeler — podręcznik eksploracji w bazie danych*.

Przykłady skryptów. Przykłady zamieszczono w publikacji *IBM SPSS Modeler — podręcznik tworzenia skryptów i automatyzacji*.

Folder Demos

Pliki danych i przykładowe strumienie używane z przykładami do aplikacji są instalowane w folderze Demos wewnątrz katalogu instalacyjnego produktu (na przykład: `C:\Program Files\IBM\SPSS\Modeler\<version>\Demos`). Dostęp do tego folderu można także uzyskać z grupy programów IBM SPSS Modeler w menu Start systemu Windows lub klikając opcję Demos na liście ostatnich katalogów w oknie dialogowym **Plik** > **Otwórz strumień**.

Monitorowanie wykorzystania licencji

Podczas pracy z produktem SPSS Modeler wykorzystanie licencji jest monitorowanie i regularnie rejestrowane. Metryka wykorzystania licencji nosi nazwę *AUTHORIZED_USER* (użytkownik autoryzowany) lub *CONCURRENT_USER* (użytkownik pracujący jednocześnie), a typ rejestrowanej metryki zależy od typu licencji na produkt SPSS Modeler, którą posiada użytkownik.

Generowane pliki dzienników mogą być przetwarzane przez program IBM License Metric Tool, z którego uzyskać można raporty o wykorzystaniu licencji.

Pliki dzienników wykorzystania licencji są tworzone w tym samym katalogu, w którym zapisywane są dzienniki klienta SPSS Modeler (domyślnie `%ALLUSERSPROFILE%/IBM/SPSS/Modeler/<wersja>/log`).

Rozdział 2. Nowe funkcje w programie IBM SPSS Modeler 18.3.0

W tej wersji programu IBM SPSS Modeler wprowadzono następujące nowe funkcje i możliwości.

- Jedną z niedogodności związanych z modelowaniem jest dezaktualizacja modeli wynikająca ze zmian w danych źródłowych zachodzących na przestrzeni czasu. Zjawisko to jest często nazywane *dryftem modelu* lub *dryfem modelu*. SPSS Modeler oferuje teraz funkcję *ciągłego zautomatyzowanego uczenia maszynowego*, która pomaga w skutecznym niwelowaniu skutków dryftu modelu. Ciągłe uczenie maszynowe bazuje na rezultatach badań IBM i jest inspirowane biologicznym zjawiskiem naturalnej selekcji. Można z niego korzystać w węzłach Auto Klasyfikacja i Auto Predykcja.
- Strumienie można teraz przesyłać bezpośrednio z klienta IBM SPSS Modeler do serwera IBM Cloud Pak for Data. Aby uzyskać więcej informacji, zobacz Rozdział 12, “Zapisywanie strumieni na serwerze IBM Cloud Pak for Data”, na stronie 221.
- Aby włączyć rejestrowanie dla klienta IBM SPSS Modeler, otwórz plik `log4j2.xml` w edytorze tekstu i zmień wpis `level="info"` na `level="debug"` w tym wierszu:

```
<Logger name="com.spss" additivity="false" level="info">
```

- Do węzła GLE na karcie Opcje budowania w sekcji Oszacowanie parametru dodano nowe ustawienie **Użyj metody najmniejszych kwadratów z ograniczeniem nieujemności**. Metoda najmniejszych kwadratów z ograniczeniem nieujemności (NNLS) to ograniczony problem najmniejszych kwadratów, w którym współczynniki nie mogą przyjmować wartości ujemnych. Nie wszystkie zbiory danych są odpowiednie do estymacji NNLS, ponieważ wymaga ona, aby między predyktorami a zmienną przewidywaną istniała korelacja dodatnia lub nie istniała żadna korelacja.
- Do węzła źródłowego programu Excel dodano nowe ustawienie **Wiersze do przeskanowania dla kolumny i typu**. Wartością domyślną nowego ustawienia jest 200. Należy pamiętać, że zwiększenie liczby wierszy danych w formacie Excel analizowanych w celu określenia typu kolumny i składowania może mieć negatywny wpływ na wydajność.
- W węźle źródłowym Cognos TM1 i węźle eksportu IBM Cognos TM1 jest dostępna nowa metoda połączenia (Połączenie z serwerem TM1 Server) umożliwiająca współpracę z usługą Planning Analytics on Cloud. Serwer administracyjny został usunięty z produktu Planning Analytics on Cloud, dlatego jeśli masz stare strumienie zawierające węzły TM1 łączące się z usługą Planning Analytics on Cloud za pośrednictwem serwera TM1 Admin Server, możesz je zmodyfikować tak, aby odwoływały się do serwera TM1 Server.
- Obsługiwana jest teraz baza danych Db2 11.5.
- Obsługiwana jest teraz baza danych Db2 Warehouse.
- Obsługiwana jest teraz baza danych Db2 Big SQL 7.1.0 na platformie Cloudera Data Platform 7.1.5.
- Obsługiwany jest teraz mechanizm Apache Hive 3.1.3 na platformie Cloudera Data Platform 7.1.5.
- Obsługiwany jest teraz mechanizm Cloudera Impala 3.4.0 na platformie Cloudera Data Platform 7.1.5.
- Obsługiwana jest teraz baza danych Informix 14.10.
- Obsługiwany jest teraz system RedHat 8.3.
- Używana jest nowa wersja środowiska R (4.0.4).

Rozdział 3. Przegląd produktu

Pierwsze kroki

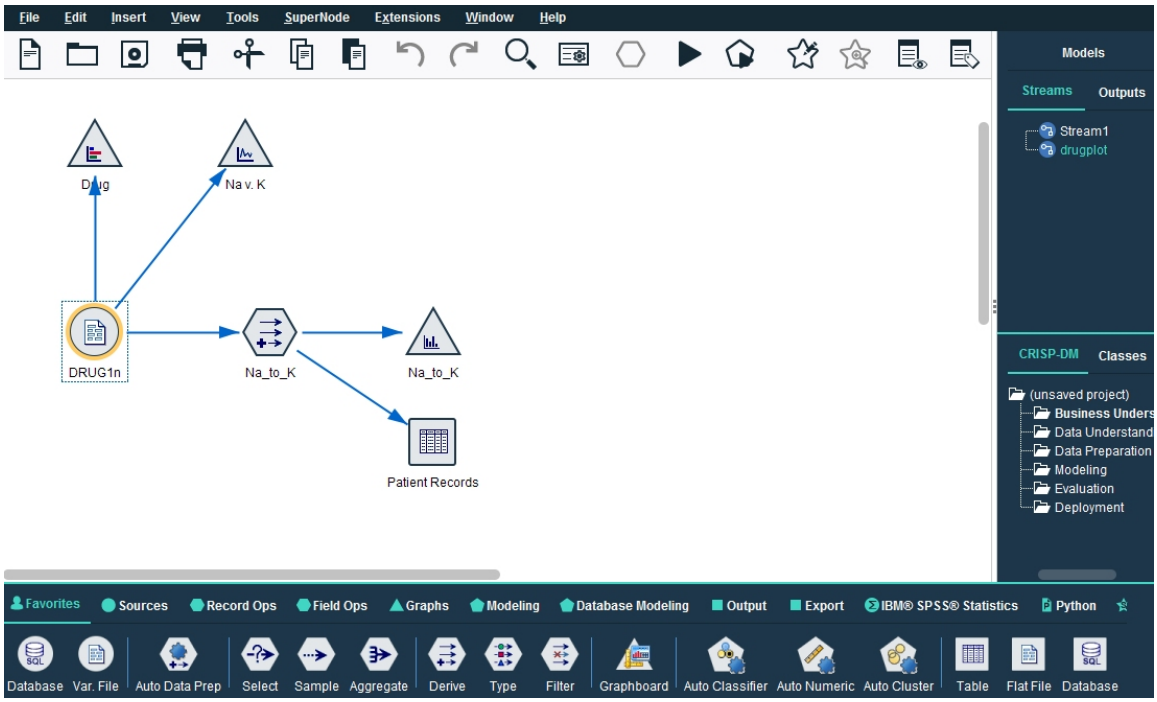
Będąc narzędziem do eksploracji danych, oprogramowanie IBM SPSS Modeler umożliwia uzyskanie użytecznych informacji o relacjach istniejących w dużych zbiorach danych. Inaczej niż w tradycyjnych metodach statystycznych użytkownik na początku nie musi wiedzieć, czego szuka. Dane można przeglądać, dopasowując różne modele, i badać różnorodne relacje do momentu znalezienia użytecznych informacji.

Uruchamianie produktu IBM SPSS Modeler

Aby uruchomić aplikację, kliknij:

Start > **[Wszystkie] Programy** > **IBM SPSS Modeler <wersja>** > **IBM SPSS Modeler <wersja>**

Okno główne zostanie wyświetlone po kilku sekundach.



Rysunek 1. Okno główne aplikacji IBM SPSS Modeler

Uruchamianie z wiersza komend

Można użyć wiersza komend systemu operacyjnego w celu uruchomienia programu IBM SPSS Modeler w następujący sposób:

1. Na komputerze, na którym zainstalowano program IBM SPSS Modeler, otwórz okno DOS lub okno wiersza komend.
2. Aby uruchomić interfejs IBM SPSS Modeler w trybie interaktywnym, wpisz komendę `modelerclient` wraz z wymaganymi argumentami, na przykład:

```
modelerclient -stream report.str -execute
```

Dostępne argumenty (flagi) umożliwiają nawiązanie połączenia z serwerem, załadowanie strumieni, uruchomienie skryptów lub określenie innych parametrów, stosownie do potrzeb.

Łączenie się z serwerem IBM SPSS Modeler Server

Serwer IBM SPSS Modeler można uruchamiać jako niezależną aplikację, jako klienta połączonego bezpośrednio z serwerem IBM SPSS Modeler Server lub jako klienta połączonego z serwerem IBM SPSS Modeler Server albo z klastrem serwerów za pośrednictwem wtyczki Coordinator of Processes programu IBM SPSS Collaboration and Deployment Services. Bieżący status połączenia jest wyświetlany w lewym dolnym rogu okna programu IBM SPSS Modeler.

Przy każdej próbie połączenia się z serwerem można ręcznie wprowadzić nazwę serwera, z którym ma zostać nawiązane połączenie, lub wybrać nazwę zdefiniowaną wcześniej. Jeśli jednak zainstalowany jest program IBM SPSS Collaboration and Deployment Services, możliwe jest przeszukiwanie listy serwerów lub klastrów serwerów z okna dialogowego Logowanie do serwera. Możliwość przeglądania usług Statistics działających w sieci jest dostępna za pośrednictwem usługi Coordinator of Processes.

Aby połączyć się z serwerem

1. W menu Narzędzia kliknij opcję **Logowanie do serwera**. Zostanie otwarte okno dialogowe Logowanie do serwera. Można też dwukrotnie kliknąć obszar statusu połączenia w oknie IBM SPSS Modeler.
2. Korzystając z okna dialogowego, podaj opcje umożliwiające nawiązanie połączenia z komputerem serwera lokalnego lub wybierz połączenie z tabeli.
   - Kliknij przycisk **Dodaj** lub **Edycja**, aby dodać lub edytować połączenie. Więcej informacji można znaleźć w temacie “Dodawanie i edytowanie połączenia z serwerem IBM SPSS Modeler Server” na stronie 11.
   - Kliknij opcję **Wyszukaj**, aby uzyskać dostęp do serwera lub klastra serwerów za pośrednictwem usługi Coordinator of Processes. Więcej informacji można znaleźć w temacie “Wyszukiwanie serwerów w programie IBM SPSS Collaboration and Deployment Services ” na stronie 11.

   Tabela serwerów. W tabeli tej znajduje się zestaw zdefiniowanych połączeń z serwerem. W tabeli wyświetlane są połączenie domyślne, nazwa serwera, opis i numer portu. Istnieje możliwość ręcznego dodania nowego połączenia, a także wyboru lub wyszukania istniejącego połączenia. Aby ustawić konkretny serwer jako połączenie domyślne, zaznacz pole wyboru w kolumnie Domyślne w tabeli.

   Domyślna ścieżka danych. Wskaż ścieżkę danych na komputerze serwera. Kliknij przycisk z wielokropkiem **(...)**, aby przejść do wymaganej lokalizacji.

   Ustaw dane uwierzytelniające. To pole wyboru powinno być niezaznaczone, aby możliwe było włączenie funkcji **pojedynczego uwierzytelniania**, która próbuje zalogować użytkownika na serwerze przy użyciu nazwy użytkownika i hasła na lokalnym komputerze. Jeśli pojedyncze logowanie nie jest możliwe, lub w przypadku zaznaczenia tego pola w celu wyłączenia pojedynczego logowania (na przykład, w celu zalogowania się na konto administratora) uaktywniane są poniższe pola umożliwiające wprowadzenie danych uwierzytelniających.

   Identyfikator użytkownika. Wprowadź nazwę użytkownika, z użyciem której należy się logować do serwera.

   Hasło. Wprowadź hasło powiązane z określoną nazwą użytkownika.

   Domena. Określ domenę używaną do zalogowania się na serwerze. Nazwa domeny jest wymagana tylko wówczas, gdy komputer serwera znajduje się w innej domenie Windows niż komputer kliencki.
3. Kliknij przycisk **OK**, aby zakończyć nawiązywanie połączenia.

Aby odłączyć się od serwera

1. W menu Narzędzia kliknij opcję **Logowanie do serwera**. Zostanie otwarte okno dialogowe Logowanie do serwera. Można też dwukrotnie kliknąć obszar statusu połączenia w oknie IBM SPSS Modeler.
2. W oknie dialogowym wybierz pozycję Lokalny serwer, a następnie kliknij przycisk **OK**.

Dodawanie i edytowanie połączenia z serwerem IBM SPSS Modeler Server

Połączenia z serwerem można edytować lub dodawać ręcznie w oknie dialogowym Logowanie do serwera. Klikając przycisk Dodaj, można wyświetlić puste okno dialogowe Dodaj/Edytuj serwer, w którym można wprowadzać szczegółowe dane połączenia z serwerem. Po wybraniu istniejącego połączenia i kliknięciu przycisku Edytuj w oknie dialogowym Logowanie do serwera zostanie otwarte okno dialogowe zawierające szczegóły tego połączenia, co pozwala wprowadzić dowolne zmiany.

Uwaga: Nie można edytować połączenia z serwerem dodanego z programu IBM SPSS Collaboration and Deployment Services, ponieważ nazwa, port i inne szczegóły są definiowane w programie IBM SPSS Collaboration and Deployment Services. Zalecane jest użycie tych samych portów do komunikacji zarówno z programem IBM SPSS Collaboration and Deployment Services, jak i z klientem SPSS Modeler Client. Można je ustawić pod parametrami `max_server_port` i `min_server_port` w pliku options.cfg.

Aby dodać połączenia z serwerem

1. W menu Narzędzia kliknij opcję **Logowanie do serwera**. Zostanie otwarte okno dialogowe Logowanie do serwera.
2. W tym oknie dialogowym kliknij przycisk **Dodaj**. Zostanie otwarte okno dialogowe Logowanie do serwera: Dodaj/Edytuj serwer.
3. Wprowadź szczegóły połączenia z serwerem, a następnie kliknij przycisk **OK**, aby zapisać połączenie i powrócić do okna dialogowego Logowanie do serwera.

- **Serwer.** Wskaż dostępny serwer lub wybierz serwer z listy. Komputer serwera można zidentyfikować za pośrednictwem nazwy alfanumerycznej (na przykład *moj_serwer*) lub adresu IP przypisanego do komputera serwera (na przykład 202.123.456.78).
- **Port.** Podaj numer portu, na którym serwer nasłuchuje. Jeśli wartość domyślna nie działa, poproś administratora systemu o podanie poprawnego numeru portu.
- **Opis.** Wprowadź opcjonalny opis dla tego połączenia z serwerem.
- **Zapewnij bezpieczne połączenie (z pomocą SSL).** Określa, czy ma zostać użyte bezpieczne połączenie SSL (**Secure Sockets Layer**). SSL jest powszechnie używanym protokołem do zarządzania bezpieczeństwem przekazywania danych w Internecie. Aby możliwe było użycie tej funkcji, protokół SSL musi być włączony na serwerze będącym hostem programu IBM SPSS Modeler Server. W razie konieczności skontaktuj się z lokalnym administratorem systemu w celu uzyskania bardziej szczegółowych informacji.

Aby edytować połączenia z serwerem

1. W menu Narzędzia kliknij opcję **Logowanie do serwera**. Zostanie otwarte okno dialogowe Logowanie do serwera.
2. W tym oknie dialogowym wybierz połączenie, które chcesz edytować, a następnie kliknij przycisk **Edycja**. Zostanie otwarte okno dialogowe Logowanie do serwera: Dodaj/Edytuj serwer.
3. Wprowadź szczegóły połączenia z serwerem, a następnie kliknij przycisk **OK**, aby zapisać zmiany i powrócić do okna dialogowego Logowanie do serwera.

Wyszukiwanie serwerów w programie IBM SPSS Collaboration and Deployment Services

Zamiast wprowadzać dane połączenia z serwerem ręcznie, można wybrać dostępny w sieci serwer lub klaster serwerów za pośrednictwem usługi Coordinator of Processes dostępnej w programie IBM SPSS Collaboration and Deployment Services. Klaster serwerów to grupa serwerów, z której usługa Coordinator of Processes określa serwer optymalny do odpowiedzi na żądanie przetwarzania.

Choć do okna dialogowego Logowanie do serwera można ręcznie dodawać serwery, wyszukiwanie dostępnych serwerów pozwala łączyć się z serwerami bez konieczności znajomości ich poprawnej nazwy i numeru portu. Informacje te są dostarczane automatycznie. Nadal jednak potrzebne są prawidłowe dane logowania: nazwa użytkownika, domena i hasło.

Uwaga: W przypadku braku dostępu do możliwości usługi Coordinator of Processes można nadal ręcznie wprowadzić nazwę serwera, z którym ma zostać nawiązane połączenie, lub można wybrać uprzednio zdefiniowaną nazwę. Więcej informacji można znaleźć w temacie "Dodawanie i edytowanie połączenia z serwerem IBM SPSS Modeler Server" na stronie 11.

Aby wyszukiwać serwery i klastry

1. W menu Narzędzia kliknij opcję **Logowanie do serwera**. Zostanie otwarte okno dialogowe Logowanie do serwera.
2. Kliknij przycisk **Szukaj** w tym oknie dialogowym, aby otworzyć okno dialogowe Szukaj serwerów. Jeśli podczas próby przeglądania danych usługi Coordinator of Processes okaże się, że użytkownik nie jest zalogowany do programu IBM SPSS Collaboration and Deployment Services, zostanie wyświetlony monit z prośbą o zalogowanie się.
3. Wybierz serwer lub klaster serwerów z listy.
4. Kliknij przycisk **OK**, aby zamknąć okno dialogowe i dodać to połączenie do tabeli w oknie dialogowym Logowanie do serwera.

Łączenie się z serwerem Analytic Server

Jeśli dostępnych jest kilka serwerów Analytic Server, użytkownik może użyć okna dialogowego Analytic Server Connection w celu zdefiniowania więcej niż jednego serwera, jaki będzie używany przez klienta IBM SPSS Modeler. Administrator mógł już skonfigurować domyślny serwer Analytic Server w pliku `<Modeler_install_path>/config/options.cfg`. Możliwe jest jednak użycie innych dostępnych serwerów po wcześniejszym ich zdefiniowaniu. Przykładowo, w przypadku użycia węzła źródłowego i węzła eksportu Analytic Server użytkownik może chcieć użyć różnych połączeń Analytic Server w różnych gałęziach strumienia, tak aby po uruchomieniu poszczególnych gałęzi korzystały one z własnego serwera Analytic Server i aby żadne dane nie były przekazywane do serwera IBM SPSS Modeler Server. Należy pamiętać, że jeśli gałąź zawiera więcej niż jedno połączenie Analytic Server, dane będą pobierane z serwerów Analytic Server na serwer IBM SPSS Modeler Server.

Aby utworzyć nowe połączenie serwera Analytic Server, należy wybrać opcje **Narzędzia** > **Połączenia programu Analytic Server** i wprowadzić wymagane informacje w następujących sekcjach okna dialogowego.

Połączenie

URL. Należy wpisać adres URL serwera Analytic Server w formacie `https://nazwa_hosta:port/obiekt_główny_kontekstu`, gdzie `nazwa_hosta` jest adresem IP lub nazwą hosta Analytic Server, `port` oznacza numer portu, a `obiekt_główny_kontekstu` oznacza obiekt główny kontekstu Analytic Server.

Podmiot użytkujący. Należy wpisać nazwę podmiotu użytkującego, którego serwer IBM SPSS Modeler Server jest elementem. Jeśli podmiot użytkujący nie jest znany, należy skontaktować się z administratorem.

Uwierzytelnianie

Dominanta. Należy wybrać jeden z następujących trybów uwierzytelniania.

- **Nazwa użytkownika i hasło** wymaga wprowadzenia nazwy użytkownika i hasła.
- **Zapisane dane uwierzytelniające** wymaga dokonania wyboru danych uwierzytelniających z repozytorium IBM SPSS Collaboration and Deployment Services Repository.
- **Kerberos** wymaga wprowadzenia głównej nazwy usługi (Service Principal Name, SPN) oraz ścieżki pliku konfiguracyjnego. Jeśli informacje te nie są znane, należy skontaktować się z administratorem.

Nazwa użytkownika. Należy wpisać nazwę użytkownika programu Analytic Server.

Dziedziny. Wybierz dziedzinę, która ma być używana dla połączenia z serwerem Analytic Server.

Password. Należy wpisać hasło do programu Analytic Server.

Połącz. Należy kliknąć przycisk **Połącz**, aby przetestować nowe połączenie.

Połączenia

Po wprowadzeniu powyższych informacji i kliknięciu przycisku **Połącz** połączenie zostanie dodane do tabeli połączeń. Jeśli konieczne jest usunięcie połączenia, należy je wybrać i kliknąć przycisk **Usuń**.

Jeśli administrator zdefiniował domyślne połączenie serwera Analytic Server w pliku `options.cfg`, można kliknąć przycisk **Dodaj domyślne połączenie**, aby je również dodać do dostępnych połączeń. Zostanie wyświetlone pytanie o nazwę użytkownika i hasło.

Zmiana katalogu tymczasowego

Wykonanie niektórych operacji w oprogramowaniu IBM SPSS Modeler Server może wymagać utworzenia plików tymczasowych. Domyślnie podczas obsługi plików tymczasowych program IBM SPSS Modeler korzysta z katalogu tymczasowego w systemie operacyjnym. Wykonując poniższe czynności, można zmienić położenie katalogu tymczasowego.

1. Utwórz nowy katalog o nazwie `spss` i podkatalog *servertemp*.
2. Edytuj plik *options.cfg* znajdujący się w katalogu */config* w katalogu instalacyjnym programu IBM SPSS Modeler. Edytuj parametr `temp_directory` zawarty w tym pliku i wprowadź w nim ciąg: `temp_directory, "C:/spss/servertemp"`.
3. Następnie ponownie uruchom usługę IBM SPSS Modeler Server. W tym celu można kliknąć kartę **Usługi** dostępną w Panelu sterowania systemu Windows. Zatrzymaj usługę i uruchom ją ponownie, aby zastosować wprowadzone zmiany. Ponowne uruchomienie komputera również spowoduje ponowne uruchomienie usługi.

Wszystkie pliki tymczasowe będą teraz zapisywane w nowym katalogu.

Uwaga: Należy stosować prawe ukośniki.

Uruchamianie wielu sesji programu IBM SPSS Modeler

Jeśli konieczne jest uruchomienie więcej niż jednej sesji programu IBM SPSS Modeler naraz, należy wprowadzić pewne zmiany w ustawieniach programu IBM SPSS Modeler i systemu Windows. Uruchomienie więcej niż jednej sesji może być konieczne na przykład w sytuacji, gdy mamy dwie odrębne licencje serwerowe i chcemy uruchomić dwa strumienie na dwóch różnych serwerach z tego samego komputera klienckiego.

Aby umożliwić jednoczesne działanie wielu sesji programu IBM SPSS Modeler:

1. Kliknij:

   Start > **[Wszystkie] Programy** > **IBM SPSS Modeler**
2. Kliknij prawym przyciskiem myszy na skrócie IBM SPSS Modeler (tym z ikoną), a następnie wybierz polecenie **Właściwości**.
3. W polu tekstowym **Element docelowy** dopisz `-noshare` na końcu łańcucha.
4. W Eksploratorze Windows wybierz:

   Narzędzia > **Opcje folderów...**
5. Na karcie Typy plików wybierz opcję Strumień IBM SPSS Modeler i kliknij przycisk **Zaawansowane**.
6. W oknie dialogowym Edytowanie typu pliku wybierz opcję Otwórz za pomocą IBM SPSS Modeler i kliknij przycisk **Edytuj**.
7. W polu tekstowym **Aplikacja używana do wykonania akcji** dopisz `-noshare` przed argumentem **-stream**.

Przegląd interfejsu oprogramowania IBM SPSS Modeler

Na każdym etapie procesu eksploracji danych użytkownik może skorzystać z prostego w obsłudze interfejsu IBM SPSS Modeler. Algorytmy modelujące, takie jak predykcja, klasyfikacja, segmentacja i wykrywanie asocjacji, pozwalają na tworzenie wydajnych i dokładnych modeli. Wyniki działania modelu można w prosty sposób wdrażać i odczytywać w bazach danych, w programie IBM SPSS Statistics i wielu innych aplikacjach.

Obsługa programu IBM SPSS Modeler polega na wykonaniu trzyetapowego procesu opracowywania danych.

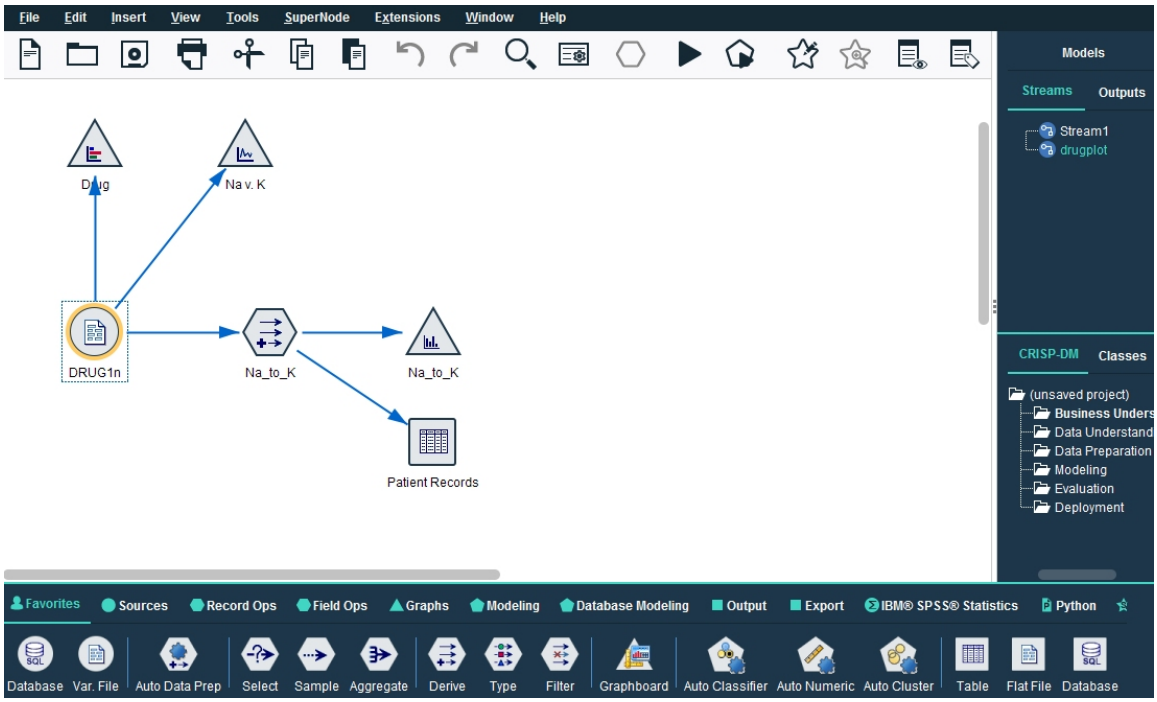
- Najpierw dane są wczytywane do oprogramowania IBM SPSS Modeler.
- Następnie na danych wykonywanych jest wiele operacji.
- Na koniec dane są przesyłane do położenia docelowego.

Ta sekwencja operacji jest nazywana **strumieniem danych**, ponieważ dane (rekord po rekordzie) wychodzą ze źródła, podlegają pewnym operacjom, a następnie są przekazywane do celu (modelu lub określonego typu danych wyjściowych).

Rysunek 2. Prosty strumień

Obszar roboczy strumienia IBM SPSS Modeler

Obszar roboczy strumienia to największa część okna IBM SPSS Modeler. W tym obszarze użytkownik tworzy strumienie danych i pracuje z nimi.



Rysunek 3. Obszar roboczy programu IBM SPSS Modeler (widok domyślny)

Tworzenie strumienia polega na narysowaniu (na głównym obszarze roboczym interfejsu) diagramu operacji biznesowych wykonywanych na danych. Każda operacja jest oznaczona ikoną lub **węzłem**, a węzły są łączone do postaci **strumieni** odzwierciedlających przepływ danych w ramach każdej operacji.

Jednocześnie w oprogramowaniu IBM SPSS Modeler można opracowywać wiele strumieni, korzystając z jednego obszaru roboczego strumienia lub wielu obszarów roboczych. W trakcie sesji strumienie są przechowywane w menedżerze strumieni dostępnym w prawym górnym narożniku okna IBM SPSS Modeler.

Uwaga: W przypadku korzystania z MacBooka z aktywowanym ustawieniem **Kliknięcie mocne i odpowiedzi haptyczne** wbudowanego gładzika, przeciąganie i upuszczanie węzłów z palety węzłów do kanwy strumienia może spowodować zduplikowanie węzłów dodawanych do kanwy. Aby uniknąć tego problemu, zalecamy wyłączenie preferencji systemu gładzika **Kliknięcie mocne i odpowiedzi haptyczne**.

Paleta węzłów

Większość narzędzi do pracy z danymi i modelowania danych w programie SPSS Modeler znajduje się na *palecie węzłów*, która rozciąga się u dołu okna pod obszarem roboczym strumienia.

Na przykład karta palety **Op. rekordu** zawiera węzły, które umożliwiają wykonywanie danych na *rekordach* danych, takich jak wybieranie, scalanie i dołączanie.

Aby dodawać węzły do obszaru roboczego, klikaj dwukrotnie ikony na palecie węzłów lub przeciągaj je do obszaru roboczego i tam upuszczaj. Następnie węzły można połączyć, tworząc *strumień* odzwierciedlający przepływ danych.

Rysunek 4. Karta Rekordy na palecie węzłów

Każda karta palety zawiera zbiór pokrewnych węzłów używanych w różnych fazach działania strumienia, na przykład:

- Węzły z grupy **Źródła** służą do wprowadzania danych do programu SPSS Modeler.
- Węzły z grupy **Rekordy** wykonują na *rekordach* danych takie operacje, jak wybieranie, scalanie i dołączanie.
- Węzły z grupy **Zmienne** wykonują operacje na *zmiennych*, np. filtrowanie, wywodzenie nowych zmiennych i określanie poziomu pomiaru danych zmiennych.
- Węzły z grupy **Wykresy** prezentują dane w formie graficznej przed modelowaniem i po nim. Dostępne formy graficzne to wykresy, histogramy, węzły sieciowe i wykresy ewaluacyjne.
- Węzły z grupy **Modelowanie** korzystają z algorytmów modelowania dostępnych w programie SPSS Modeler, takich jak sieci neuronowe, drzewa decyzyjne, algorytmy grupowania i określanie sekwencji danych.
- Węzły z grupy **Modelowanie w bazie** korzystają z algorytmów modelowania dostępnych w bazach danych Microsoft SQL Server, IBM Db2, Oracle i Netezza.
- Węzły z grupy **Wynik** generują różne wyniki w postaci danych, wykresów i wyników modeli, które można przeglądać w programie SPSS Modeler.
- Węzły z grupy **Eksport** generują różne dane wyjściowe, które można przeglądać w aplikacjach zewnętrznych, takich jak IBM SPSS Data Collection lub Excel.
- Węzły **IBM SPSS Statistics** importują dane z programu IBM SPSS Statistics i eksportują do niego dane, a także umożliwiają uruchamianie procedur IBM SPSS Statistics.
- Węzły z palety **Python** umożliwiają wykonywanie algorytmów zapisanych w języku Python.
- Węzły z grupy **Spark** umożliwiają wykonywanie algorytmów środowiska Spark.

W miarę nabywania doświadczenia w pracy z programem SPSS Modeler, można dostosowywać zawartość palet do indywidualnych potrzeb.

Po lewej stronie palety węzłów można filtrować wyświetlane węzły, wybierając kategorię: Nadzorowane, Związek lub Segmentacja.

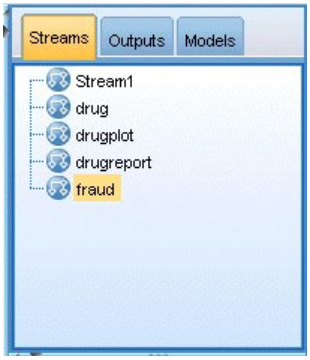
Pod paletą węzłów znajduje się panel raportów, na którym wyświetlane są informacje o postępach w realizacji różnych operacji, np. wczytywania danych do strumienia. Poniżej palety węzłów znajduje się także panel statusu, na którym wyświetlane są informacje o tym, co w danej chwili robi aplikacja, a także komunikaty informujące o konieczności wprowadzenia danych przez użytkownika.

Uwaga: W przypadku korzystania z MacBooka z aktywowanym ustawieniem **Kliknięcie mocne i odpowiedzi haptyczne** wbudowanego gładzika, przeciąganie i upuszczanie węzłów z palety węzłów do kanwy strumienia może spowodować zduplikowanie węzłów dodawanych do kanwy. Aby uniknąć tego problemu, zalecamy wyłączenie preferencji systemu gładzika **Kliknięcie mocne i odpowiedzi haptyczne**.

Panele zarządzania w programie IBM SPSS Modeler

W prawym górnym narożniku okna znajduje się panel zarządzania. Ten obszar składa się z trzech kart służących do zarządzania strumieniami, wynikami i modelami.

Za pomocą karty Strumienie można otwierać, zapisywać i usuwać strumienie otwarte w sesji oraz zmieniać ich nazwy.



Rysunek 5. Karta Strumienie

Rysunek 6. Karta Wyniki

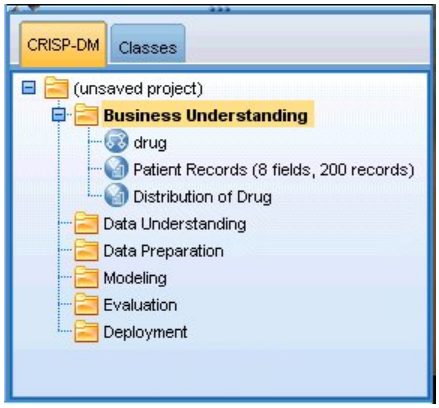
Na karcie Wyniki znajduje się wiele plików, takich jak wykresy i tabele, utworzonych w ramach obsługi strumieni w oprogramowaniu IBM SPSS Modeler. Użytkownik może wyświetlać, zapisywać i zamykać tabele, wykresy oraz raporty znajdujące się na tej karcie, jak również zmieniać ich nazwy.

Rysunek 7. Karta Modele zawierająca modele użytkowe

Karta Modele to najbardziej funkcjonalna z kart zarządzania. Na tej karcie znajdują się wszystkie **modele użytkowe**, które zawierają modele wygenerowane w bieżącej sesji oprogramowania IBM SPSS Modeler. Modele te można przeglądać bezpośrednio z poziomu karty Modele, można je również dodawać do strumienia w obszarze roboczym.

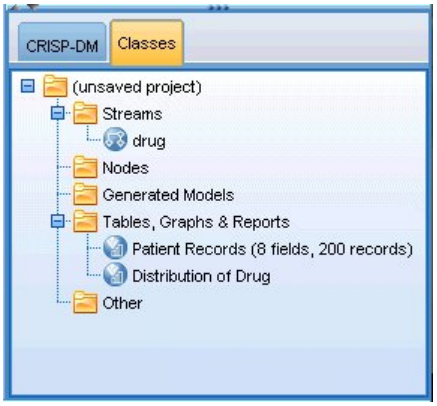
Projekty w programie IBM SPSS Modeler

W prawym dolnym narożniku okna znajduje się panel projektów służący do tworzenia **projektów** eksploracji danych (grup plików powiązanych z zadaniem eksploracji danych) oraz zarządzania nimi. Projekty utworzone w oprogramowaniu IBM SPSS Modeler można wyświetlać na dwa sposoby: w widoku Klas oraz CRISP-DM.



Rysunek 8. widok CRISP-DM

Karta CRISP-DM pozwala na organizowanie projektów zgodnie z formatem Cross-Industry Standard Process for Data Mining, będącym sprawdzoną, branżową metodologią działań. Zarówno w przypadku doświadczonych, jak i młodych stażem eksploratorów danych, narzędzie CRISP-DM ułatwia przeprowadzanie operacji organizacyjnych i usprawnia uzyskiwanie określonych wyników pracy.



Rysunek 9. widok Klasy

Karta Klasy umożliwia organizowanie pracy wg kategorii w oprogramowaniu IBM SPSS Modeler, czyli wg typów tworzonych obiektów. Ten widok jest przydatny podczas tworzenia zestawienia danych, strumieni i modeli.

Pasek narzędzi programu IBM SPSS Modeler

W górnej części okna IBM SPSS Modeler znajduje się pasek narzędzi umożliwiający wykonanie wielu przydatnych funkcji. Poniżej przedstawiono przyciski widoczne na pasku narzędzi wraz z opisami ich działania.

	Utwórz nowy strumień		Otwórz strumień
	Zapisz strumień		Wydrukuj bieżący strumień
	Wytnij i przenieś do schowka		Kopiuj do schowka
	Wklej zaznaczenie		Cofnij ostatnią czynność
	Powtórz		Wyszukaj węzły
	Edytuj właściwości strumienia		Wyświetl podgląd tworzenia kodu SQL
	Wykonaj bieżący strumień		Wykonaj wybrany fragment strumienia
	Zatrzymaj strumień (funkcja dostępna wyłącznie, gdy strumień jest wykonywany)		Dodaj Superwęzeł

 Powiększ (tylko Superwęzły)

 Pomniejsz (tylko Superwęzły)

 Brak znacznika w strumieniu

 Wstaw komentarz

 Ukryj znacznik strumienia (o ile istnieje)

 Pokaż ukryte znaczniki strumienia

 Otwórz strumień w oprogramowaniu IBM SPSS Modeler Advantage

Znaczniki strumienia składają się z komentarzy, łączy modelu i wskaźników gałęzi oceniania.

Łącza modelu opisano w publikacji *IBM SPSS - węzły modelowania*.

Dostosowywanie paska narzędzi

Użytkownik może zmienić wiele aspektów paska narzędzi:

- Wyświetlanie
- Dostępność podpowiedzi ikon
- Wielkość ikon

Włączanie i wyłączanie paska narzędzi:

1. W menu głównym kliknij opcje:

   Widok > **Pasek narzędzi** > **Wyświetlanie**

Zmiana ustawień rozmiarów ikon lub podpowiedzi:

1. W menu głównym kliknij opcje:

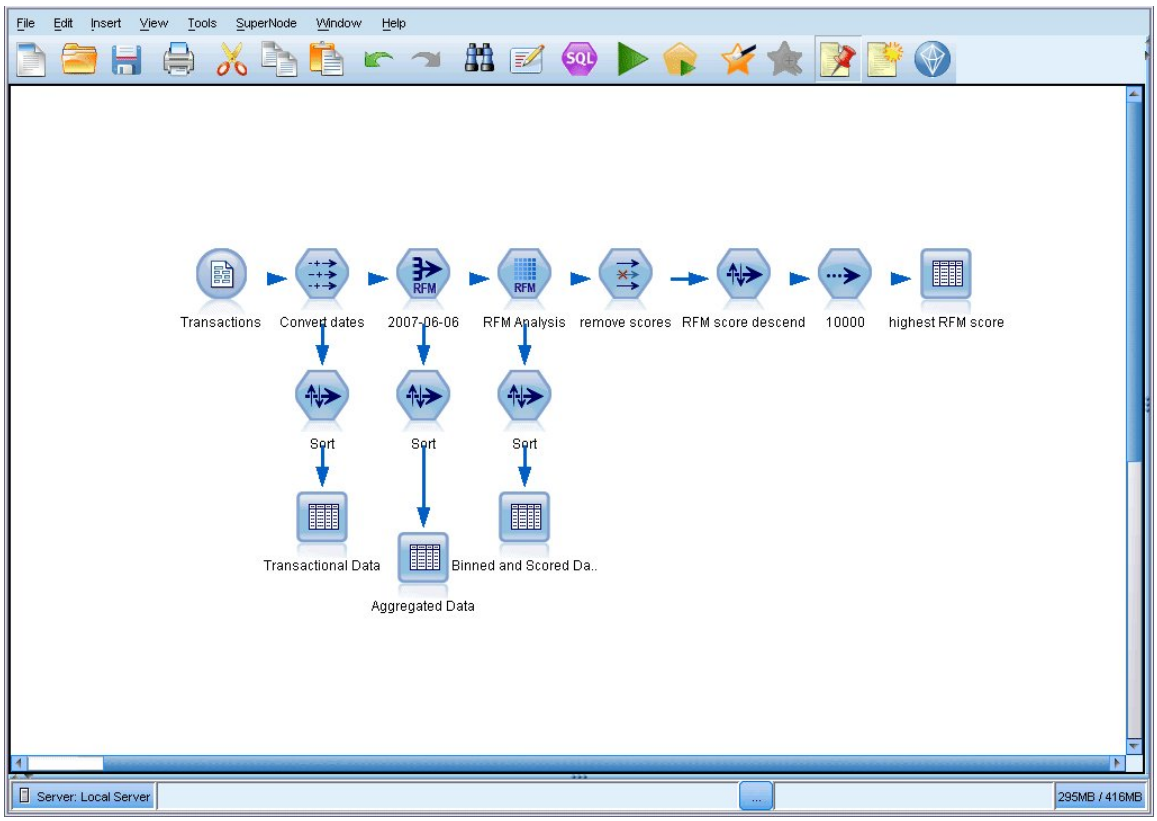
   Widok > **Paski narzędzi** > **Dostosuj**

Kliknij pozycję **Pokaż podpowiedzi** lub **Duże przyciski**.

Dostosowywanie okna programu IBM SPSS Modeler

Korzystając z linii podziału między różnymi fragmentami interfejsu programu SPSS Modeler, można zmieniać rozmiar lub zamykać narzędzia, aby dopasować środowisko pracy do swoich preferencji. Na przykład w przypadku pracy z dużym strumieniem można użyć małych strzałek znajdujących się na każdej z linii podziału w celu zamknięcia palety węzłów, panelu zarządzania oraz panelu projektu. Powoduje to maksymalizację obszaru roboczego strumienia, udostępniając wystarczający obszar roboczy dla jednego dużego lub wielu strumieni.

W menu Widok można też kliknąć pozycję **Paleta węzłów**, **Zarządzanie** lub **Projekt**, aby włączyć lub wyłączyć wyświetlanie tych elementów.



Rysunek 10. Zmaksymalizowany obszar roboczy strumienia

Jako alternatywy wobec zamknięcia palety węzłów oraz paneli zarządzania i projektu można użyć obszaru roboczego strumienia jako strony z możliwością przewijania; w tym celu należy użyć pasków przewijania w pionie i w poziomie znajdujących się z boku i u dołu okna SPSS Modeler.

Można także sterować wyświetlaniem znaczników ekranowych, składających się z komentarzy do strumieni, łączy modelu i oznaczeń gałęzi oceniania. W celu włączenia lub wyłączenia wyświetlania kliknij opcję:

Widok > **Adnotacja strumienia**

Zmiana rozmiaru ikony strumienia

Użytkownik może zmienić rozmiar ikon strumienia, korzystając z poniższych metod:

- za pomocą ustawienia właściwości strumienia,
- za pomocą menu podręcznego w strumieniu,
- korzystając z klawiatury.

Cały widok strumienia można przeskalować do dowolnego rozmiaru w zakresie od 8% do 200% w odniesieniu do standardowego rozmiaru ikony.

Skalowanie całego strumienia (właściwości strumienia)

1. Z menu głównego wybierz pozycje

   Narzędzia > **Właściwości strumienia** > **Opcje** > **Układ.**
2. Z menu Rozmiar ikony wybierz żądany rozmiar.
3. Kliknij przycisk **Zastosuj**, aby wyświetlić wynik operacji.
4. Kliknij przycisk **OK**, aby zapisać zmiany.

Skalowanie całego strumienia (menu)

1. Prawym przyciskiem myszy kliknij tło strumienia w obszarze roboczym.
2. Wybierz pozycję **Rozmiar ikony** i wybierz żądany rozmiar.

Skalowanie całego strumienia (klawiatura)

1. Na klawiaturze głównej naciśnij klawisze Ctrl + [-], aby zredukować skalę powiększenia do kolejnej mniejszej wartości.
2. Na klawiaturze głównej naciśnij klawisze Ctrl + Shift + [+], aby zwiększyć skalę powiększenia do kolejnej większej wartości.

   Należy zwrócić uwagę, że ta metoda powiększania może nie działać z niektórymi systemami operacyjnymi i klawiaturami.

Ta funkcja jest szczególnie przydatna w celu wyświetlenia ogólnego widoku złożonego strumienia. Dzięki niej można również zmniejszyć liczbę stron, na których zostanie wydrukowany strumień.

Korzystanie z myszy w oprogramowaniu IBM SPSS Modeler

Najczęstsze zastosowania myszy w oprogramowaniu IBM SPSS Modeler:

- **Kliknięcie.** Za pomocą lewego lub prawego przycisku można wybierać opcje z menu, otwierać menu podręczne i obsługiwać elementy sterujące oraz opcje. Kliknięcie i przytrzymanie przycisku pozwala na przenoszenie i przeciąganie węzłów.
- **Kliknięcie dwukrotne.** Kliknięcie dwukrotne lewym przyciskiem myszy powoduje wstawianie węzłów w obszarze roboczym strumienia i pozwala na edytowanie istniejących węzłów.
- **Kliknięcie środkowym przyciskiem.** Kliknięcie środkowym przyciskiem myszy i przeciągnięcie kursora powoduje połączenie węzłów w obszarze roboczym strumienia. Węzeł można odłączyć, klikając go dwukrotnie środkowym przyciskiem myszy. Jeśli użytkownik nie dysponuje myszą trzyprzyciskową, to trzeci przycisk można emulować, przytrzymując klawisz Alt podczas klikania lub przeciągania.

Używanie skrótów klawiaturowych

Wiele operacji programistycznych w środowisku wizualnym IBM SPSS Modeler można wykonać za pomocą przypisanych do nich skrótów klawiaturowych. Przykładowo: węzeł można usunąć, klikając go i naciskając klawisz Delete. Podobnie strumień można szybko zapisać, przytrzymując klawisz Ctrl i naciskając klawisz S. Komendy sterujące są uruchamiane za pomocą kombinacji klawisza Ctrl i innych klawiszy, np. Ctrl+S.

Używanych jest również wiele standardowych skrótów klawiaturowych systemu Windows, takich jak Ctrl+X (wycinanie). Są one obsługiwane w oprogramowaniu IBM SPSS Modeler wraz ze skrótami dostępnymi tylko w tym programie.

Uwaga: Niekiedy stare skróty klawiaturowe IBM SPSS Modeler powodują konflikt ze standardowymi skrótami klawiaturowymi systemu Windows. Stare skróty są uruchamiane w oprogramowaniu z użyciem klawisza Alt. Przykładowo: skrót Ctrl+Alt+C służy do włączania i wyłączania pamięci podręcznej.

Tabela 1. Obsługiwane skróty

Klawisz skrótu	**Funkcja**
Ctrl+A	Zaznacz wszystko
Ctrl+X	Wytnij
Ctrl+N	Nowy strumień
Ctrl+O	Otwórz strumień
Ctrl+P	Drukuj

Tabela 1. Obsługiwane skróty (kontynuacja)	
Klawisz skrótu	**Funkcja**
Ctrl+C	Kopiuj
Ctrl+V	Wklej
Ctrl+Z	Cofnij
Ctrl+Q	Zaznacz wszystkie węzły znajdujące się poniżej danego węzła
Ctrl+W	Usuń zaznaczenie wszystkich poniższych węzłów (przełączanie za pomocą klawiszy Ctrl+Q)
Ctrl+E	Wykonaj od wybranego węzła
Ctrl+S	Zapisz bieżący strumień
Alt+klawisze strzałek	Przenieś węzeł wybrany w obszarze roboczym strumienia w kierunku wskazanym za pomocą strzałki
Shift+F10	Otwórz menu podręczne wybranego węzła

Tabela 2. Obsługiwane skróty - stare klawisze skrótów	
Klawisz skrótu	**Funkcja**
Ctrl+Alt+D	Duplikuj węzeł
Ctrl+Alt+L	Załaduj węzeł
Ctrl+Alt+R	Zmień nazwę węzła
Ctrl+Alt+U	Utwórz węzeł danych użytkownika
Ctrl+Alt+C	Wł./wył. pamięć podręczną
Ctrl+Alt+F	Opróżnij pamięć podręczną
Ctrl+Alt+X	Rozwiń Superwęzeł
Ctrl+Alt+Z	Powiększ/pomniejsz
Usuń	Usuń węzeł lub połączenie

Drukowanie

W programie IBM SPSS Modeler możliwe jest wydrukowanie następujących obiektów:

- Diagramy strumienia
- Wykresy
- Tabele
- Raporty (z węzła Raport oraz z raportów projektu)
- Skrypty (z Właściwości strumienia, z okien dialogowych Skrypt samodzielny lub Skrypt superwęzła)
- Modele (przeglądarki modeli, karty okien dialogowych z aktualnym fokusem, widoku drzewa)
- Adnotacje (za pośrednictwem karty Adnotacje dla danych wynikowych)

Aby wydrukować projekt:

- Aby wydrukować bez podglądu, kliknij przycisk Drukuj na pasku narzędzi.
- Aby skonfigurować stronę przed przystąpieniem do drukowania, wybierz opcję **Ustawienia strony** z menu Plik.

- Aby przed drukowaniem wyświetlić podgląd, wybierz pozycję **Podgląd wydruku** z menu Plik.
- Aby wyświetlić standardowe okno dialogowe z opcjami wyboru drukarek oraz aby określić opcje wyglądu, wybierz pozycję **Drukuj** z menu Plik.

Automatyzacja procesów programu IBM SPSS Modeler

Ponieważ zaawansowana eksploracja danych może być procesem złożonym i niekiedy dość długotrwałym, program IBM SPSS Modeler oferuje różne możliwości programowania i automatyzacji.

- Język CLEM (ang. **Control Language for Expression Manipulation**) to język do analizowania i manipulowania danymi przechodzącymi przez strumienie IBM SPSS Modeler. Narzędzia do eksploracji danych wykorzystują język CLEM w szerokim zakresie w operacjach strumieniowych, do wykonywania zadań zarówno tak prostych, jak wyliczanie zysku na podstawie kosztów i przychodów, jak i tak złożonych, jak transformowanie treści blogów w zestaw zmiennych i rekordów zawierający użyteczne informacje.
- **Skrypty** są potężnym narzędziem automatyzacji procesów interfejsu użytkownika. Skrypty umożliwiają wykonywanie tego samego rodzaju czynności, jakie wykonywane są przez użytkownika za pomocą myszy lub klawiatury. Można także określić dane wynikowe i manipulować wygenerowanymi modelami.

Rozdział 4. Omówienie eksploracji danych

Eksploracja danych — przegląd

Eksploracja danych — poprzez zastosowanie różnych technik — identyfikuje modele użytkowe informacji w korpusach danych. Eksploracja danych wyodrębnia informacje, by były praktycznie użyteczne w takich dziedzinach, jak wsparcie procesów decyzyjnych, predykcja, prognozy i estymacje. Często zdarza się, że dane w formie nieprzetworzonej są bardzo obszerne, ale mają niską wartość i słabo nadają się do bezpośredniego wykorzystania. Realną wartość mają natomiast informacje, które są w tych danych ukryte.

Sukces eksploracji danych jest pochodną wiedzy użytkownika (lub eksperta pracującego dla użytkownika) na temat źródłowych danych w połączeniu z zaawansowanymi technikami aktywnej analizy, w których komputer wykrywa relacje i predyktory ukryte w danych. W procesie eksploracji danych na podstawie danych historycznych tworzone są modele, które następnie wykorzystuje się do predykcji, wykrywania wzorców i innych zastosowań. Technika budowania tych modeli nazywana jest **uczeniem maszynowym** (lub uczeniem się maszyn) bądź **modelowaniem**.

Techniki modelowania

IBM SPSS Modeler oferuje szereg technik uczenia maszynowego, które można ogólnie pogrupować według typów problemów, jakie w zamierzeniu mają rozwiązywać.

- Do metod modelowania predykcyjnego zalicza się drzewa decyzyjne, sieci neuronowe i modele statystyczne.
- Modele grupowania koncentrują się na identyfikacji grup o podobnych rekordach i oznaczania rekordów etykietami zgodnie z grupą, do której należą. Do metod grupowania zaliczamy metodę Kohonen, *k*-średnich i Dwustopniowa.
- Reguły asocjacyjne kojarzą konkretny wniosek (np. decyzję o zakupie konkretnego produktu) ze zbiorem warunków (zakupem kilku innych produktów).
- Modele monitorujące umożliwiają monitorowanie danych w poszukiwaniu zmiennych i rekordów, które z największym prawdopodobieństwem okażą się interesujące podczas modelowania, oraz identyfikację wartości odstających, które nie pasują do znanych wzorów. Do dostępnych metod należy wybór predyktorów i wykrywanie anomalii.

Manipulowanie danymi i wykrywanie danych

IBM SPSS Modeler zawiera także szereg mechanizmów, które umożliwiają użytkownikowi zastosowanie posiadanej wiedzy specjalistycznej wobec danych:

- **Manipulowanie danymi.** Możliwe jest budowanie nowych elementów danych na podstawie elementów istniejących i rozbijanie danych na praktycznie użyteczne podzbiory. Można scalać i filtrować dane pochodzące z wielu różnych źródeł.
- **Przeglądanie i wizualizacja.** Oprogramowanie prezentuje aspekty danych przy użyciu węzła Audyt danych, umożliwiając wykonanie początkowego audytu, w tym uzyskanie wykresów i statystyk. Zaawansowane techniki wizualizacji obejmują interaktywne postacie graficzne, które można eksportować w celu włączenia do raportów z projektu.
- **Statystyki.** Techniki te umożliwiają też potwierdzenie podejrzeń co do istnienia relacji między zmiennymi w danych. W programie IBM SPSS Modeler można też używać statystyk z programu IBM SPSS Statistics.
- **Testowanie hipotez.** Pozwala tworzyć modele określające sposób zachowywania się danych i sprawdza poprawność tych modeli.

Zwykle narzędzi tych używa się do rozpoznania obiecującego zbioru atrybutów w danych. Atrybuty te można następnie wykorzystać w procesie modelowania, w którym podjęta zostanie próba rozpoznania reguł i relacji ukrytych w danych.

Typowe zastosowania

Oto niektóre typowe zastosowania technik eksploracji danych:

Mailing bezpośredni. Eksploracja danych służy do ustalenia, w których grupach demograficznych wskaźnik odpowiedzi będzie najwyższy. Informacje takie można wykorzystać do optymalizacja wskaźnika odpowiedzi na przyszłe kampanie mailingowe.

Ocena kredytowa. Podejmowanie decyzji kredytowych na podstawie historii kredytowej osób.

Kadry. Analiza praktyk stosowanych dotychczas przy zatrudnianiu i stworzenie reguł decyzyjnych, które usprawnią proces rekrutacyjny.

Naukowe badania medyczne. Tworzenie reguł decyzyjnych, które będą podpowiadały odpowiednie procedury na podstawie dowodów medycznych.

Analiza rynku. Określenie, które zmienne, takie jak położenie geograficzne, cena i charakterystyka klienta, są powiązane ze sprzedażą.

Kontrola jakości. Analiza danych z procesu wytwarzania produktów i ustalenie, od których zmiennych zależą wady produktów.

Badania nad strategiami. Wykorzystanie danych ankietowych do formułowania strategii poprzez zastosowanie reguł decyzyjnych w celu wybrania najważniejszych zmiennych.

Opieka medyczna. Ankiety i dane kliniczne można wykorzystać łącznie do ustalenia, które zmienne mają wpływ na stan zdrowia.

Pojęcia

Terminy **atrybut**, **zmienna** i **pole** oznaczają jeden element danych wspólny dla wszystkich rozważanych obserwacji. Zbiór wartości atrybutów dotyczących konkretnej obserwacji nazywamy **rekordem**, **przykładem** lub **obserwacją**.

Uzyskiwanie dostępu do danych

Eksploracja danych prawdopodobnie nie okaże się owocna, o ile dane, które mają zostać użyte, nie będą spełniać określonych kryteriów. W poniższych sekcjach przedstawiono niektóre z aspektów dotyczących danych i ich zastosowań, które należy mieć na uwadze.

Upewnianie się, że dane są dostępne

Może się to wydawać oczywiste, lecz należy pamiętać, że choć dane mogą być udostępnione, mogą mieć one format, w którym nie da się ich w łatwy sposób użyć. Program IBM SPSS Modeler umożliwia importowanie danych z baz danych (za pośrednictwem ODBC) lub z plików. Dane mogą być jednak zapisane na komputerze w innej postaci, w której nie jest możliwy do nich bezpośredni dostęp. Będzie wówczas konieczne ich pobranie lub zapisanie w odpowiedniej postaci, zanim będą one mogły być używane. Możliwe jest również, że dane są rozproszone między różne bazy danych i źródła i będzie konieczne ich zgromadzenie w jednym miejscu. Dane mogą także być niedostępne online. Jeśli mają one wyłącznie postać papierową, przed przystąpieniem do eksploracji danych może okazać się konieczne ich wprowadzenie w postaci cyfrowej.

Sprawdzanie, czy dane obejmują właściwe atrybuty

Celem eksploracji danych jest identyfikacja odpowiednich atrybutów, tak że wykonywanie tej operacji sprawdzającej może początkowo wydawać się dziwne. Jednak bardzo użyteczne jest przeanalizowanie, które dane są dostępne, i podjęcie próby zidentyfikowania odpowiednich czynników, które nie zostały dotąd zarejestrowane. Przy próbie przewidywania sprzedaży lodów użytkownik może dysponować dużą ilością informacji na temat punktów sprzedaży lub historią sprzedaży, lecz może nie dysponować danymi na temat pogody i temperatury, która z dużym prawdopodobieństwem będzie odgrywać znaczącą rolę. Brak atrybutów niekoniecznie oznacza, że eksploracja danych nie wygeneruje użytecznych wyników; może ona jednak ograniczyć dokładność otrzymanych predykcji.

Prostym sposobem oceny sytuacji jest przeprowadzenie na posiadanych danych wyczerpującego audytu. Przed podjęciem dalszych działań należy jednak rozważyć podłączenie węzła Audyt danych do źródła danych i uruchomienie go w celu wygenerowania pełnego raportu.

Uwaga na dane zaszumione

Dane często zawierają błędy, mogą też zawierać subiektywne, a zatem odmienne, opinie. Zjawiska te są zbiorczo zwane **szumami**. Niekiedy szumy w danych są normalne. Mogą również występować niewidoczne reguły, lecz mogą one nie być przestrzegane w 100% obserwacji.

Zwykle im więcej szumów znajduje się w danych, tym trudniej jest uzyskać dokładne wyniki. Metody uczenia maszynowego w programie IBM SPSS Modeler są jednak zdolne do obsługi danych z szumami i okazują się skuteczne w przypadku zbiorów danych zawierających niemal 50% szumów.

Upewnianie się, że dane są wystarczające

Podczas eksploracji danych to nie określenie wielkości zbioru danych okazuje się najistotniejsze. Reprezentatywność zbioru danych jest znacznie istotniejsza, podobnie jak pokrycie możliwych danych wynikowych i kombinacji zmiennych.

Zwykle im więcej atrybutów jest rozważanych, tym więcej rekordów będzie potrzebnych do uzyskania reprezentatywnego pokrycia.

Jeśli dane są reprezentatywne i istnieją ogólne, podstawowe reguły, może się również okazać, że próba obejmująca kilka tysięcy (a nawet kilkaset) rekordów pozwoli uzyskać równie dobre wyniki, co próba zawierająca milion rekordów, zaś wynik będzie dostępny znacznie szybciej.

Poszukiwanie specjalistów znających dane

W wielu przypadkach użytkownik będzie pracował na własnych danych, przez co będzie zaznajomiony z ich treścią i znaczeniem. Jednak w przypadku pracy z danymi innego działu organizacji lub z danymi klienta wysoce pożądany jest dostęp do specjalistów zaznajomionych z tymi danymi. Mogą oni udzielić wskazówek co do identyfikacji właściwych atrybutów oraz pomóc w interpretacji wyników eksploracji danych, w odróżnieniu do prawdziwych modeli użytkowych zawierających od modeli pozornie istotnych czy też artefaktów spowodowanych anomaliami w zestawach danych.

Strategia eksploracji danych

Podobnie jak większość przedsięwzięć biznesowych eksploracja danych udaje się najlepiej, gdy prowadzi się ją w sposób planowy i systematyczny. Nawet najnowocześniejsze narzędzia do eksploracji danych, takie jak IBM SPSS Modeler, nie będą wiele warte bez nadzoru kompetentnego analityka biznesowego, który będzie dbał o utrzymanie projektu na właściwym torze. Planowanie należy zacząć od odpowiedzi na następujące pytania:

- Jaki istotny problem chcemy rozwiązać?
- Jakie źródła danych są dostępne i jakie części danych są istotne dla rozwiązania naszego problemu?
- W jaki sposób musimy wstępnie przetworzyć i oczyścić dane zanim rozpoczniemy eksploracje?
- Jakich technik eksploracji danych użyjemy?
- W jaki sposób będziemy oceniać wyniki eksploracji danych?
- W jaki sposób można najlepiej wykorzystać informacje uzyskane w wyniku eksploracji danych?

Typowy proces eksploracji danych bardzo szybko się komplikuje. Musimy pamiętać o wielu rzeczach naraz — złożonych problemach biznesowych, wielu źródłach danych, różnicach w jakości danych między tymi źródłami, technikach eksploracji danych, różnych metodach pomiaru jakości uzyskanych wyników itd.

Aby się nie pogubić, warto mieć wyraźnie zdefiniowany model procesu eksploracji danych. Model procesu pomoże w udzieleniu odpowiedzi na pytania wymienione wcześniej w tej sekcji oraz zapewni uwzględnienie wszystkich ważnych aspektów. Model procesu jest swego rodzaju mapą drogową eksploracji danych, dzięki której nie zgubimy się w procesie mimo wysokiej złożoności danych.

Modelem procesu rekomendowanym do zastosowania z programem SPSS Modeler jest międzybranżowy standardowy model procesu eksploracji danych — Cross-Industry Standard Process for Data Mining (CRISP-DM). Zgodnie ze swą nazwą, model ten został opracowany jako model ogólny, który można stosować w różnych branżach do rozwiązywania różnych procesów biznesowych.

Model procesu CRISP-DM

Ogólny model procesu CRISP-DM obejmuje sześć faz uwzględniających główne problemy eksploracji danych. Fazy te są wzajemnie połączone w proces cykliczny, który w zamierzeniu ma umożliwić włączenie eksploracji danych do szerzej rozumianych procesów biznesowych.

Oto sześć faz opisywanego modelu:

- **Zapoznanie z zadaniem.** Jest to prawdopodobnie najważniejsza faza eksploracji danych. Zapoznanie z zadaniem polega na określeniu celów biznesowych, ocenie sytuacji, ustaleniu celów eksploracji danych i opracowaniu planu projektu.
- **Zrozumienie danych.** Dane są „surowcem" eksploracji danych. W tej fazie staramy się zrozumieć, jakie zasoby danych mamy do dyspozycji, i jaka jest charakterystyka tych zasobów. Faza ta obejmuje zbieranie początkowych danych, opisywanie danych, zdobywanie orientacji w danych i weryfikowanie jakości danych. Węzeł Audyt danych dostępny w palecie węzłów Wynik jest niezastąpionym narzędziem, które pomaga zrozumieć dane.
- **Przygotowanie danych.** Po skatalogowaniu dostępnych zasobów danych musimy przygotować je do eksploracji. Przygotowania obejmują wybór, czyszczenie, budowanie, integrowanie i formatowanie danych.
- **Modelowanie.** Jest to najbardziej efektowna część eksploracji danych, w której stosujemy wyrafinowane metody analityczne w celu wydobycia z danych wartościowych informacji. Faza ta obejmuje wybór technik modelowania, generowanie projektów testowych, budowanie oraz ocenę modeli.
- **Ewaluacja.** Po wybraniu modeli jesteśmy gotowi do oceny przydatności potencjalnych wyników eksploracji danych w kontekście celów biznesowych. Do elementów tej fazy należy ewaluacja wyników, przegląd procesu eksploracji danych i określenie następnych kroków.
- **Wdrożenie.** Poprzednie pracochłonne fazy miały charakter inwestycji — a teraz pora zebrać owoce naszych wysiłków. W tej fazie skupiamy się na integracji zdobytej wiedzy z codziennymi procesami biznesowymi w celu rozwiązania pierwotnego problemu biznesowego. Faza ta obejmuje wdrożenie, monitorowanie i aktualizowanie planu, wygenerowanie końcowego raportu i recenzję projektu.

W opisywanym modelu procesu należy zwrócić uwagę na kilka ważnych kwestii. Po pierwsze, choć z reguły kolejne kroki procesu wykonywane są w opisanej powyżej kolejności, istnieje szereg punktów, w których fazy wzajemnie oddziałują na siebie w sposób nieliniowy. Na przykład przygotowanie danych zwykle poprzedza modelowanie. Jednak decyzje podjęte i informacje zebrane w fazie modelowania mogą skłonić nas do innego niż dotąd spojrzenia na fazę przygotowania danych, powrotu do niej i powtórzenia modelowania. Wyniki obu tych faz wpływają na siebie wzajemnie i fazy powtarzają się, aż do uzyskania oczekiwanego rezultatu. Podobnie, faza ewaluacji może skłonić nas do weryfikacji pierwotnych opinii na temat uwarunkowań biznesowych i doprowadzić do wniosku, że próbujemy odpowiedzieć na źle postawione pytanie. Na tym etapie możemy na nowo przemyśleć uwarunkowania biznesowe i powtórzyć cały proces — tym razem już z lepiej sformułowanym celem.

Druga ważna kwestia to iteracyjny charakter eksploracji danych. Rzadko, jeśli w ogóle, projekt eksploracji jest zdarzeniem jednorazowym: planujemy, realizujemy, odbieramy wyniki — i gotowe. W rzeczywistości eksploracja danych ukierunkowana na rzeczywiste oczekiwania klientów jest procesem ciągłym. Wiedza uzyskana w jednym cyklu eksploracji danych niemal zawsze doprowadzi nas do nowych pytań, nowych problemów i nowych okazji, by rozpoznać potrzeby klientów i je zaspokoić. Odpowiedzią na te nowe pytania, problemy i szanse jest zwykle ponowna eksploracja danych. Proces eksploracji i odkrywania nowych szans powinien stać się fundamentem nowego spojrzenia na całą działalność i strategię biznesową.

Niniejsze wprowadzenie jest jedynie krótkim omówieniem modelu procesu CRISP-DM. Szczegółowe informacje na jego temat można znaleźć w następujących źródłach:

- *Podręczniku CRISP-DM*, który znajduje się, podobnie jak pozostała dokumentacja, w folderze *\Documentation* na dysku instalacyjnym;
- w systemie Pomocy modelu CRISP-DM, który jest dostępny w menu Start lub po wybraniu pozycji **Metodologia CRISP-DM** z menu Pomoc programu IBM SPSS Modeler.

Typy modeli

Oprogramowanie IBM SPSS Modeler umożliwia korzystanie z wielu metod modelowania opartych na sztucznej inteligencji, uczeniu maszynowym i statystykach. Metody dostępne na palecie Modelowanie pozwalają na ekstrahowanie nowych informacji z danych i tworzenie modeli predykcyjnych. Każda metoda ma określone mocne strony i jest dostosowana do rozwiązywania określonych problemów.

Publikacja *IBM SPSS Modeler — podręcznik zastosowań* zawiera przykłady zastosowania wielu z tych metod wraz z ogólnym wprowadzeniem do procesu modelowania. Ten podręcznik jest dostępny jako samouczek online oraz jako plik w formacie PDF. Więcej informacji można znaleźć w temacie "Przykłady aplikacji" na stronie 4.

Metody modelowania dzielą się na niniejsze kategorie:

- Nadzorowane
- Związek
- Segmentacja

Modele nadzorowane

Modele nadzorowane korzystają z wartości jednej lub większej liczby zmiennych **wejściowych** do przewidywania wartości jednej lub większej liczby zmiennych wyjściowych lub **przewidywanych**. Niektóre z przykładów takich technik to: drzewa decyzyjne (drzewo C&R, QUEST, CHAID i algorytmy C5.0), regresja (liniowa, logistyczna, uogólniona liniowa oraz algorytmy regresji Coksa), sieci neuronowe, algorytmy SVM oraz sieci Bayesowskie.

Modele nadzorowane pomagają organizacjom przewidywać znany wynik, np. fakt albo rezygnacji z zakupu bądź też dopasowania transakcji do znanego wzorca oszustwa. Techniki modelowania obejmują także uczenie maszynowe, wywodzenie reguł, identyfikację podgrup, metody statystyczne i generowanie wielu modeli.

Węzły nadzorowane

Węzeł Auto Klasyfikacja tworzy i porównuje różne modele pod kątem wyników binarnych (tak lub nie, odejścia lub brak odejścia itd.), umożliwiając użytkownikowi wybór optymalnego podejścia do danej analizy. Obsługiwana jest pewna liczba algorytmów modelowania, co umożliwia wybór metod, które mają zostać użyte, konkretnych opcji dla każdej z nich oraz kryteriów porównywania wyników. Węzeł generuje zestaw modeli w oparciu o określone opcje i nadaje rangi najlepszym kandydatom wybranym według wskazanych kryteriów.

Węzeł Auto Predykcja estymuje i porównuje modele zwracające wyniki w formie ciągłego przedziału liczbowego, korzystając z szeregu różnych metod. Węzeł działa tak samo, jak węzeł Auto Klasyfikacja, umożliwiając użytkownikowi wybór używanych algorytmów oraz eksperymentowanie z wieloma kombinacjami opcji w pojedynczym przebiegu modelowania. Obsługiwane algorytmy obejmują sieci neuronowe, drzewo C&R, CHAID, regresję liniową, uogólnioną regresję liniową oraz algorytmy SVM. Modele można porównywać na podstawie korelacji, błędu względnego lub liczby używanych zmiennych.

Węzeł Klasyfikacja i regresja (C&R) generuje drzewo decyzyjne umożliwiające predykcję lub klasyfikację przyszłych obserwacji. W metodzie tej stosowany jest rekursywny podział rekordów na segmenty przez minimalizację zanieczyszczeń w każdym kroku, przy czym węzeł w drzewie jest uważany za „czysty”, jeśli 100% obserwacji w węźle przypada na konkretną kategorię zmiennej przewidywanej. Zmienne przewidywana i wejściowa mogą być zakresami liczbowymi lub jakościowymi (nominalnymi, porządkowymi lub flagami); wszystkie podziały są binarne (tylko dwie podgrupy).

Węzeł QUEST oferuje metodę klasyfikacji binarnej służącą do budowania drzew decyzyjnych, zaprojektowaną w celu redukcji czasu przetwarzania analiz dużych drzew decyzyjnych C&R, a jednocześnie w celu redukcji tendencji obecnej w metodach drzew klasyfikacji do preferowania danych wejściowych dopuszczających więcej podziałów. Zmienne wejściowe mogą być zakresami liczbowymi (ciągłymi), lecz zmienna przewidywana musi być jakościowa. Wszystkie podziały są binarne.

Węzeł CHAID generuje drzewa decyzyjne, korzystając ze statystyk chi-kwadrat w celu identyfikacji optymalnych podziałów. W odróżnieniu od węzłów drzewa C&R i węzłów QUEST, CHAID może generować drzewa niebinarne, co oznacza, że niektóre podziały mają więcej niż dwie gałęzie. Zmienne przewidywana i wejściowa mogą być zakresami liczbowymi (ciągłymi) lub jakościowymi. Wyczerpujący CHAID stanowi modyfikację CHAID umożliwiającą dokładniejsze badanie wszystkich możliwych podziałów, lecz obliczenia w jego przypadku zajmują więcej czasu.

Węzeł C5.0 tworzy drzewo decyzyjne lub zestaw reguł. Model działa w oparciu o podział próby na podstawie zmiennej oferującej maksimum korzyści z informacji na każdym z poziomów. Zmienna przewidywana musi być jakościowa. Dozwolonych jest wiele podziałów na więcej niż dwie podgrupy.

Węzeł Lista decyzyjna identyfikuje podgrupy lub segmenty wskazujące wyższe lub niższe prawdopodobieństwo danego wyniku binarnego względem całej populacji. Można na przykład wyszukać klientów, których prawdopodobieństwo odejścia jest niewielkie, lub którzy z dużym prawdopodobieństwem pozytywnie zareagują na kampanię. Istnieje możliwość zastosowania posiadanej wiedzy biznesowej w modelu przez dodanie własnych, niestandardowych segmentów i przejrzenie modeli alternatywnych jeden obok drugiego w celu porównania wyników. Modele Lista decyzyjna składają się z list reguł, w których każda reguła ma warunek i wynik. Reguły są stosowane w kolejności wprowadzania, a pierwsza reguła spełniona określa wynik.

Modele regresji liniowej przewidują docelową wartość ilościową na podstawie liniowych relacji między docelową wartością ilościową a jednym lub większą liczbą predyktorów.

Węzeł analizy PCA/czynnikowej udostępnia wydajne techniki redukcji danych pozwalające obniżyć stopień ich złożoności. Analiza głównych składowych (ang. Principal Components Analysis, PCA) znajduje kombinacje liniowe zmiennych wejściowych, które umożliwiają określenie wariancji w całym zestawie zmiennych, pod warunkiem że składowe są zlokalizowane ortogonalnie (prostopadle) do siebie. Analiza czynnikowa próbuje zidentyfikować współczynniki objaśniające wzory korelacji występujące w ramach zbiorów obserwowanych zmiennych. W przypadku obu podejść celem jest znalezienie niewielkiej liczby zmiennych wyliczanych w efektywny sposób, która podsumowuje informacje w oryginalnym zestawie zmiennych.

Węzeł Dobór predyktorów przegląda zmienne wejściowe do usunięcia w oparciu o zbiór kryteriów (takich jak procent braków danych); następnie nadaje rangę istotności pozostałych danych wejściowych względem określonej zmiennej przewidywanej. Na przykład, jeśli mamy zbiór danych z setkami potencjalnych danych wejściowych, to które z nich z dużym prawdopodobieństwem okażą się użyteczne w modelowaniu wyników leczenia pacjenta?

Analiza dyskryminacyjna opiera się na ściślejszych założeniach niż regresja logistyczna, lecz może stanowić wartościową alternatywę lub uzupełnienie analizy metodą regresji logistycznej w przypadku spełnienia tych założeń.

Regresja logistyczna to technika statystyczna umożliwiająca klasyfikację rekordów na podstawie wartości zmiennych wejściowych. Jest ona analogiczna do regresji liniowej, lecz bazuje na przewidywanej zmiennej jakościowej zamiast na przedziale liczbowym.

Procedura ogólnych modeli liniowych rozszerza ogólny model liniowy w taki sposób, że zmienna zależna jest liniowo powiązana z czynnikami i współzmiennymi za pośrednictwem określonej funkcji łączenia. Model pozwala ponadto, aby zmienna zależna nie miała rozkładu normalnego. Obejmuje ona funkcjonalność dużej liczby modeli statystycznych, m.in. regresji liniowej, regresji logistycznej, modeli logarytmiczno-liniowych dla danych o liczebności.

Uogólniony liniowy model mieszany (GLMM) stanowi wersję modelu liniowego rozszerzoną w taki sposób, że zmienna przewidywana może mieć rozkład inny niż normalny, jest liniowo powiązana z czynnikami i współzmiennymi za pośrednictwem określonej funkcji łączenia, a obserwacje mogą być skorelowane. Uogólnione liniowe modele mieszane obejmują szeroki wachlarz modeli, począwszy od prostych modeli regresji liniowej, aż po złożone wielopoziomowe modele dla danych z obserwacji długofalowych nieposiadających rozkładu normalnego.

Węzeł regresji Coksa umożliwia utworzenie modelu przeżycia dla danych określających czasy do wystąpienia zdarzeń i zawierających ocenzurowane rekordy. Model generuje funkcję przeżycia przewidującą prawdopodobieństwo, że zdarzenie będące przedmiotem zainteresowania wystąpiło w określonym czasie (*t*) dla danych wartości zmiennych wejściowych.

Węzeł Algorytm SVM umożliwia szybką klasyfikację danych do jednej lub dwu grup bez przeuczenia. Algorytm SVM działa prawidłowo dla szerokiego zbioru danych, na przykład takiego o bardzo dużej liczbie zmiennych wejściowych.

Węzeł Sieć Bayesowska umożliwia utworzenie modelu prawdopodobieństwa przez połączenie zaobserwowanych i zarejestrowanych dowodów z wiedzą rzeczywistą w celu ustanowienia prawdopodobieństwa występowania. Węzeł koncentruje się na sieciach Tree Augmented Naïve Bayes (TAN) i Markov Blanket, używanych głównie podczas klasyfikacji.

Węzeł Model odpowiedzi samonauczania (SLRM) umożliwia utworzenie modelu, w którym pojedyncza nowa obserwacja lub niewielka liczba nowych obserwacji może zostać użyta do ponownej oceny modelu bez konieczności ponownego uczenia modelu z wykorzystaniem wszystkich danych.

Węzeł Szereg czasowy umożliwia estymację modelu wygładzania wykładniczego, modelu autoregresyjnej zintegrowanej średniej ruchomej (ARIMA) jednej zmiennej oraz modelu ARIMA wielu zmiennych (lub funkcji przenoszenia) dla danych szeregów czasowych i generuje prognozy przyszłych wyników. Węzeł Szereg czasowy jest podobny do dotychczasowego węzła Szereg czasowy, który w wersji 18 programu SPSS Modeler ma status nieaktualnego. Nowszy węzeł Szeregi czasowe umożliwia wykorzystanie potencjału produktu IBM SPSS Analytic Server przy przetwarzaniu wielkich zbiorów danych i prezentuje wynikowy model w przeglądarce wyników dodanej w wersji 17 programu SPSS Modeler.

Węzeł *k*-najbliższego sąsiedztwa (KNN) wiąże nową obserwację z kategorią lub wartością *k* (gdzie *k* jest liczbą całkowitą) najbliższych obiektów w przestrzeni predyktora. Podobne obserwacje znajdują się blisko siebie, a niepodobne — daleko.

Węzeł predykcji przestrzenno-czasowej używa danych zawierających informacje o lokalizacji, zmiennych wejściowych predykcji (predyktorów), zmiennej czasu i zmiennej przewidywanej. W danych z każdą lokalizacją powiązany jest szereg wierszy, które odzwierciedlają wartości predyktorów w różnych punktach w czasie. Po przeanalizowaniu danych mogą być one używane do przewidywania wartości w dowolnej lokalizacji w danych kształtu używanych w analizie.

Modele asocjacyjne

Modele asocjacyjne znajdują wzorce w danych, w których jeden lub więcej obiektów (takich jak zdarzenia, zakupy czy atrybuty) jest powiązanych z jednym lub większą liczbą z pozostałych obiektów. Modele te tworzą zestawy reguł definiujące relacje. W tym miejscu zmienne w ramach danych pełnią rolę zarówno danych wejściowych, jak i docelowych. Związki te można znaleźć również ręcznie, lecz algorytmy reguł asocjacyjnych pozwalają wykonać te operacje znacznie szybciej i umożliwiają eksplorację bardziej złożonych wzorców. Modele Apriori i Carma stanowią przykłady użycia takich algorytmów. Jednym z kolejnych typów modeli asocjacyjnych jest model wykrywania kolejności, który znajduje wzorce sekwencyjne w danych ustrukturyzowanych względem czasu.

Modele asocjacyjne są najbardziej użyteczne w przypadku przewidywania wielokrotnych danych wynikowych — na przykład klienci, którzy kupili produkt X, kupili także produkty Y i Z. Modele asocjacyjne umożliwiają powiązanie konkretnego wniosku (np. decyzji o zakupie) z zestawem warunków. Przewagą algorytmów reguł asocjacyjnych wobec bardziej standardowych algorytmów drzew decyzyjnych (C5.0 i C&RT) jest fakt, że dozwolone są w nich związki między dowolnymi atrybutami. Algorytm drzewa decyzyjnego pozwala utworzyć reguły z tylko jednym wnioskiem, podczas gdy algorytmy powiązań próbują znaleźć wiele reguł, z których każda może mieć inny wniosek.

Węzły powiązań

Węzeł Apriori pozwala wyodrębnić zestaw reguł na podstawie danych, pobierając reguły o najwyższej możliwej zawartości informacji. Apriori oferuje pięć różnych metod wybierania reguł i korzysta ze złożonego schematu indeksowania do efektywnego przetwarzania dużych zbiorów danych. W przypadku dużych problemów czas uczenia Apriori jest zwykle krótszy. Brak jest arbitralnego limitu co do liczby reguł do utrzymania, możliwa jest obsługa reguł z maksymalnie 32 predykcjami. Apriori wymaga, aby wszystkie zmienne wejściowe i wyjściowe były zmiennymi jakościowymi, lecz oferuje wyższą wydajność z uwagi na optymalizację pod kątem tego typu danych.

Model CARMA pozwala wyodrębnić zestaw reguł na podstawie danych bez konieczności określania zmiennych wejściowych lub przewidywanych. Inaczej niż Apriori, węzeł CARMA oferuje ustawienia tworzenia dla obsługi reguł (pokrycie poprzedników i następników) zamiast pokrycia tylko poprzedników. Oznacza to, że wygenerowane reguły mogą być używane w szerszym spektrum zastosowań — na przykład w celu znalezienia listy produktów lub usług (poprzedników), z których wynikać będzie decyzja o promowaniu konkretnego produktu (następnika) w tegorocznym sezonie świątecznym.

Węzeł Sekwencje wykrywa reguły asocjacyjne w danych sekwencyjnych lub zorientowanych czasowo. Sekwencja to lista zbiorów elementów z tendencją do występowania w przewidywalnej kolejności. Na przykład klient dokonujący zakupu brzytwy i balsamu po goleniu przy następnej wizycie w sklepie może dokonać zakupu kremu po goleniu. Węzeł Sekwencje bazuje na algorytmie reguł asocjacyjnych CARMA, który korzysta z efektywnej metody dwu przejść do znajdowania sekwencji.

Węzeł Reguły asocjacyjne jest podobny do węzła Apriori; jednak inaczej niż w przypadku Apriori, węzeł Reguły asocjacyjne umożliwia przetwarzanie danych w postaci listy. Ponadto węzeł Reguły asocjacyjne może być używany wraz z IBM SPSS Analytic Server do przetwarzania dużych zbiorów danych i korzystania z szybszego przetwarzania równoległego.

Modele segmentacji

Modele segmentacji dzielą dane na segmenty lub grupy rekordów o podobnych wzorcach zmiennych wejściowych. Ponieważ modele segmentacji przetwarzają jedynie zmienne wejściowe, nie mają one żadnych informacji na temat zmiennych wyjściowych ani przewidywanych. Przykłady modeli segmentacji to sieci Kohonen, grupowanie K-średnich, grupowanie dwustopniowe i wykrywanie anomalii.

Modele segmentacji (zwane również „modelami grupowania”) są szczególnie przydatne w przypadkach, gdzie konkretny wynik jest nieznany (na przykład, przy identyfikacji nowych wzorców oszustw lub identyfikacji będących potencjalnie przedmiotem zainteresowania grup w bazie danych klientów). Modele skupień koncentrują się na identyfikacji grup o podobnych rekordach i oznaczaniu rekordów etykietami zgodnie z grupą, do której należą. Jest to realizowane mimo braku wstępnej wiedzy o grupach i ich charakterystykach, i pozwala odróżnić modele grupowania od innych technik modelowania, w których brak wstępnie zdefiniowanego wyniku czy zmiennej docelowej dla modelu objętego predykcją. W przypadku tych modeli nie ma poprawnych czy niepoprawnych odpowiedzi. Ich wartość określana jest przez zdolność do przechwytywania interesujących skupień w danych i oferowania użytecznych opisów tych skupień. Modele grupowania są często używane do tworzenia grup lub segmentów, które są następnie często używane jako dane wejściowe w kolejnych analizach (na przykład dzięki segmentacji potencjalnych klientów w jednorodne podgrupy).

Węzły segmentacji

Węzeł Autogrupowanie szacuje i porównuje modele skupień identyfikujące grupy rekordów o podobnej charakterystyce. Węzeł działa tak samo, jak pozostałe zautomatyzowane węzły modelowania, umożliwiając eksperymentowanie z wieloma kombinacjami opcji w pojedynczym przebiegu modelowania. Modele można porównywać, korzystając z miar bazowych, które pozwalają podejmować próby filtrowania i oceny przydatności modelu skupień oraz udostępniają miary bazujące na istotności poszczególnych zmiennych.

Węzeł K-średnie grupuje zbiór danych w osobne grupy (lub skupienia). Metoda ta definiuje stałą liczbę skupień, w sposób iteracyjny przypisuje rekordy do skupień i dopasowuje centra skupień do chwili, gdy dalsze pokrycie nie będzie miało wpływu na ulepszenie modelu. Zamiast prób predykcji danych wynikowych *k*-średnia korzysta z procesu znanego jako nienadzorowane uczenie w celu ujawnienia wzorców w zbiorze zmiennych wejściowych.

Węzeł Kohonen generuje typ sieci neuronowej, którą można wykorzystać do grupowania zbioru danych w osobne grupy. Po pełnym przeszkoleniu sieci rekordy podobne do siebie powinny znajdować się blisko siebie na mapie wyników, podczas gdy rekordy różne od siebie powinny znajdować się daleko od siebie. Na podstawie liczby obserwacji przechwyconych przez każdą jednostkę w modelu użytkowym można rozpoznać silne jednostki. Może to dać pojęcie o odpowiedniej liczbie skupień.

Węzeł Dwustopniowa korzysta z dwustopniowej metody grupowania. Pierwszy krok stanowi pojedynczy przebieg danych z myślą o kompresji surowych danych wejściowych w łatwy w zarządzaniu zestaw podgrup. Drugi krok korzysta z hierarchicznej metody grupowania w celu progresywnego scalania podgrup w coraz większe grupy. Metoda Dwustopniowa oferuje korzyści wynikające z automatycznego szacowania optymalnej liczby grup na potrzeby danych szkoleniowych. Pozwala ona skutecznie obsługiwać mieszane typy zmiennych i duże zbiory danych.

Węzeł Wykrywanie anomalii umożliwia identyfikację nietypowych obserwacji lub wartości odstających, które są niezgodne z wzorcami dla „normalnych” danych. Korzystając z tego węzła, można zidentyfikować wartości odstające nawet, jeśli nie pasują one do żadnego z wcześniej znanych wzorców oraz jeśli brak pewności co do charakteru poszukiwanych danych.

Modele eksploracji w bazie danych

Program IBM SPSS Modeler oferuje integrację narzędzi do eksploracji danych i modelowania, dostępnych w bazach danych, takich jak Oracle Data Miner i Microsoft Analysis Services. Można tworzyć, oceniać i składować modele w bazie danych — a wszystko to w ramach aplikacji IBM SPSS Modeler. Bardziej szczegółowe informacje można znaleźć w publikacji *IBM SPSS Modeler — podręcznik eksploracji w bazie danych*.

Modele IBM SPSS Statistics

Jeśli na komputerze zainstalowano kopię produktu IBM SPSS Statistics z licencją, wówczas w celu tworzenia i oceny modeli źródłowych można uzyskiwać dostęp do konkretnych procedur programu IBM SPSS Statistics i uruchamiać je z poziomu programu IBM SPSS Modeler.

Przykłady eksploracji danych

Najlepszym sposobem zapoznania się z operacjami dotyczącymi eksploracji danych jest wykonanie przykładowych ćwiczeń. Kilka ćwiczeń zawarto w publikacji *IBM SPSS Modeler Applications — podręcznik*, gdzie przedstawiono skrócone informacje wprowadzające związane z konkretnymi metodami i technikami modelowania. Więcej informacji można znaleźć w temacie “Przykłady aplikacji” na stronie 4.

Rozdział 5. Tworzenie strumieni

Tworzenie strumienia — przegląd

Podczas wykonywania operacji eksploracji danych w oprogramowaniu IBM SPSS Modeler dane przepływają przez szereg węzłów tworzących tzw. **strumień**. Ten szereg odzwierciedla operacje wykonywane na danych, a połączenia między węzłami wskazują kierunek przepływu danych. Zwykle strumień danych służy do wczytywania danych do oprogramowania IBM SPSS Modeler, przeprowadzania na nich szeregu manipulacji i wysyłania do położenia docelowego, np. do tabeli lub przeglądarki.

Przykładowo: użytkownik chce otworzyć źródło danych, dodać nową zmienną, wybrać rekordy w oparciu o wartości w nowej zmiennej, a następnie wyświetlić wyniki w tabeli. W tym przypadku strumień danych będzie się składał z czterech węzłów:

Węzła pliku zmiennego skonfigurowanego w celu odczytania danych ze źródła danych.

Węzła wyliczenia używanego do dodania nowej, obliczonej zmiennej do zbioru danych.

Węzła wyboru służącego do skonfigurowania kryteriów wyboru umożliwiających wykluczenie rekordów ze strumienia danych.

Węzła tabeli umożliwiającego wyświetlenie wyników manipulacji na ekranie.

Tworzenie strumieni danych

Unikalny interfejs SPSS Modeler pozwala na wizualne eksplorowanie danych z użyciem diagramów strumieni danych. Najbardziej podstawowy strumień danych można utworzyć w następujący sposób:

- Dodaj węzły do obszaru roboczego strumienia.
- Połącz węzły w celu utworzenia strumienia.
- Określ opcje strumienia lub węzła.
- Uruchom strumień.

W tej sekcji przedstawiono bardziej szczegółowe informacje na temat pracy z węzłami, które pozwolą na utworzenie złożonych strumieni. Omówiono również opcje i ustawienia węzłów oraz strumieni. Aby zapoznać się ze szczegółowymi informacjami na temat tworzenia strumienia z zastosowaniem danych dostarczonych wraz z oprogramowaniem SPSS Modeler (w folderze Demos w katalogu instalacyjnym programu), patrz rozdział “Przykłady aplikacji” na stronie 4.

Praca z węzłami

Węzły są używane w programie IBM SPSS Modeler, aby ułatwić eksplorację danych. Różne węzły w obszarze roboczym reprezentują różne obiekty i czynności. Paleta u dołu okna IBM SPSS Modeler zawiera wszystkie możliwe węzły użyte do tworzenia modelu.

Istnieje kilka typów węzłów. **Węzły źródłowe** służą do wprowadzania danych do strumienia i znajdują się na karcie Źródła w palecie węzłów. **Węzły procesowe** pozwalają na wykonywanie operacji w

poszczególnych rekordach danych i zmiennych; można je znaleźć na kartach Rekordy i Zmienne w palecie. **Węzły wyników** tworzą różne wyniki dla danych, wykresów i wyników dla modelu i znajdują się na kartach Wykresy, Wynik i Eksportuj w palecie węzłów. **Węzły modelowania** używają algorytmów statystycznych do tworzenia modeli użytkowych i znajdują się na karcie Modelowanie oraz na karcie Modelowanie w bazie (o ile została aktywowana) w palecie węzłów. Więcej informacji można znaleźć w temacie "Paleta węzłów" na stronie 15.

Łączenie węzłów pozwala na tworzenie strumieni, które po uruchomieniu umożliwiają wizualizację relacji i wyciąganie wniosków. Strumienie są podobne do skryptów — można je zapisać i użyć ponownie w innych plikach danych.

Możliwy do uruchomienia węzeł, który przetwarza dane strumienia, jest znany jako **węzeł końcowy**. Węzeł modelowania lub wyniku jest węzłem końcowym, jeśli znajduje się na końcu strumienia lub gałęzi strumienia. Nie można połączyć dalszych węzłów z węzłem końcowym.

Uwaga: Paletę Węzły można dostosować do indywidualnych potrzeb. Więcej informacji można znaleźć w temacie "Dostosowywanie palety węzłów" na stronie 244.

Dodawanie węzłów do strumienia

Węzły do strumienia można dodać z poziomu palety na kilka sposobów:

- Kliknij dwukrotnie węzeł na palecie. *Uwaga*: kliknięcie dwukrotne węzła powoduje jego automatyczne połączenie z bieżącym strumieniem. Więcej informacji można znaleźć w temacie "Łączenie węzłów w strumieniu" na stronie 36.
- Przeciągnij i upuść węzeł z palety na obszar roboczy strumienia.
- Kliknij węzeł na palecie, a następnie kliknij obszar roboczy strumienia.
- Wybierz odpowiednią opcję z menu Wstaw w programie IBM SPSS Modeler.

Po dodaniu węzła do obszaru roboczego strumienia kliknij go dwukrotnie w celu wyświetlenia okna dialogowego. Zakres dostępnych opcji jest zależny od typu dodawanego węzła. Informacje na temat określonych elementów sterujących dostępnych w oknie dialogowym są dostępne po kliknięciu przycisku **Pomoc**.

Usuwanie węzłów

Aby usunąć węzeł ze strumienia danych, kliknij go i naciśnij klawisz Delete lub kliknij prawym przyciskiem myszy i z menu wybierz polecenie **Usuń**.

Łączenie węzłów w strumieniu

Węzły dodane do obszaru roboczego strumienia nie tworzą strumienia danych, dopóki nie zostaną połączone. Połączenia między węzłami wskazują kierunek przepływu danych z jednej operacji do drugiej. Istnieje kilka sposobów łączenia węzłów w celu utworzenia strumienia: dwukrotne kliknięcie, użycie środkowego przycisku myszy lub metoda ręczna.

Aby dodać i połączyć węzły za pomocą dwukrotnego kliknięcia

Najprostszy sposób utworzenia strumienia to dwukrotne kliknięcie węzłów na palecie. Metoda ta umożliwia automatyczne łączenie nowego węzła z wybranym węzłem w obszarze roboczym strumienia. Na przykład, jeśli obszar roboczy zawiera węzeł Baza danych, można wybrać ten węzeł, a następnie dwukrotnie kliknąć następny węzeł z palety, taki jak węzeł wyliczeń. Ta czynność umożliwia automatyczne połączenie węzła wyliczania z istniejącym węzłem bazy danych. Proces ten można powtarzać aż do momentu osiągnięcia węzła końcowego, takiego jak węzeł Histogram lub Tabela, w którym to punkcie wszystkie nowe węzły zostaną połączone z ostatnim węzłem niebędącym węzłem końcowym, idąc w górę strumienia.

Aby połączyć węzły środkowym przyciskiem myszy

W obszarze roboczym strumienia można kliknąć i przeciągnąć z jednego węzła do drugiego, korzystając ze środkowego przycisku myszy. (Jeśli mysz nie ma środkowego przycisku, można go zasymulować, naciskając klawisz Alt podczas przeciągania myszą z jednego węzła do drugiego.)

Aby ręcznie połączyć węzły

Jeśli nie masz środkowego przycisku myszy i preferujesz ręczne połączenie węzłów, możesz użyć menu podręcznego węzła, podłączając go do innego węzła już znajdującego się w obszarze roboczym.

1. Kliknij prawym przyciskiem myszy węzeł, z którego chcesz rozpocząć łączenie. Spowoduje to otwarcie menu węzła.
2. W menu kliknij opcję **Połącz**.
3. Ikona łączenia jest wyświetlana zarówno w węźle początkowym, przy kursorze. Kliknij drugi węzeł w obszarze roboczym, aby utworzyć połączenie między obydwoma węzłami.

Podczas łączenia węzłów należy przestrzegać wytycznych. W przypadku próby utworzenia dowolnego z poniższych typów połączeń zostanie wyświetlony komunikat o błędzie:

- Połączenie prowadzące do węzła źródłowego
- Połączenie prowadzące od węzła końcowego
- Węzeł z większą niż maksymalna dla niego liczbą połączeń wejściowych
- Połączenie między dwoma węzłami, które są już połączone
- Odwołanie cykliczne (dane są zwracane do węzła, z którego przepłynęły)

Pomijanie węzłów w strumieniu

W przypadku pominięcia węzła w strumieniu danych wszystkie jego połączenia wejściowe i wyjściowe są zastępowane połączeniami prowadzącymi bezpośrednio z jego węzłów wejściowych do węzłów wyjściowych. Jeśli węzeł nie zawiera połączeń wejściowych oraz wyjściowych, wszystkie jego połączenia zostają usunięte i poprowadzone na nowo.

Przykładowo: może istnieć węzeł wyliczający nową zmienną, zmienne filtrów i badający wyniki tej operacji w histogramie i tabeli. Jeśli użytkownik chce również wyświetlić te same wykresy i tabele w przypadku danych *przed* odfiltrowaniem zmiennych, do strumienia można dodać nowy węzeł Histogram lub Tabela bądź pominąć węzeł Filtr. W przypadku pominięcia węzła Filtr połączenia z wykresem i tabelą wychodzą bezpośrednio z węzła Wyliczenie. Węzeł Filtr jest odłączony od strumienia.

Pomijanie węzła

1. Z poziomu obszaru roboczego strumienia za pomocą środkowego przycisku myszy kliknij dwukrotnie węzeł, który zostanie pominięty. Można również przytrzymać klawisz Alt i kliknąć dwukrotne.

Uwaga: tę operację można cofnąć, klikając polecenie **Cofnij** dostępne w menu Edycja lub naciskając klawisze Ctrl+Z.

Wyłączanie węzłów w strumieniu

Węzły procesowe z pojedynczymi danymi wejściowymi w ramach strumienia można wyłączyć, w efekcie czego węzły takie są ignorowane podczas uruchamiania strumienia. Eliminuje to konieczność usunięcia lub ominięcia węzła i oznacza, że może on pozostać połączony z pozostałymi węzłami. Nadal jednak można otwierać i edytować ustawienia węzła; zmiany nie odniosą jednak skutku do chwili ponownego włączenia węzła.

Na przykład można użyć strumienia filtrującego kilka zmiennych, a następnie tworzącego modele ze zredukowanym zestawem danych. W przypadku chęci utworzenia także tych samych modeli bez filtrowania zmiennych, w celu sprawdzenia, czy umożliwiają one poprawę wyników dla modelu, można wyłączyć węzeł Filtrowanie. Po wyłączeniu węzła Filtrowanie połączenia z węzłami modelowania przechodzą bezpośrednio z węzła Wyliczanie do węzła Typ.

Aby wyłączyć węzeł

1. W obszarze roboczym strumienia kliknij prawym przyciskiem myszy węzeł, który chcesz wyłączyć.
2. Kliknij opcję **Wyłącz węzeł** w menu podręcznym.

Można też kliknąć opcję **Węzeł** > **Wyłącz węzeł** w menu Edycja. Jeśli chcesz uwzględnić węzeł ponownie w strumieniu, kliknij opcję **Włącz węzeł** w ten sam sposób.

Uwaga: tę operację można cofnąć, klikając polecenie **Cofnij** dostępne w menu Edycja lub naciskając klawisze Ctrl+Z.

Dodawanie węzłów w istniejących połączeniach

Nowy węzeł można dodać między dwa połączone węzły, przeciągając strzałkę łączącą te dwa węzły.

1. Następnie należy kliknąć środkowym przyciskiem myszy i przeciągnąć strzałkę połączenia, do której zostanie wstawiony węzeł. Można również przytrzymać klawisz Alt podczas klikania i przeciągania, aby zasymulować kliknięcie środkowym przyciskiem myszy.
2. Przeciągnij połączenie do węzła, który zostanie uwzględniony i zwolnij przycisk myszy.

Uwaga: Nowe połączenia można usuwać z węzła i przywracać oryginalne, **pomijając** węzeł.

Usuwanie połączeń między węzłami

Aby usunąć połączenie między dwoma węzłami:

1. Kliknij prawym przyciskiem myszy strzałkę połączenia.
2. W menu kliknij opcję **Usuń połączenie**.

Aby usunąć wszystkie połączenia do oraz z węzła, wykonaj jedną z następujących czynności:

- Wybierz węzeł i naciśnij klawisz F3.
- Wybierz węzeł, a następnie w menu głównym kliknij:

Edycja > **Węzeł** > **Odłącz**

Określanie opcji węzłów

Po utworzeniu i połączeniu węzłów dostępnych jest kilka opcji pozwalających na dostosowanie węzłów. Kliknij węzeł prawym przyciskiem myszy i wybierz jedną z opcji menu.

- Kliknij opcję **Edytuj**, aby otworzyć okno dialogowe dla wybranego węzła.
- Kliknij opcję **Połącz**, aby ręcznie połączyć węzły.
- Kliknij opcję **Odłącz**, aby usunąć wszystkie połączenia prowadzące do i z węzła.
- Kliknij opcję **Zmień nazwę i skomentuj**, aby otworzyć kartę Adnotacje edytowanego okna dialogowego.
- Kliknij opcję **Nowy komentarz**, aby dodać komentarz związany z węzłem. Więcej informacji można znaleźć w temacie “Dodawanie komentarzy i adnotacji do węzłów i strumieni” na stronie 56.
- Kliknij opcję **Wyłącz węzeł**, aby „ukryć” węzeł podczas przetwarzania. Aby węzeł ponownie był widoczny podczas przetwarzania, kliknij opcję **Włącz węzeł**. Więcej informacji można znaleźć w temacie “Wyłączanie węzłów w strumieniu” na stronie 37.
- Kliknij opcję **Wytnij** lub **Usuń**, aby usunąć wybrane węzły z obszaru roboczego strumienia. *Uwaga*: Kliknięcie opcji **Wytnij** pozwala na wklejenie węzłów; opcja **Usuń** nie daje takiej możliwości.
- Kliknij opcję **Kopiuj węzeł**, aby utworzyć kopię węzła bez połączeń. Może on zostać dodany do nowego lub istniejącego strumienia.
- Kliknij opcje **Załaduj węzeł**, aby otworzyć wcześniej zapisany węzeł i załadować jego opcje do aktualnie wybranego węzła. Węzły muszą być takiego samego typu.
- Kliknij opcję **Pobierz węzeł**, aby pobrać węzeł z połączonego repozytorium IBM SPSS Collaboration and Deployment Services Repository.
- Kliknij opcję **Zapisz węzeł**, aby zapisać szczegóły węzła w pliku. Szczegóły węzła można załadować tylko do innego węzła tego samego typu.
- Kliknij opcję **Składuj węzeł**, aby zapisać wybrany węzeł w połączonym repozytorium IBM SPSS Collaboration and Deployment Services Repository.
- Kliknij opcję **Pamięć podręczna**, aby rozwinąć menu z opcjami zapisywania wybranego węzła w pamięci podręcznej.

- Kliknij opcję **Mapowanie danych**, aby rozwinąć menu z opcjami mapowania danych w nowym źródle lub określania pól obowiązkowych.
- Kliknij opcję **Utwórz Superwęzeł**, aby rozwinąć menu z opcjami tworzenia Superwęzła w bieżącym strumieniu.
- Kliknij opcję **Generuj węzeł danych użytkownika**, aby zastąpić wybrany węzeł. Przykłady wygenerowane za pośrednictwem tego węzła będą zawierały takie same zmienne, jak bieżący węzeł.
- Kliknij opcję **Wykonaj stąd**, aby uruchomić wszystkie węzły końcowe poniżej wybranego węzła.

Opcje buforowania węzłów

Aby zoptymalizować wykonywanie strumienia, można ustawić *pamięć podręczną* w dowolnym węźle niebędącym węzłem końcowym. Po skonfigurowaniu pamięci podręcznej dla węzła pamięć podręczna jest wypełniana danymi przechodzącymi przez węzeł przy następnym uruchomieniu strumienia danych. Od tego momentu dane są odczytywane z pamięci podręcznej (która jest przechowywana w katalogu tymczasowym na dysku), nie zaś ze źródła danych.

Buforowanie jest szczególnie przydatne w przypadku przeprowadzenia wcześniej czasochłonnej operacji takiej jak sortowanie, scalanie czy agregacja. Załóżmy na przykład, że mamy węzeł źródłowy z ustawieniem odczytu danych sprzedażowych z bazy danych oraz węzeł Agregacja podsumowujący sprzedaż wg lokalizacji. Istnieje możliwość skonfigurowania pamięci podręcznej w węźle Agregacja zamiast w węźle źródłowym, tak aby w pamięci podręcznej były przechowywane dane zagregowane, nie zaś cały zbiór danych.

Uwaga: Buforowanie w węzłach źródłowych, którego istotą jest po prostu przechowywanie kopii oryginalnych danych wczytywanych do IBM SPSS Modeler, w większości przypadków nie wpływa dodatnio na wydajność.

Węzły z włączonym buforowaniem są wyświetlane z niewielką ikoną dokumentu w prawym górnym rogu. W przypadku zbuforowania danych w węźle ikona dokumentu ma kolor zielony.

Aby włączyć pamięć podręczną

1. W obszarze roboczym strumienia kliknij węzeł prawym przyciskiem myszy i w menu wybierz pozycję **Pamięć podręczna**.
2. W podmenu buforowania kliknij opcję **Włącz**.
3. Pamięć podręczną można wyłączyć, klikając węzeł prawym przyciskiem myszy i klikając opcję **Wyłącz** w podmenu buforowania.

Buforowanie węzłów w bazie danych

W przypadku strumieni uruchamianych w bazie danych dane mogą być buforowane na bieżąco w tabeli tymczasowej w bazie danych zamiast w systemie plików. W przypadku połączenia z optymalizacją SQL może to skutkować znaczącymi korzyściami, jeśli chodzi o wydajność. Na przykład dane wynikowe ze strumienia scalającego wiele tabel z myślą o stworzeniu widoku eksploracji bazy danych mogą zostać zbuforowane i wykorzystane ponownie w razie potrzeby. Wydajność można dodatkowo podnieść, automatycznie generując kod SQL dla wszystkich kolejnych węzłów.

Aby korzystać z buforowania w bazie danych, należy włączyć zarówno optymalizację SQL, jak i buforowanie w bazie danych. Należy pamiętać, że ustawienia optymalizacji serwera zastępują odpowiadające im ustawienia klienta. Więcej informacji można znaleźć w temacie “Ustawianie opcji optymalizacji strumieni” na stronie 44.

Jeśli włączono buforowanie w bazie danych, należy po prostu kliknąć prawym przyciskiem myszy dowolny węzeł niebędący węzłem końcowym celu zbuforowania danych w tym punkcie. Spowoduje to automatyczne utworzenie pamięci podręcznej bezpośrednio w bazie danych przy następnym uruchomieniu strumienia. Jeśli nie włączono buforowania bazy danych lub optymalizacji SQL, wówczas pamięć podręczna zostanie zapisana w systemie plików.

Uwaga: Następujące bazy danych obsługują tymczasowe tabele do celów buforowania: Db2, Oracle, SQL Server i Teradata. W przypadku innych baz danych, takich jak Netezza, na potrzeby buforowania w bazie

danych będzie używana normalna tabela. Kod SQL można dostosować do konkretnych baz danych — w celu uzyskania pomocy należy skontaktować się z działem usług.

Opróżnianie pamięci podręcznej

Biała ikona dokumentu przy węźle oznacza, że pamięć podręczna jest pusta. Gdy pamięć podręczna jest wypełniona, ikona jest wyświetlana w kolorze zielonym. Aby zastąpić zawartość pamięci podręcznej, należy ją opróżnić, a następnie wypełnić ponownie, uruchamiając jeszcze raz strumień danych.

1. W obszarze roboczym strumienia kliknij węzeł prawym przyciskiem myszy i w menu wybierz pozycję **Pamięć podręczna**.
2. W podmenu pamięci podręcznej kliknij polecenie **Opróżnij**.

Zapisywanie pamięci podręcznej

Zawartość pamięci podręcznej można zapisać w postaci pliku danych IBM SPSS Statistics (`*.sav`). Następnie plik ten można wczytać ponownie jako pamięć podręczną lub skonfigurować węzeł korzystający z tego pliku jako źródła danych. Można również załadować pamięć podręczną zapisaną w innym projekcie.

1. W obszarze roboczym strumienia kliknij węzeł prawym przyciskiem myszy i w menu wybierz pozycję **Pamięć podręczna**.
2. W podmenu pamięci podręcznej kliknij przycisk **Zapisz pamięć podręczną**.
3. W oknie dialogowym Zapisz pamięć podręczną przejdź do położenia, w którym zostanie zapisany plik pamięci podręcznej.
4. W polu tekstowym Nazwa pliku wprowadź nazwę.
5. Upewnij się, że na liście typów plików wybrany jest typ ***.sav** i kliknij przycisk **Zapisz**.

Ładowanie pamięci podręcznej

Jeśli plik pamięci podręcznej został zapisany przed usunięciem go z węzła, ten plik można załadować ponownie.

1. W obszarze roboczym strumienia kliknij węzeł prawym przyciskiem myszy i w menu wybierz pozycję **Pamięć podręczna**.
2. W podmenu pamięci podręcznej kliknij polecenie **Załaduj pamięć podręczną**.
3. W oknie dialogowym Załaduj pamięć podręczną przejdź do położenia, w którym znajduje się plik pamięci podręcznej i kliknij przycisk **Załaduj**.

Podgląd danych w węzłach

Aby podczas tworzenia strumienia mieć pewność, że dane są zmieniane w oczekiwany sposób, można uruchomić dane za pośrednictwem węzła Tabela na każdym istotnym etapie. Aby tego uniknąć, można wygenerować podgląd z każdego węzła, który będzie przedstawiał przykładowe dane, jakie zostaną utworzone; pozwala to skrócić czas potrzebny na utworzenie modelu.

W przypadku węzłów poprzedzających model użytkowy podgląd będzie zawierał zmienne wejściowe; dla modelu użytkowego lub węzłów poniżej tego modelu (z wyjątkiem węzłów końcowych) podgląd będzie zawierał zmienne wejściowe i utworzone.

Domyślnie wyświetlanych jest 10 wierszy; ustawienie to można jednak zmienić we właściwościach strumienia. Więcej informacji można znaleźć w temacie “Określanie opcji ogólnych strumienia” na stronie 42.

Korzystając z menu **Utwórz**, można utworzyć kilka typów węzłów.

Uwaga: Podczas podglądu danych tworzonych przez ten węzeł wszystkie zmiany właściwości będą stosowane do węzła i nie można ich będzie anulować (tak jak po naciśnięciu przycisku **Zastosuj**).

Blokowanie węzłów

Aby uniemożliwić innym użytkownikom wprowadzanie zmian w ustawieniach jednego lub kilku węzłów w strumieniu, można ukryć węzeł lub węzły w specjalnym typie węzła o nazwie Superwęzeł, a następnie zablokować Superwęzeł, stosując zabezpieczenie hasłem.

Praca ze strumieniami

Strumień powstaje po połączeniu węzłów źródła, procesu i węzła końcowego w obszarze roboczym strumienia. Strumień, jako zbiór węzłów, można zapisywać, opatrywać adnotacjami i dodawać do projektów. Można także określać różne opcje dotyczące strumieni, np. ustawienia optymalizacji, daty i godziny, parametry i skrypty. Właściwości te omówione zostaną w następnych tematach.

W jednej sesji pracy z programem IBM SPSS Modeler można używać więcej niż jednego strumienia danych i modyfikować więcej niż jeden strumień programu IBM SPSS Modeler. Po prawej stronie okna głównego znajduje się panel zarządzania, który ułatwia nawigację między otwartymi strumieniami, wynikami i modelami. Jeśli panel zarządzania nie jest widoczny, kliknij polecenie **Zarządzanie** w menu Widok, a następnie kliknij kartę **Strumienie**.

Karta ta umożliwia:

- Uzyskiwanie dostępu do strumieni.
- Zapisywanie strumieni.
- Zapisywanie strumieni do bieżącego projektu.
- Zamykanie strumieni.
- Otwieranie nowych strumieni.
- Zapisywanie strumieni w repozytorium IBM SPSS Collaboration and Deployment Services i pobieranie ich z repozytorium (jeśli jest ono dostępne w środowisku użytkownika). Więcej informacji można znaleźć w temacie “Informacje o programie IBM SPSS Collaboration and Deployment Services Repository ” na stronie 203.

Aby uzyskać dostęp do tych opcji, kliknij strumień prawym przyciskiem myszy.

Ustawianie opcji strumieni

W bieżącym strumieniu można zastosować wiele opcji. Opcje te można również zapisać jako ustawienia domyślne z przeznaczeniem do zastosowania we wszystkich strumieniach. Poniżej przedstawiono dostępne opcje.

- **Ogólne.** Różnorodne opcje, takie jak symbole i kodowanie tekstu używane w strumieniu. Więcej informacji można znaleźć w temacie “Określanie opcji ogólnych strumienia” na stronie 42.
- **Oblicz datę i czas.** Opcje dotyczące formatu wyrażeń daty i czasu. Więcej informacji można znaleźć w temacie “Określanie daty i czasu strumieni” na stronie 43.
- **Formaty liczb.** Opcje sterujące formatem wyrażeń liczbowych. Więcej informacji można znaleźć w temacie “Określanie opcji formatu liczb dla strumieni” na stronie 44.
- **Optymalizacja.** Opcje służące do optymalizacji wydajności strumienia. Więcej informacji można znaleźć w temacie “Ustawianie opcji optymalizacji strumieni” na stronie 44.
- **Rejestr i status.** Opcje sterujące rejestrowaniem SQL i statusem rekordów. Więcej informacji można znaleźć w temacie “Konfiguracja opcji sterujących rejestrowaniem SQL i statusem rekordów w strumieniach” na stronie 46.
- **Układ.** Opcje dotyczące układu strumienia na obszarze roboczym. Więcej informacji można znaleźć w temacie “Określanie opcji układu strumieni” na stronie 46.
- **Analytic Server.** Opcje dotyczące używania oprogramowania Analytic Server wraz z SPSS Modeler. Więcej informacji można znaleźć w temacie “Właściwości strumienia Analytic Server” na stronie 47.
- **Geoprzestrzenne.** Opcje dotyczące formatowania danych geoprzestrzennych w celu wykorzystania w strumieniu. Więcej informacji można znaleźć w temacie “Konfigurowanie opcji geoprzestrzennych strumieni” na stronie 47.

Konfigurowanie opcji strumienia

1. W menu Plik kliknij pozycję **Właściwości strumienia** (lub wybierz strumień na karcie Strumienie w panelu zarządzania, klikając prawym przyciskiem myszy i wybierając z menu podręcznego pozycję **Właściwości strumienia**).
2. Kliknij kartę **Opcje**.

W menu Narzędzia można również kliknąć:

Właściwości strumienia > **Opcje**

Określanie opcji ogólnych strumienia

Opcje ogólne to zestaw opcji dotyczących różnorodnych aspektów bieżącego strumienia.

W sekcji **Podstawowe** dostępne są następujące opcje podstawowe:

- **Separator dziesiętny.** Wybierz przecinek (,) lub kropkę (.) jako separator dziesiętny.
- **Symbol grupowania.** W celu określenia formatów wyświetlania liczb wybierz symbol używany do grupowania wartości (np. przecinek: 3,000.00). Dostępne opcje to: brak, kropka, przecinek, spacja, symbol zdefiniowany w ustawieniach regionalnych (wtedy wybierane jest ustawienie domyślne opcji regionalnych).
- **Kodowanie.** Określ domyślną metodę kodowania tekstu w strumieniu. (*Uwaga*: dotyczy tylko węzła źródłowego pliku zmiennego i węzła eksportu pliku płaskiego. Żaden inny węzeł nie korzysta z tego ustawienia. W większości przypadków informacje o kodowaniu są osadzone w pliku danych). Można wybrać domyślne ustawienie systemowe lub UTF-8. Domyślne ustawienia systemowe są określone w Panelu sterowania systemu Windows (lub w przypadku trybu rozproszonego — na serwerze). Więcej informacji można znaleźć w temacie “Obsługa kodowania Unicode w IBM SPSS Modeler ” na stronie 267.
- **Ocena zestawu reguł.** Określa sposób oceny w modelach zestawów reguł. Domyślnie zestawy reguł korzystają z opcji **Głosowanie** w celu łączenia predykcji z poszczególnych ról i określania ostatecznej predykcji. Aby mieć pewność, że zestawy reguł domyślnie korzystają z reguły pierwszego trafienia, należy wybrać opcję **Pierwsze trafienie**. Ta opcja nie wpływa na modele Listy decyzyjnej, które zawsze korzystają z pierwszego trafienia zgodnie z definicją określoną w algorytmie.

Maksymalna liczba wierszy wyświetlana w podglądzie danych. Określ liczbę pokazywanych wierszy, gdy wymagane jest zaprezentowanie danych węzła. Więcej informacji można znaleźć w temacie “Podgląd danych w węzłach” na stronie 40.

Maksimum członków dla zmiennych nominalnych. Wybierz tę opcję, aby określić maksymalną liczbę członków dla zmiennych nominalnych (zestaw), po przekroczeniu której dane zaczynają być określane jako **Bez typu**. Ta opcja jest przydatna podczas pracy z dużymi zmiennymi nominalnymi. *Uwaga*: gdy dla poziomu pomiaru zmiennej wybrano ustawienie **Bez typu**, dla roli jest automatycznie określane ustawienie **Brak**. Oznacza to, że zmienne nie podlegają modelowaniu.

Ogranicz liczbę kategorii zmiennej dla modelowania Kohonena i K-średnich. Wybierz, aby określić maksymalną liczbę członków dla zmiennych nominalnych używanych w modelowaniu sieci Kohonena i *K*-średnich. Wielkością domyślną zestawu jest 20; po przekroczeniu tej wartości zmienna jest ignorowana i wyświetlane jest ostrzeżenie z informacją o danej zmiennej.

W celu zachowania zgodności ta opcja również jest stosowana w przypadku starego węzła Sieć neuronowa. Ten węzeł został zastąpiony w wersji 14. oprogramowania IBM SPSS Modeler. Jednak niektóre starsze strumienie mogą nadal zawierać ten węzeł.

Odśwież węzły źródłowe danych przy wykonywaniu strumienia. Wybierz, aby automatycznie odświeżyć wszystkie węzły źródłowe podczas wykonywania bieżącego strumienia. Ta czynność odpowiada kliknięciu przycisku **Odśwież** dostępnego w węźle źródłowym, jednak zapewnia ona automatyczne odświeżenie wszystkich węzłów źródłowych (poza węzłami danych użytkownika) bieżącego strumienia.

Uwaga: Użycie tej opcji powoduje opróżnienie pamięci podręcznej kolejnych węzłów, nawet jeśli dane nie uległy zmianie. W przypadku zastosowania opcji **RuWykonaj bieżący strumień** opróżnianie jest

wykonywane jeden raz na każde wykonanie strumienia, co oznacza, że nadal można korzystać z kolejnych pamięci podręcznych jako magazynów tymczasowych w przypadku pojedynczych wykonań. Załóżmy, że utworzyliśmy pamięć podręczną w środku strumienia, już za złożoną operacją wyliczenia, a za tym węzłem wyliczenia znajduje się kilka dołączonych wykresów oraz raportów. Podczas wykonywania strumienia pamięć podręczna w węźle wyliczenia zostanie opróżniona i wypełniona ponownie, jednak tylko dla pierwszego wykresu lub raportu. Kolejne węzły końcowe odczytają dane z pamięci podręcznej węzła wyliczenia. Należy zwrócić uwagę, że jeśli każdy węzeł końcowy będzie wykonywany osobno (gdy istnieje więcej niż jeden), a nie za pośrednictwem opcji **Wykonaj bieżący strumień**, to pamięć podręczna będzie opróżniana przy każdym wykonaniu węzła końcowego.

Wyświetlaj w wynikach etykiety zmiennych i wartości. Wyświetla etykiety wartości i zmiennych w tabelach, wykresach i innych danych wyjściowych. Jeśli tabele nie istnieją, wyświetlone zostaną nazwy zmiennych i wartości danych. Etykiety są domyślnie wyłączone, jednak można je włączyć pojedynczo w oprogramowaniu IBM SPSS Modeler. Użytkownik może również wyświetlić etykiety w oknie wynikowym za pomocą przełącznika dostępnego na pasku narzędzi.

Rysunek 11. Ikona na pasku narzędzi służąca do przełączania widoczności etykiet wartości i zmiennych

Wyświetl czasy obliczeń. Wyświetla poszczególne czasy wykonania dla węzłów strumienia na karcie Czasy wykonania po wykonaniu strumienia. Więcej informacji można znaleźć w temacie "Wyświetlanie czasu wykonywania węzła" na stronie 49.

Sekcja **Automatyczne tworzenie węzłów** zawiera poniższe opcje pozwalające na automatyczne tworzenie węzłów w poszczególnych strumieniach. Opcje te sterują wstawianiem modeli użytkowych w obszarze roboczym strumienia podczas tworzenia nowych modeli użytkowych. Domyślnie opcje te odnoszą się do strumieni utworzonych w wersji 16 lub nowszej. W oprogramowaniu IBM SPSS Modeler 16 lub nowszym po otwarciu strumienia utworzonego w wersji 15 lub starszej i wykonaniu węzła modelowania model użytkowy nie zostanie umieszczony w obszarze roboczym strumienia, tak jak to miało miejsce w przypadku poprzednich wydań. Utworzenie nowego strumienia w oprogramowaniu IBM SPSS Modeler 16 lub nowszym i wykonanie węzła modelowania powoduje umieszczenie utworzonego modelu użytkowego w obszarze roboczym strumienia. Jest to zamierzone, ponieważ np. użycie opcji **Utwórz w strumieniu węzły modeli dla nowych wyników modelowania** prawdopodobnie spowoduje awarię strumieni utworzonych w wersjach starszych niż 16. uruchamianych wsadowo w oprogramowaniu IBM SPSS Collaboration and Deployment Services oraz w innych środowiskach pozbawionych interfejsu użytkownika klienta IBM SPSS Modeler Server.

- **Utwórz w strumieniu węzły modeli dla nowych wyników modelowania.** Automatycznie tworzy węzły stosowania modelu dla nowych wyników modelowania. Po wybraniu tej opcji w obszarze **Utwórz połączenia modeli z węzłami modelowania** można również określić, czy łącza są włączone lub wyłączone, bądź nie tworzyć ich.

  Gdy zostanie utworzony węzeł stosowania modelu lub węzeł źródłowy, opcje łącza w menu rozwijanym umożliwiają określenie, czy zostaną utworzone łącza aktualizacji między węzłem tworzenia i nowym węzłem. Jeśli zostaną one utworzone, można określić tryb tych łączy. W przypadku utworzenia łączy użytkownik prawdopodobnie będzie je chciał włączyć, a dzięki tym opcjom użytkownik dysponuje pełną kontrolą nad tymi elementami.
- **Aktualizuj węzły źródłowe danymi z węzłów tworzących źródła.** Automatycznie tworzy węzły źródłowe na podstawie węzłów tworzących źródła. Podobnie jak w przypadku poprzedniej opcji, po wybraniu tej opcji za pomocą menu rozwijanego **Utwórz połączenia aktualizujące źródła** można określić, czy łącza odświeżania są włączone lub wyłączone, bądź nie tworzyć ich.

Zapisz jako domyślne. Określone opcje dotyczą tylko bieżącego strumienia. Kliknij ten przycisk, aby określić wybrane opcje jako ustawienia domyślne wszystkich strumieni.

Określanie daty i czasu strumieni

Te opcje umożliwiają określanie formatu wyrażeń daty i godziny w bieżącym strumieniu.

Importuj datę/czas jako Określ, czy korzystać z typu daty/czasu w przypadku zmiennych daty/czasu czy zaimportować te elementy jako zmienne łańcuchowe.

Format daty Wybierz format daty, który będzie używany w przypadku zmiennych typu data lub gdy łańcuchy są interpretowane jako daty przez funkcje dat CLEM.

Format czasu Wybierz format czasu, który będzie używany w przypadku zmiennych typu czas lub gdy łańcuchy są interpretowane jako czas przez funkcje czasu CLEM.

Cofnij do dni/godzin W przypadku formatów czasu określ, czy ujemne różnice czasu będą interpretowane jako odniesienie do poprzedniego dnia lub poprzedniej godziny.

Data bazowa (1 stycznia) Wybierz lata bazowe (zawsze 1 stycznia), które będą używane przez funkcje daty CLEM obsługujące pojedynczą datę.

2-cyfrowe daty zaczynają się od Określ rok odcięcia w celu dodania cyfr wieku w przypadku lat prezentowanych tylko za pomocą dwóch cyfr. Przykładowo: określenie wartości 1930 jako roku odcięcia spowoduje określenie daty 05/11/02 jako daty w roku 2002. To samo ustawienie spowoduje przyjęcie 20. wieku w przypadku dat po 30; więc przyjmuje się, że data 05/11/73 przypada w roku 1973.

Strefa czasowa Wybierz sposób wyboru strefy czasowej do użytku z wyrażeniem datetime_now CLEM.

- Jeśli wybierzesz **Serwer**, wówczas strefa czasowa będzie zależna od następujących elementów:
  - Jeśli aktualny strumień korzysta ze źródła danych produktu Analytic Server, wówczas wyrażenie datetime_now używa czasu z produktu Analytic Server; domyślnie serwer stosuje czas uniwersalny.
  - Jeśli aktualny strumień korzysta z węzła źródłowego bazy danych, wówczas obsługiwane bazy danych używają analizy wstępnej SQL, a wyrażenie datetime_now używa czasu bazy danych.
  - W przypadku wszystkich innych strumieni strefa czasowa używa czasu z serweraSPSS Modeler Server.
- Jeśli wybierzesz opcję **Klient Modeler**, strefa czasowa będzie odzwierciedlać szczegóły strefy czasowej z komputera, na którym zainstalowany jest produkt SPSS Modeler.
- Alternatywnie dla strefy czasowej można wybrać dowolną z wartości czasu uniwersalnego.

Zapisz jako domyślne. Określone opcje dotyczą tylko bieżącego strumienia. Kliknij ten przycisk, aby określić wybrane opcje jako ustawienia domyślne wszystkich strumieni.

Określanie opcji formatu liczb dla strumieni

Te opcje umożliwiają określanie formatu różnorodnych wyrażeń liczbowych w bieżącym strumieniu.

Format wyświetlania liczb. Użytkownik może wybrać format wyświetlania: standardowy (**####.###**), naukowy (**#.###E+##**) lub walutowy (**$###.##**).

Miejsca dziesiętne (standardowy, naukowy, walutowy). W przypadku formatów wyświetlania liczb określa liczbę miejsc dziesiętnych, które zostaną użyte podczas wyświetlania lub drukowania liczb rzeczywistych. Tę opcję należy ustawić oddzielnie dla każdego formatu wyświetlania.

Wyliczenia w. Jako jednostkę miary w przypadku wyrażeń trygonometrycznych CLEM, wybierz opcję **Radiany** lub **Stopnie**. Więcej informacji można znaleźć w temacie “Funkcje trygonometryczne” na stronie 175.

Zapisz jako domyślne. Określone opcje dotyczą tylko bieżącego strumienia. Kliknij ten przycisk, aby określić wybrane opcje jako ustawienia domyślne wszystkich strumieni.

Ustawianie opcji optymalizacji strumieni

Ustawienia Optymalizacja umożliwiają optymalizację wydajności strumienia. Należy zwrócić uwagę na fakt, że ustawienia wydajności i optymalizacji na serwerze IBM SPSS Modeler Server (jeśli jest on używany) zastępują odpowiadające im ustawienia w komputerze klienckim. Jeśli te ustawienia są wyłączone na serwerze, wówczas nie ma możliwości włączenia ich z komputera klienckiego. Z kolei, jeśli są one włączone na serwerze, istnieje możliwość wyłączenia ich na komputerze klienckim.

Uwaga: Modelowanie bazy danych i optymalizacja SQL wymagają włączenia na komputerze z programem IBM SPSS Modeler możliwości połączenia z serwerem IBM SPSS Modeler Server. Po włączeniu tej opcji można uzyskać dostęp do algorytmów baz danych, wstawić SQL do kolejki bezpośrednio z programu IBM SPSS Modeler i uzyskać dostęp do programu IBM SPSS Modeler Server. W celu sprawdzenia bieżącego statusu licencji należy wybrać z menu programu IBM SPSS Modeler następujące opcje.

Pomoc > **Informacje o programie** > **Dodatkowe szczegóły**

Po włączeniu możliwości połączenia na karcie Status licencji widoczna jest opcja **Aktywacja serwera**.

Więcej informacji zawiera temat "Łączenie się z serwerem IBM SPSS Modeler Server " na stronie 10.

Uwaga: Obsługa funkcji analiz wstępnych SQL oraz optymalizacji zależy od rodzaju używanej bazy danych. W celu uzyskania najnowszych informacji na temat obsługiwanych i przetestowanych pod kątem współpracy z produktem IBM SPSS Modeler baz danych i sterowników ODBC należy odwiedzić korporacyjny serwis wsparcia pod adresem http://www.ibm.com/support.

Zezwalaj na przepisywanie strumienia w celu optymalizacji. Tę opcję należy wybrać, aby zezwolić na przepisywanie strumienia w produkcie IBM SPSS Modeler. Dostępne są cztery rodzaje przepisywania, z których można wybrać co najmniej jeden. Przepisywanie strumienia zmienia kolejność węzłów w strumieniu w tle, podnosząc efektywność przetwarzania bez modyfikowania semantyki strumienia.

- **Optymalizuj operacje generujące kod SQL.** Ta opcja umożliwia zmianę kolejności węzłów w strumieniu tak, aby możliwe było przygotowanie kodu SQL do wykonania w bazie danych większej liczby operacji. W przypadku znalezienia węzła, którego nie można wyrazić w formie kodu SQL, optymalizator będzie antycypować obecność kolejnych węzłów dających się wyrazić w języku SQL i bezpiecznie przenieść przed węzeł z problemem bez wpływu na semantykę strumienia. Baza danych wykonuje operacje bardziej efektywnie niż program IBM SPSS Modeler wykonywanie operacji, a ponadto wstępne przekształcenie węzła do postaci kodu SQL ogranicza objętość danych zwracanych do programu IBM SPSS Modeler celem przetwarzania. To zaś może spowodować zmniejszenie ruchu w sieci i przyspieszenie operacji na strumieniach. Należy zwrócić uwagę na fakt, że pole wyboru **Generuj kod SQL** musi być zaznaczone, aby optymalizacja SQL dała jakikolwiek efekt.
- **Optymalizuj wyrażenie CLEM.** Ta opcja umożliwia wyszukiwanie wyrażeń CLEM przez optymalizator, tak aby mogły być one przetworzone wstępnie przed uruchomieniem strumienia w celu zwiększenia prędkości przetwarzania. A oto prosty przykład: w przypadku wyrażenia takiego jak *log(wynagrodzenie)* optymalizator oblicza rzeczywistą wartość wynagrodzenia i przekazuje ją to dalszego przetwarzania. Umożliwia to zarówno poprawę analiz wstępnych SQL jak i optymalizację wydajności programu IBM SPSS Modeler Server.
- **Optymalizuj wykonywanie komend.** Ta metoda przepisywania strumienia zwiększa efektywność operacji obejmujących więcej niż jeden węzeł zawierający komendy IBM SPSS Statistics. Optymalizacja odbywa się przez połączenie komend w pojedynczą operację zamiast uruchomienia każdej z nich jako osobnej operacji.
- **Optymalizuj inne wykonywane operacje.** Ta metoda przepisywania strumienia zwiększa efektywność operacji, których nie można przekazać do bazy danych. Optymalizację osiąga się, redukując ilość danych w strumieniu tak szybko, jak to tylko możliwe. Strumień jest przepisywany (z zachowaniem integralności danych), co pozwala na zbliżenie operacji do źródła danych, a to z kolei umożliwia ograniczenie ilości danych podlegających kosztownym operacjom, takim jak łączenia, w dalszych węzłach.

Włącz przetwarzanie równoległe. W przypadku korzystania z komputera wieloprocesorowego opcja ta pozwala na zrównoważenie obciążenia tych procesorów, co może skutkować jego szybszym działaniem. Przetwarzanie równoległe może okazać się korzystne w przypadku używania wielu węzłów lub używania następujących pojedynczych węzłów: C5.0, Łączenie (wg klucza), Sortowanie, Kategoria (metody rangi i N-tyla) oraz Agregacja (z użyciem jednej lub większej liczby zmiennych kluczowych).

Generuj kod SQL. Wybierz tę opcję, aby włączyć przekształcanie węzłów w kod SQL, który następnie będzie przekazywany do bazy danych celem wykonania z większą wydajnością. W celu dalszego zwiększenia wydajności można także wybrać opcję **Optymalizuj generowanie kodu SQL** w celu zmaksymalizowania liczby operacji kierowanych z powrotem do bazy danych. Gdy operacje na węźle zostały skierowane do bazy danych, węzeł ten zostanie podświetlony na purpurowo po uruchomieniu strumienia.

- **Buforowanie w bazie danych.** W przypadku strumieni generujących kod SQL do wykonania w bazie danych dane mogą być buforowane na bieżąco w tabeli tymczasowej w bazie danych zamiast w systemie plików. W przypadku połączenia z optymalizacją SQL może to skutkować znaczącymi korzyściami, jeśli chodzi o wydajność. Na przykład dane wynikowe ze strumienia scalającego wiele tabel z myślą o stworzeniu widoku eksploracji bazy danych mogą zostać zbuforowane i wykorzystane ponownie w razie potrzeby. Jeśli włączono buforowanie w bazie danych, należy po prostu kliknąć prawym przyciskiem myszy dowolny węzeł niebędący węzłem końcowym w celu zbuforowania danych w tym punkcie. Spowoduje to automatyczne utworzenie pamięci podręcznej bezpośrednio w bazie danych przy następnym uruchomieniu strumienia. Umożliwia to wygenerowanie kodu SQL dla wszystkich kolejnych węzłów, co dodatkowo zwiększa wydajność. Alternatywnie tę opcję można w razie potrzeby wyłączyć — na przykład w sytuacji, gdy polityki lub uprawnienia wykluczają zapisywanie danych w bazie danych. Jeśli nie włączono buforowania bazy danych lub optymalizacji SQL, wówczas pamięć podręczna zostanie zapisana w systemie plików. Więcej informacji można znaleźć w temacie "Opcje buforowania węzłów" na stronie 39.
- **Użyj swobodnego przekształcenia formatów.** Ta opcja umożliwia przekształcenie danych z łańcuchów na liczby lub z liczb na łańcuchy, o ile są one zapisywane w odpowiednim formacie. Na przykład, jeśli dane są przechowywane w bazie danych jako łańcuchy, lecz w rzeczywistości zawierają znaczącą liczbę, wówczas można przekształcić je w celu ich wykorzystania podczas wstawiania do kolejki.

Uwaga: Z powodu niewielkich różnic dotyczących wdrażania kodu SQL strumienie uruchamiane w bazie danych mogą zwracać nieznacznie inne wyniki niż te zwracane podczas uruchamiania w programie IBM SPSS Modeler. Z podobnych przyczyn różnice te mogą się również występować w zależności od dostawcy bazy danych.

Zapisz jako domyślne. Określone opcje dotyczą tylko bieżącego strumienia. Kliknij ten przycisk, aby określić wybrane opcje jako ustawienia domyślne wszystkich strumieni.

Konfiguracja opcji sterujących rejestrowaniem SQL i statusem rekordów w strumieniach

Te ustawienia obejmują różnorodne opcje pozwalające na sterowanie wyświetlaniem instrukcji SQL generowanych przez strumień oraz sterowanie wyświetlaniem liczby rekordów przetwarzanych przez strumień.

Wyświetlaj kod SQL jako komunikaty w trakcie wykonywania strumienia. Określa, czy kod SQL wygenerowany podczas wykonywania strumienia jest przekazywany do dziennika komunikatów.

Pokaż w komunikatach szczegóły kodu SQL w fazie przygotowania strumienia. Podczas wyświetlania podglądu strumienia określa, czy podgląd kodu SQL, który zostanie wygenerowany, będzie przekazany do dziennika komunikatów.

Wyświetl kod SQL. Określa, czy kod SQL wyświetlany w dzienniku będzie zawierać natywne funkcje SQL czy standardowe funkcje ODBC w postaci `{fn FUNC(...)}` tak jak zostały wygenerowane w oprogramowaniu SPSS Modeler. Działanie drugiej opcji opiera się na funkcjach sterownika ODBC, które mogły nie zostać zaimplementowane.

Zmień format SQL dla zwiększenia czytelności. Określa, czy kod SQL wyświetlany w dzienniku będzie sformatowany w celu zwiększenia jego czytelności.

Pokaż status dla rekordów. Określa, kiedy należy zgłaszać rekordy wpływające do węzłów końcowych. Określ liczbę używaną podczas aktualizowania statusu co *N* rekordów.

Zapisz jako domyślne. Określone opcje dotyczą tylko bieżącego strumienia. Kliknij ten przycisk, aby określić wybrane opcje jako ustawienia domyślne wszystkich strumieni.

Określanie opcji układu strumieni

Te ustawienia zapewniają dostęp do wielu opcji wyświetlania i obsługi obszarów roboczych strumieni.

Minimalna szerokość obszaru roboczego strumienia. Określ minimalną szerokość obszaru roboczego strumienia w pikselach.

Minimalna wysokość obszaru roboczego strumienia. Określ minimalną wysokość obszaru roboczego strumienia w pikselach.

Szybkość przewijania strumienia. Określ szybkość przewijania obszaru roboczego strumienia w celu skonfigurowania prędkości przesuwania panelu obszaru roboczego strumienia podczas przeciągania węzła na obszarze roboczym. Wyższe wartości oznaczają większą szybkość przewijania.

Maksymalna długość nazwy ikony. Określ ograniczenie liczby znaków w nazwie węzła w obszarze roboczym strumienia.

Rozmiar ikony. Cały widok strumienia można przeskalować do dowolnego rozmiaru w zakresie od 8% do 200% w odniesieniu do standardowego rozmiaru ikony.

Rozmiar komórki siatki. Na liście wybierz rozmiar komórki siatki. Wybrana wartość służy do wyrównywania położenia węzłów na obszarze roboczym strumienia z zastosowaniem niewidocznej siatki. Domyślna wielkość komórki siatki to 0,25.

Wyrównaj do siatki. Za pomocą tej opcji można wyrównać ikony względem niewidocznej siatki (opcja domyślnie włączona).

Umieszczenie ikon generowanych węzłów. Wybierz położenie na obszarze roboczym, w którym zostaną umieszczone ikony węzłów utworzonych na podstawie modeli użytkowych. Położeniem domyślnym jest lewy górny narożnik.

Zapisz jako domyślne. Określone opcje dotyczą tylko bieżącego strumienia. Kliknij ten przycisk, aby określić wybrane opcje jako ustawienia domyślne wszystkich strumieni.

Właściwości strumienia Analytic Server

Ustawienia te oferują szereg sposobów pracy z programem Analytic Server.

Maksymalna liczba rekordów do przetworzenia poza programem Analytic Server
Określ maksymalną liczbę rekordów do zaimportowania na serwer SPSS Modeler ze źródła danych Analytic Server.

Powiadomienie, gdy węzła nie można przetworzyć w programie Analytic Server
Ustawienie to określa, co dzieje się, gdy strumień, który powinien zostać przesłany do programu Analytic Server, zawiera węzeł, którego nie można przetworzyć w programie Analytic Server. Zdecyduj, czy powinno zostać wygenerowane ostrzeżenie, a następnie powinna być możliwa kontynuacja przetwarzania strumienia, czy też powinien zostać wygenerowany błąd i przetwarzanie powinno zostać zatrzymane.

Ustawienia składowania modelu rozdzielonego

Składuj modele jako referencje do modeli w Analytic Server Analytic Server, gdy ich wielkość przekracza (MB)
Modele użytkowe są zwykle zapisywane jako część strumienia. Modele rozdzielone z wieloma podziałami mogą generować duże modele użytkowe, zaś przenoszenie modelu użytkowego między strumieniem a programem Analytic Server może negatywnie wpłynąć na wydajność. W efekcie, gdy model rozdzielony przekracza określoną wielkość, jest przechowywany na serwerze Analytic Server, zaś model użytkowy w programie SPSS Modeler zawiera odwołanie do modelu.

Domyślny folder składowania modeli referencyjnych na serwerze Analytic Server po zakończeniu wykonywania
Wskaż ścieżkę domyślną na serwerze Analytic Server, w której mają być przechowywane modele rozdzielone. Ścieżka powinna rozpoczynać się od poprawnej nazwy projektu Analytic Server.

Folder przechowywania awansowanych modeli
Wskaż ścieżkę domyślną na serwerze, w której mają być przechowywane „wypromowane” modele. Model wypromowany nie jest kasowany po zakończeniu sesji SPSS Modeler.

Konfigurowanie opcji geoprzestrzennych strumieni

Ze wszystkimi zmiennymi geoprzestrzennymi np. kształtami, współrzędnymi lub wartościami pojedynczymi osi (x lub y bądź szerokość i długość geograficzna) powiązane są układy współrzędnych.

Ten układ współrzędnych określa atrybuty, takie jak punkt początkowy (0,0) i jednostki powiązanie z wartościami.

Istnieje wiele układów współrzędnych oraz dwa typy: geograficzne i rzutowane. Wszystkich funkcji przestrzennych w oprogramowaniu SPSS Modeler można używać tylko w rzutowanym układzie współrzędnych.

Ze względu na charakter układów współrzędnych scalanie lub dołączanie danych z dwóch oddzielnych źródeł danych geoprzestrzennych wymaga, aby te źródła korzystały z tego samego układu współrzędnych. Dlatego też należy określić ustawienie współrzędnych w przypadku wszystkich danych geoprzestrzennych używanych w strumieniu.

W następujących sytuacjach dane są automatycznie ponownie odwzorowane w celu korzystania z wybranego układu współrzędnych strumienia:

- W przypadku funkcji przestrzennych (np. area, closeto, within) parametr przekazywany do funkcji jest automatycznie rzutowany ponownie, jednak oryginalne dane wiersza nie ulegają zmianie.
- Podczas używania węzłów budowania lub oceniania (model użytkowy) w ramach predykcji przestrzenno-czasowej (STP) zmienna lokalizacji jest automatycznie rzutowana ponownie. Podczas oceniania lokalizacja pochodząca z modelu użytkowego stanowi lokalizację oryginalną.
- Podczas używania węzła Wizualizacja na mapie.

Układ współrzędnych obowiązujący w strumieniu. Ta opcja jest dostępna wyłącznie po zaznaczeniu pola wyboru. Kliknij pozycję **Zmień**, aby wyświetlić listę dostępnych odwzorowanych układów współrzędnych, i wybierz układ, który zostanie użyty w bieżącym strumieniu.

Zapisz jako domyślne. Wybrany układ współrzędnych jest stosowany tylko w przypadku bieżącego strumienia. Kliknij ten przycisk, aby określić układ jako domyślny dla wszystkich strumieni.

Wybór geoprzestrzennych układów współrzędnych

Wszystkie funkcje przestrzenne w oprogramowaniu SPSS Modeler mogą być używane tylko w rzutowanym układzie współrzędnych.

Okno dialogowe **Układ współrzędnych obowiązujący w strumieniu** zawiera listę wszystkich rzutowanych układów współrzędnych, które można wybrać dla dowolnych danych geoprzestrzennych używanych w strumieniu.

Dla każdego układu współrzędnych wymieniono następujące informacje.

- **WKID** Dobrze znany identyfikator, unikalny dla każdego układu współrzędnych.
- **Nazwa** Nazwa układu współrzędnych.
- **Jednostki** Jednostka miary powiązana z układem współrzędnych.

Poza listą wszystkich układów współrzędnych w oknie dialogowym dostępny jest element sterujący **Filtrowanie**. Jeśli znasz całą nazwę układu współrzędnych lub tylko jej część, wpisz ją w polu **Nazwa** w dolnej części okna dialogowego. Lista układów współrzędnych, z której można dokonać wyboru, jest filtrowana automatycznie, co pozwala wyświetlić tylko systemy o nazwach zawierających wprowadzony tekst.

Wyświetlanie komunikatów o działaniu strumienia

Komunikaty dotyczące działań w ramach strumienia, np. wykonywanie, optymalizacja i czas budowania i oceniania modelu, można wyświetlać z poziomu karty Komunikaty dostępnej w oknie dialogowym właściwości strumienia. W tej samej tabeli są również wyświetlane komunikaty o błędach.

Wyświetlanie komunikatów - strumienie

1. W menu Plik kliknij pozycję **Właściwości strumienia** (lub wybierz strumień na karcie Strumienie w panelu zarządzania, klikając prawym przyciskiem myszy i wybierając z menu podręcznego pozycję **Właściwości strumienia**).
2. Kliknij kartę **Komunikaty**.

W menu Narzędzia można również kliknąć:

Właściwości strumienia > **Wykonywanie**

W tym obszarze dostępne są komunikaty dotyczące operacji związanych ze strumieniami oraz komunikaty o błędach. Gdy wykonywanie strumienia zostanie przerwane z powodu błędu, w tym oknie dialogowym zostanie otwarta karta Komunikaty zawierająca odpowiedni komunikat o błędzie. Ponadto węzeł z błędami w obszarze roboczym strumienia zostanie zaznaczony kolorem czerwonym.

Jeśli w oknie dialogowym Opcje użytkownika wybrano opcje rejestrowania i optymalizacji SQL, wyświetlane są również informacje o wygenerowanym kodzie SQL. Więcej informacji można znaleźć w temacie “Ustawianie opcji optymalizacji strumieni” na stronie 44.

Tutaj można zapisać komunikaty zgłaszane przez strumień. W tym celu należy kliknąć pozycję **Zapisz komunikaty** dostępną w menu rozwijanym przycisku Zapisz (po lewej stronie pod kartą Komunikaty).

Za pomocą opcji **Wyczyść komunikaty** dostępnej w menu rozwijanym można wyczyścić wszystkie komunikaty dotyczące danego strumienia.

Czas procesora to czas, przez który proces serwera korzysta z procesora. Czas, który upłynął, to całkowity czas między rozpoczęciem i zakończeniem wykonywania, więc okres ten obejmuje także takie działania, jak przenoszenie plików i prezentowanie wyników. Wartość czasu procesora może być większa niż wartość czasu, który upłynął, gdy strumień korzysta z wielu procesorów (działanie równoległe). Gdy strumień zostanie w całości przekazany z powrotem do bazy danych będącej źródłem danych, czas procesora będzie wynosił zero.

Wyświetlanie czasu wykonywania węzła

Na karcie Komunikaty można wyświetlić pozycję Czasy wykonania. Umożliwia ona zapoznanie się z poszczególnymi czasami wykonywania węzłów w strumieniach, które są wykonywane w oprogramowaniu IBM SPSS Modeler Server. Czasy mogą być niedokładne w przypadku strumieni uruchamianych w innych obszarach, np. jako programy R lub Analytic Server. Możliwe jest również obliczanie czasu wykonywania niektórych węzłów.

Uwaga: Działanie tej funkcji jest uzależnione od zaznaczenia opcji **Wyświetl czasy obliczeń** w ustawieniach **Ogólne** na karcie **Opcje**.

W tabeli czasów wykonywania węzłów dostępne są następujące kolumny. Kliknięcie nagłówka kolumny pozwala na posortowanie wpisów w kolejności rosnącej lub malejącej (w ten sposób można np. wyświetlić węzły, których czas wykonywania jest najdłuższy).

Węzeł końcowy. Identyfikator gałęzi, do której należy węzeł. Identyfikator to nazwa węzła końcowego znajdującego się na zakończeniu gałęzi.

Etykieta węzła. Nazwa węzła, do którego odnosi się czas wykonywania.

ID węzła. Niepowtarzalny identyfikator węzła, do którego odnosi się czas wykonywania. Identyfikator jest tworzony przez system podczas tworzenia węzła.

Czasy wykonania. Czas w sekundach wymagany do wykonania tego węzła. Należy zwrócić uwagę, że czas wykonania może różnić się od podanego w ogólnym komunikacie, ponieważ podczas wykonywania strumienia pewien czas zajmuje również przygotowanie danych i pobranie wyników. Tego dodatkowego czasu nie można obliczyć.

Konfigurowanie parametrów strumienia i sesji

Parametry można definiować w celu wykorzystania ich w wyrażeniach CLEM oraz w skryptach. Stanowią one w efekcie zdefiniowane przez użytkownika zmienne, zapisane i utrwalone w bieżącym strumieniu, sesji lub superwęźle, i można uzyskać do nich dostęp za pośrednictwem interfejsu użytkownika, a także skryptów. W przypadku zapisania strumienia wszelkie parametry ustawione dla tego strumienia również są zapisywane. (To odróżnia je od lokalnych zmiennych skryptu, które mogą być używane tylko w skrypcie, w którym zostały zadeklarowane). Parametry są często używane w skryptach do sterowania zachowaniem skryptów dzięki dostarczaniu informacji na temat zmiennych i wartości niewymagających ustawienia na stałe w skrypcie.

Zasięg parametru zależy od tego, gdzie jest on ustawiany:

- Parametry strumienia można ustawić w skrypcie strumienia lub we właściwościach strumienia, są one również dostępne we wszystkich węzłach w strumieniu. Są one wyświetlane na liście Parametry Konstruktora wyrażeń.
- Parametry sesji można ustawić w autonomicznym skrypcie lub w oknie dialogowym parametrów sesji. Są one dostępne we wszystkich strumieniach używanych w bieżącej sesji (we wszystkich węzłach wymienionych na karcie Strumienie w panelu Zarządzanie).

Parametry można także ustawić dla superwęzłów — w takim przypadku są one widoczne tylko dla węzłów opakowanych w ramach tego superwęzła.

Konfigurowanie parametrów strumienia i sesji za pomocą interfejsu użytkownika

1. Aby określić parametry strumienia, w menu głównym kliknij:

   Narzędzia > **Właściwości strumienia** > **Parametry**
2. Aby ustawić parametry sesji, w menu Narzędzia wybierz pozycję **Parametry sesji**.

Pytanie. Zaznacz to pole wyboru, aby podczas każdego wykonywania był wyświetlany monit o wprowadzenie wartości tego parametru.

Nazwa. Nazwy parametrów wymieniono tutaj. Nowy parametr można utworzyć, wprowadzając nazwę w tym polu. Na przykład, aby utworzyć parametr dla temperatury minimalnej, można wpisać: `minvalue`. Nie należy uwzględniać prefiksu $P-, który określa parametr w wyrażeniach CLEM. Nazwa ta jest także używana do wyświetlania w Konstruktorze wyrażeń CLEM.

Długa nazwa. Przedstawia nazwy opisowe każdego utworzonego parametru.

Składowanie. Umożliwia wybór typu składowania z listy. Składowanie wskazuje, w jaki sposób wartości danych są składowane w parametrze. Na przykład w przypadku pracy z wartościami zawierającymi wiodące zera, które użytkownik chce zachować (takie jak 008), należy jako typ składowania wybrać **Łańcuch**. W przeciwnym wypadku zera zostaną odrzucone. Dostępne typy składowania to: łańcuch, liczba całkowita, liczba rzeczywista, czas, data oraz znacznik czasu. W przypadku parametrów typu data należy zwrócić uwagę, że wartości należy określić, korzystając z zapisu standardowego ISO zgodnie z informacją w następnym akapicie.

Value. Zawiera listę bieżących wartości dla każdego parametru. Parametr należy dostosować odpowiednio do potrzeb. Należy zwrócić uwagę, że w przypadku parametrów typu data wartości muszą być zapisane w notacji określonej normą ISO, to jest: `RRRR-MM-DD`. Daty w innych formatach nie będą akceptowane.

Typ (opcja). Jeśli planuje się wdrożenie strumienia do aplikacji zewnętrznej, należy wybrać poziom pomiaru z listy. W przeciwnym wypadku zaleca się pozostawienie kolumny *Typ* bez zmian. W przypadku określenia ograniczeń wartości dla parametru, takich jak dolna i górna granica przedziału liczbowego, należy wybrać z listy opcję **Określ**.

Należy pamiętać, że opcje długiej nazwy, składowania i typu parametru można ustawić tylko za pośrednictwem interfejsu użytkownika. Opcji tych nie można ustawić za pomocą skryptów.

Kliknij strzałki po prawej stronie, aby przenieść wybrany parametr w górę lub w dół listy dostępnych parametrów. Użyj przycisku usuwania (oznaczonego symbolem *X*) w celu usunięcia wybranego parametru.

Określanie monitów wykonywania w przypadku wartości parametrów

Jeśli użytkownik dysponuje strumieniami wymagającymi wprowadzania różnych wartości w przypadku tego samego parametru, zależnie od okoliczności, można skonfigurować monity o jeden lub wiele parametrów strumienia bądź sesji. Monity te będą wyświetlane w trakcie wykonywania strumienia.

Parametry. (Opcjonalnie) Wprowadź wartość parametru lub pozostaw wartość domyślną, jeśli jest ona dostępna.

Wyłącz te monity. Zaznacz to pole wyboru, aby monity nie były wyświetlane po wykonaniu strumienia. Wyświetlanie monitów można włączyć ponownie, zaznaczając pole wyboru **Pytanie** dostępne w

oknie dialogowym właściwości strumienia lub sesji (zależnie od położenia, w którym zdefiniowano te parametry). Więcej informacji można znaleźć w temacie “Konfigurowanie parametrów strumienia i sesji” na stronie 49.

Określanie ograniczeń wartości w przypadku typu parametru

Ograniczenia wartości parametru można udostępnić podczas wdrożenia strumienia w aplikacji zewnętrznej odczytującej strumienie modelowania danych. Za pomocą tego okna dialogowego można określić wartości dostępne w aplikacji zewnętrznej wykonującej strumień. Zależnie od typu danych ograniczenia wartości są zmieniane dynamicznie w oknie dialogowym. Widoczne tutaj opcje są zgodne z opcjami dostępnymi dla wartości z poziomu węzła Typ.

Typ. Wyświetla aktualnie wybrany poziom pomiaru. Tę wartość można zmienić w celu odzwierciedlania zamierzonego sposobu użycia parametru w oprogramowaniu IBM SPSS Modeler.

Składowanie. Wyświetla typ składowania, o ile jest on znany. Na typy składowania nie wpływa poziom pomiaru (ciągły, nominalny lub flaga) wybrany z przeznaczeniem do wykorzystywania w IBM SPSS Modeler. Użytkownik może zmienić typ składowania na karcie głównej parametrów.

Dolna część okna dialogowego podlega dynamicznym zmianom zależnie od poziomu pomiaru wybranego w zmiennej **Typ**.

Ciągłe poziomy pomiaru

Dolne. Określ ograniczenie dolne wartości parametru.

Górne. Określ ograniczenie górne wartości parametru.

Etykiety. Użytkownik może określić etykiety dowolnej wartości zmiennej przedziału. Kliknij przycisk **Etykiety**, aby otworzyć oddzielne okno dialogowe umożliwiające określenie etykiet wartości.

Nominalne poziomy pomiaru

Wartości. Za pomocą tej opcji można określić wartości parametru, które będą używane jako zmienna nominalna. Wartości nie zostaną wymuszone w strumieniu IBM SPSS Modeler, ale zostaną użyte w menu rozwijanym w przypadku zewnętrznych aplikacji. Używając przycisków strzałek i usuwania, można modyfikować istniejące wartości oraz zmieniać ich kolejność i usuwać je.

Poziomy pomiaru - flaga

Prawda. Określ wartość flagi dla parametru, gdy warunek zostanie spełniony.

Fałsz. Określ wartość flagi dla parametru, gdy warunek nie zostanie spełniony.

Etykiety. Użytkownik może określić etykiety wartości zmiennej typu flaga.

Opcje wdrażania strumieni

Karta Wdrożenie dostępna w oknie dialogowym właściwości strumienia umożliwia określenie opcji wdrożenia strumienia w oprogramowaniu IBM SPSS Collaboration and Deployment Services. Następnie można wykorzystać go do odświeżania modelu lub automatyzowania harmonogramu zadań. Wszystkie strumienie muszą mieć wyznaczoną gałąź oceniającą, zanim będzie możliwe ich wdrożenie. Więcej informacji można znaleźć w temacie “Przechowywanie i wdrażanie obiektów repozytorium” na stronie 204.

Wykonywanie strumieni w pętli

Za pomocą karty Wykonywanie dostępnej w oknie dialogowym właściwości strumienia można skonfigurować warunki działania w pętli, co pozwala na zautomatyzowanie powtarzalnych zadań w bieżącym strumieniu.

Po określeniu tych warunków można ich używać jako podstawy do utworzenia skryptów, ponieważ korzystanie z tej funkcji powoduje wypełnienie okna skryptów, które następnie można modyfikować, tworząc udoskonalone skrypty. Więcej informacji można znaleźć w temacie “Funkcje globalne” na stronie 197.

Konfigurowanie wykonywania strumienia w pętli

1. W menu Plik kliknij pozycję **Właściwości strumienia** (lub wybierz strumień na karcie Strumienie w panelu zarządzania, klikając prawym przyciskiem myszy i wybierając z menu podręcznego pozycję **Właściwości strumienia**).
2. Kliknij kartę **Wykonywanie**.
3. Wybierz tryb wykonywania **Wykonanie w pętli/warunkowe**.
4. Kliknij kartę **W pętli**.

W menu Narzędzia można również kliknąć:

Właściwości strumienia > **Wykonywanie**

Ponadto można też kliknąć węzeł prawym przyciskiem myszy i z menu podręcznego wybrać:

Wykonanie w pętli/warunkowe > **Edytuj ustawienia pętli**

Iteracja. Tego numeru wiersza nie można edytować, ale można dodawać, usuwać lub przenosić iteracje (w górę lub w dół), korzystając z przycisków dostępnych po prawej stronie tabeli.

Nagłówki tabli. Odzwierciedlają one klucz iteracji i zmienne utworzone podczas konfiguracji pętli.

Wyświetlanie wartości globalnych strumienia

Za pomocą karty Globalne dostępnej w oknie dialogowym właściwości strumienia można wyświetlać wartości globalne określone w bieżącym strumieniu. Wartości globalne są tworzone za pomocą węzła ustawień globalnych i służą do określania statystyk, takich jak średnia, suma i odchylenie standardowe wybranych zmiennych.

Po uruchomieniu węzła ustawień globalnych te wartości można wykorzystywać do wielu zastosowań w operacjach wykonywanych na strumieniu. Więcej informacji można znaleźć w temacie “Funkcje globalne” na stronie 197.

Wyświetlanie wartości globalnych strumienia

1. W menu Plik kliknij pozycję **Właściwości strumienia** (lub wybierz strumień na karcie Strumienie w panelu zarządzania, klikając prawym przyciskiem myszy i wybierając z menu podręcznego pozycję **Właściwości strumienia**).
2. Kliknij kartę **Globalne**.

W menu Narzędzia można również kliknąć:

Właściwości strumienia > **Globalne**

Dostępne globalne. Dostępne wartości globalne są zawarte w tej tabeli. Wartości globalnych nie można edytować w tym obszarze, ale można usunąć je wszystkie ze strumienia za pomocą przycisku Wyczyść wszystkie dostępnego po prawej stronie tabeli.

Wyszukiwanie węzłów w strumieniu

Węzły w strumieniu można wyszukiwać, określając kryteria wyszukiwania, np. nazwę węzła, kategorię i identyfikator. Ta funkcja jest niezwykle użyteczna w przypadku złożonych strumieni zawierających wiele węzłów.

Wyszukiwanie węzłów w strumieniu

1. W menu Plik kliknij pozycję **Właściwości strumienia** (lub wybierz strumień na karcie Strumienie w panelu zarządzania, klikając prawym przyciskiem myszy i wybierając z menu podręcznego pozycję **Właściwości strumienia**).
2. Kliknij zakładkę **Wyszukaj**.

W menu Narzędzia można również kliknąć:

Właściwości strumienia > **Wyszukaj**

Użytkownik może określić wiele opcji umożliwiających ograniczenie wyszukiwania. Należy jednak pamiętać, że wyszukiwanie wg identyfikatora węzła (z zastosowaniem zmiennej **ID węzła równe**) uniemożliwia stosowanie innych opcji.

Etykieta węzła zawiera. Zaznacz to pole i wprowadź całą etykietę węzła (lub jej część) w celu wyszukania określonego węzła. Podczas wyszukiwania nie są rozróżniane małe i wielkie litery, a wiele słów jest traktowanych jako pojedynczy ciąg tekstu.

Kategoria węzła. Zaznacz to pole i kliknij kategorię na liście, aby wyszukać określony typ węzła. **Węzeł procesowy** to węzeł dostępny na karcie Rekordy lub Zmienne w palecie węzłów. Opcja **Węzły modeli** dotyczy modelu użytkowego.

Słowa kluczowe zawierają. Zaznacz to pole i wpisz co najmniej jedno pełne słowo kluczowe, aby wyszukać węzły zawierające wpisany tekst w zmiennej Słowa kluczowe na karcie Adnotacje w oknie dialogowym węzła. Wprowadzone słowo kluczowe musi być wyszukane dokładnie w postaci wpisanej przez użytkownika. W celu wyszukania alternatyw wiele słów kluczowych należy oddzielić średnikami (np. wprowadzenie ciągu `proton;neutron` spowoduje wyszukanie wszystkich węzłów zawierających co najmniej jedno z tych słów kluczowych). Więcej informacji można znaleźć w temacie “Adnotacje” na stronie 60.

Adnotacje zawierają. Zaznacz to pole i wpisz co najmniej jedno słowo, aby wyszukać węzły zawierające wpisany tekst w głównym obszarze tekstowym na karcie Adnotacje w oknie dialogowym węzła. Podczas wyszukiwania nie są rozróżniane małe i wielkie litery, a wiele słów jest traktowanych jako pojedynczy ciąg tekstu. Więcej informacji można znaleźć w temacie “Adnotacje” na stronie 60.

Węzeł tworzy zmienną. Zaznacz to pole i wprowadź nazwę utworzonej zmiennej (np.: $C-Drug). Za pomocą tej opcji można wyszukiwać węzły modelowania tworzące określoną zmienną. Należy wprowadzić tylko jedną nazwę zmiennej, która musi stanowić dokładne dopasowanie.

ID węzła równe. Zaznacz to pole i wprowadź id. węzła, aby wyszukać określony węzeł za pomocą tego identyfikatora (wybranie tej opcji powoduje wyłączenie wszystkich wcześniejszych opcji). Identyfikatory węzłów są przypisywane przez system podczas tworzenia węzła i mogą służyć jako odniesienie do węzła w przypadku korzystania ze skryptów lub funkcji automatyzujących pracę. Należy wprowadzić tylko jeden id. węzła, który musi stanowić dokładne dopasowanie. Więcej informacji można znaleźć w temacie “Adnotacje” na stronie 60.

Wyszukaj w Superwęzłach. To pole jest domyślnie zaznaczone, co oznacza, że wyszukiwanie jest wykonywane w węzłach wewnątrz Superwęzłów i poza nimi. Usuń zaznaczenie tego pola wyboru, aby wyszukiwać tylko w węzłach poza Superwęzłami, na szczycie strumienia.

Znajdź. Po określeniu wszystkich żądanych opcji kliknij ten przycisk, aby rozpocząć wyszukiwanie.

Węzły odpowiadające określonym opcjom znajdują się w dolnej części okna dialogowego. Zaznacz węzeł na liście, aby wyróżnić go w obszarze roboczym strumienia.

Zmiana nazw strumieni

Za pomocą karty Adnotacje dostępnej w oknie dialogowym właściwości strumienia można dodawać adnotacje strumieni i wprowadzać niestandardowe nazwy strumieni. Opcje te są przydatne zwłaszcza podczas tworzenia raportów strumieni dodawanych w panelu projektu. Więcej informacji można znaleźć w temacie “Adnotacje” na stronie 60.

Opisy strumieni

Dla każdego utworzonego strumienia oprogramowanie IBM SPSS Modeler tworzy opis zawierający informacje dotyczące jego zawartości. Opisy są przydatne, gdy użytkownik chce dowiedzieć się, co robi strumień, ale nie ma zainstalowanego programu IBM SPSS Modeler, na przykład gdy uzyskuje dostęp do strumienia za pośrednictwem IBM SPSS Collaboration and Deployment Services.

Opis strumienia jest wyświetlany jako dokument HTML składający się z wielu sekcji.

Ogólne informacje o strumieniu

Ta sekcja zawiera nazwę strumienia oraz dane dotyczące daty utworzenia i ostatniego zapisu strumienia.

Opis i komentarze

Ta sekcja zawiera:

- Adnotacje strumienia (patrz: "Adnotacje" na stronie 60)
- Komentarze niezwiązane z konkretnymi węzłami
- Komentarze związane z węzłami w gałęzi modelowania i oceniania w strumieniu

Informacje o ocenie

Ta sekcja zawiera informacje dotyczące gałęzi oceniania strumienia uszeregowane wg nagłówków.

- **Komentarze.** Zawiera komentarze powiązane tylko z węzłami w gałęzi oceniania.
- **Zmienne wejściowe.** Lista zmiennych wejściowych wraz z typami składowania (np. łańcuch, liczba całkowita, liczba rzeczywista itd.).
- **Zmienne wyjściowe.** Wyświetla listę zmiennych wyjściowych wraz z dodatkowymi zmiennymi (utworzonymi przez węzeł modelowania) oraz typami składowania.
- **Parametry.** Wyświetla listę parametrów dotyczących gałęzi oceniania strumienia, które można przeglądać i edytować podczas oceniania modelu. Identyfikacja parametrów odbywa się po kliknięciu przycisku **Parametry oceniania** dostępnego na karcie **Wdrożenie** w oknie dialogowym właściwości strumienia.
- **Węzeł modelu.** Wyświetla nazwę i typ modelu (np. Sieć neuronowa, drzewo C&R itd.). Jest to model użytkowy wybrany dla pola **Węzeł modelu** na karcie **Wdrożenie** w oknie dialogowym właściwości strumienia.
- **Szczegóły modelu.** Przedstawia szczegóły modelu użytkowego zidentyfikowanego w poprzednim nagłówku. Jeśli to tylko możliwe, uwzględniane są wykresy ewaluacyjne i ważności predyktorów modelu.

Informacje o modelowaniu

Zawiera informacje dotyczące gałęzi modelowania strumienia.

- **Komentarze.** Wyświetla listę komentarzy lub adnotacji połączonych z węzłami w gałęzi modelowania.
- **Zmienne wejściowe.** Wyświetla listę zmiennych wejściowych wraz z ich rolami w gałęzi modelowania (w postaci wartości roli zmiennej, np. Wejściowa, Przewidywana, Rozdzielona itd.).
- **Parametry.** Wyświetla listę parametrów dotyczących gałęzi modelowania strumienia, które można przeglądać i edytować podczas aktualizacji modelu. Identyfikacja parametrów odbywa się po kliknięciu przycisku **Parametry modelu** dostępnego na karcie **Wdrożenie** w oknie dialogowym właściwości strumienia.
- **Węzeł modelowania.** Wyświetla nazwę i typ węzła modelowania użytego do utworzenia lub zaktualizowania modelu.

Podgląd opisu strumienia

Użytkownik może wyświetlić treść opisu strumienia w przeglądarce internetowej, klikając opcję w oknie dialogowym właściwości strumienia. Treść opisu jest uzależniona od opcji określonych na karcie Wdrożenie w oknie dialogowym. Więcej informacji można znaleźć w temacie "Opcje wdrażania strumieni" na stronie 216.

Aby wyświetlić opis strumienia:

1. W głównym menu IBM SPSS Modeler kliknij:

   Narzędzia > **Właściwości strumienia** > **Wdrożenie**

2. Ustaw typ wdrożenia, określony węzeł oceniający i parametry oceniania.
3. Jeśli typem wdrożenia jest Odświeżenie modelu, opcjonalnie można wybrać:
   - Węzeł modelowania i dowolne parametry budowania modelu
   - Model użytkowy na gałęzi oceniania strumienia
4. Kliknij przycisk **Podgląd opisu strumienia**.

Eksportowanie opisu strumienia

Treść opisu strumienia można wyeksportować do pliku HTML.

Aby wyeksportować opis strumienia:

1. W menu głównym kliknij opcje:

   Plik > **Eksportuj opis strumienia**
2. Wprowadź nazwę pliku HTML i kliknij przycisk **Zapisz**.

Uruchamianie strumieni

Po skonfigurowaniu opcji węzłów i połączeniu żądanych węzłów można uruchomić strumień, przeprowadzając dane przez węzły w strumieniu. Strumień w oprogramowaniu IBM SPSS Modeler można uruchomić na kilka sposobów. Można:

- Kliknąć przycisk **Wykonaj** dostępny w menu Narzędzia.
- Kliknąć jeden z przycisków **Wykonaj...** dostępnych na pasku narzędzi. Za pomocą tych przycisków można uruchomić cały strumień lub wybrany węzeł końcowy. Więcej informacji można znaleźć w temacie “ Pasek narzędzi programu IBM SPSS Modeler” na stronie 18.
- Uruchomić pojedynczy strumień danych, klikając prawym przyciskiem myszy węzeł końcowy i w menu podręcznym wybierając pozycję **Wykonaj**.
- Uruchomić część strumienia danych, klikając prawym przyciskiem myszy węzeł inny niż końcowy i w menu podręcznym wybierając pozycję **Wykonaj stąd**. Wykonanie tej operacji spowoduje uruchomienie działań następujących po wybranym węźle.

Aby zatrzymać wykonywanie operacji trwających w strumieniu, na pasku narzędzi kliknij czerwony przycisk **Stop** na pasku narzędzi lub w menu Narzędzia wybierz pozycję **Zatrzymaj wykonywanie**. Jednokrotne naciśnięcie przycisku Stop nakazuje programowi Modeler zatrzymanie przetwarzania na serwerze Modeler Server. W niektórych przypadkach wykonywanie zostanie zatrzymane natychmiast, ale niekiedy konieczne jest ukończenie bieżącego kroku, zanim będzie można zatrzymać całe wykonanie. Dlatego czas potrzebny na zatrzymanie będzie różny. Dwukrotne kliknięcie przycisku spowoduje przerwanie połączenia z serwerem i nawiązanie nowego połączenia. W większości przypadków spowoduje to zamknięcie wszystkich procesów serwera (i może potrwać pewien czas). Jednak w pewnych przypadkach niektóre procesy serwera nie zostaną zatrzymane.

Jeśli wykonywanie dowolnego strumienia trwa dłużej niż trzy sekundy, wyświetlane jest okno dialogowe Informacje o stanie procesu zawierające informacje o postępie operacji.

W przypadku niektórych węzłów dostępne są inne pozycje zapewniające dodatkowe informacje o wykonywaniu strumienia. Aby wyświetlić te informacje, wybierz odpowiedni wiersz w oknie dialogowym. Pierwszy wiersz jest wybrany automatycznie.

Praca z modelami

Jeśli strumień zawiera węzeł modelowania (węzeł z karty Modelowanie lub Modelowanie w bazie danych na palecie węzłów), po uruchomieniu strumienia tworzony jest **model użytkowy**. Model użytkowy to kontener na **model**, czyli zestaw reguł, formuł lub równań umożliwiających generowanie predykcji względem źródła danych. Jest to istota analizy predykcyjnej.

Rysunek 12. Model użytkowy

Po uruchomieniu węzła modelowania na obszarze roboczym strumienia umieszczany jest odpowiedni model użytkowy. Jest on oznaczany za pomocą złotej ikony w kształcie rombu. Model użytkowy można otworzyć i przejrzeć jego zawartość, aby zapoznać się ze szczegółami modelu. Aby wyświetlić predykcje, należy dołączyć i uruchomić co najmniej jednej węzeł końcowy. Raporty wynikowe pochodzące z tego węzła zawierają predykcje zapisane w postaci, którą można odczytać.

Typowy strumień modelowania składa się z dwóch gałęzi. **Gałąź modelowania** zawiera węzeł modelowania oraz poprzedzające go węzły źródłowe i węzły przetwarzania. **Gałąź oceny** jest tworzona po uruchomieniu węzła modelowania. Zawiera ona model użytkowy i węzły końcowe służące do wyświetlania predykcji.

Dodatkowe informacje zawarto w dokumentacji *IBM SPSS Modeler - węzły modelowania*.

Dodawanie komentarzy i adnotacji do węzłów i strumieni

Konieczne może być utworzenie opisu strumienia dla innych użytkowników w organizacji. Można tym samym dołączyć do strumieni, węzłów i modeli użytkowych komentarze z objaśnieniem.

Inni użytkownicy będą mogli zobaczyć te komentarze ekranowe lub można wydrukować obraz strumienia, który będzie zawierał komentarze.

Można wyświetlić wszystkie komentarze dla strumienia lub superwęzła, zmieniać kolejność komentarzy na liście, edytować treść komentarza oraz zmieniać kolor na pierwszym planie lub w tle komentarza. Więcej informacji można znaleźć w temacie “Wyświetlanie listy komentarzy w strumieniu” na stronie 59.

Do strumieni, węzłów i modeli użytkowych można również dodawać uwagi w postaci adnotacji tekstowych, korzystając w tym celu z karty Adnotacje dostępnej w oknie dialogowym właściwości strumienia, w oknie dialogowym węzła lub w oknie modelu użytkowego. Uwagi są widoczne tylko po otwarciu karty Adnotacje; nie dotyczy to adnotacji dla strumienia, które mogą być również wyświetlane jako komentarze ekranowe. Więcej informacji można znaleźć w temacie “Adnotacje” na stronie 60.

Komentarze

Komentarze mają postać pól tekstowych, w których można wprowadzać dowolnie długi tekst. Możliwe jest dodawanie dowolnie wielu komentarzy. Komentarz może być autonomiczny (niedołączony do żadnego obiektu strumienia) lub powiązany z jednym bądź wieloma węzłami lub modelami użytkowymi w strumieniu. Komentarze autonomiczne używane są zazwyczaj do opisywania ogólnego przeznaczenia strumienia; komentarze powiązane opisują węzły lub modele użytkowe, do których są dołączone. Do jednego węzła lub modelu użytkowego może być dołączony więcej niż jeden komentarz, a strumień może mieć dowolnie wiele komentarzy autonomicznych.

Uwaga: Również adnotacje do strumienia można uwidocznić jako komentarze ekranowe, jednak nie mogą one być dołączane do węzłów ani modeli użytkowych. Więcej informacji można znaleźć w temacie “Konwertowanie adnotacji na komentarze” na stronie 60.

Wygląd pola tekstowego odzwierciedla bieżący tryb komentarza (lub adnotacji wyświetlanej jako komentarz), zgodnie z poniższą tabelą.

Tabela 3. Tryby pola tekstowego komentarzy i adnotacji

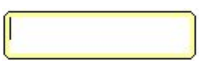
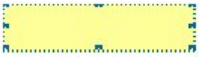
Pole tekstowe komentarza	Pole tekstowe adnotacji	Tryb	Oznacza	Jak uzyskać
		Edycja	Komentarz jest otwarty do edycji.	Utwórz nowy komentarz lub adnotację albo kliknij dwukrotnie istniejący komentarz lub adnotację.

Tabela 3. Tryby pola tekstowego komentarzy i adnotacji (kontynuacja)				
Pole tekstowe komentarza	**Pole tekstowe adnotacji**	**Tryb**	**Oznacza**	**Jak uzyskać**
		Ostatnio wybrany	Komentarz można przesunąć lub usunąć, można też zmienić jego rozmiar.	Kliknij tło strumienia po edycji lub jeden raz kliknij istniejący komentarz lub adnotację.
		Widok	Edycja jest zakończona.	Po edycji kliknij inny węzeł, komentarz lub adnotację.

Nowo utworzony komentarz autonomiczny jest początkowo wyświetlany w lewym górnym rogu obszaru roboczego strumienia.

Jeśli dołączasz komentarz do węzła lub modelu użytkowego, komentarz ten jest początkowo wyświetlany nad obiektem strumienia, do którego jest dołączony.

Pole tekstowe ma kolor biały, co oznacza, że możliwe jest wprowadzanie tekstu. Po wprowadzeniu tekstu kliknij miejsce poza polem tekstowym. Tło komentarza przyjmie kolor żółty, co sygnalizuje zakończenie wprowadzania tekstu. Komentarz pozostaje zaznaczony i można go przesuwać, usunąć lub zmienić jego rozmiar.

Po ponownym kliknięciu ramka przyjmuje postać linii ciągłych, co sygnalizuje, że edytowanie zostało zakończone.

Dwukrotne kliknięcie komentarza powoduje przełączenie pola tekstowego w tryb edycji — tło przyjmuje kolor biały i można edytować tekst komentarza.

Komentarze można także dołączać do Superwęzłów.

Operacje uwzględniające komentarze

Użytkownik może wykonać wiele operacji na komentarzach. Można:

- Dodać swobodny komentarz
- Dołączyć komentarz do węzła lub modelu użytkowego
- Edytować komentarz
- Zmienić rozmiar komentarza
- Przenieść komentarz
- Odłączyć komentarz
- Usunąć komentarz
- Pokazać lub ukryć wszystkie komentarze w strumieniu

Dodawanie swobodnego komentarza

1. Upewnij się, że w strumieniu nie jest wybrany żaden element.
2. Wykonaj jedną z czynności:
   - W menu głównym kliknij opcje:

     Wstaw > **Nowy komentarz**
   - Prawym przyciskiem myszy kliknij tło strumienia i z menu podręcznego wybierz pozycję **Nowy komentarz**.
   - Kliknij przycisk **Nowy komentarz** dostępny na pasku narzędzi.
3. Wpisz tekst komentarza (lub wklej tekst ze schowka).
4. Kliknij węzeł w strumieniu, aby zapisać komentarz.

Dołączanie komentarza do węzła lub modelu użytkowego

1. Wybierz co najmniej jeden węzeł lub model użytkowy w obszarze roboczym strumienia.
2. Wykonaj jedną z czynności:
   - W menu głównym kliknij opcje:

     Wstaw > **Nowy komentarz**
   - Prawym przyciskiem myszy kliknij tło strumienia i z menu podręcznego wybierz pozycję **Nowy komentarz**.
   - Kliknij przycisk **Nowy komentarz** dostępny na pasku narzędzi.
3. Wpisz tekst komentarza.
4. Kliknij inny węzeł w strumieniu, aby zapisać komentarz.

   Można również:
5. Wstawić swobodny komentarz (patrz punkt poprzedni).
6. Wykonaj jedną z czynności:

- Wybierz komentarz, naciśnij klawisz F2, a następnie wybierz węzeł lub model użytkowy.
- Wybierz węzeł lub model użytkowy, naciśnij klawisz F2, a następnie wybierz komentarz.
- (Tylko przy użyciu myszy wyposażonej w trzy przyciski) Ustaw wskaźnik myszy na komentarzu, przytrzymaj środkowy przycisk, przeciągnij wskaźnik na węzeł lub model użytkowy i zwolnij przycisk myszy.

Dołączanie komentarza do dodatkowego węzła lub modelu użytkowego

Jeśli komentarz jest już dołączony do węzła lub modelu użytkowego lub jeśli znajduje się on aktualnie na poziomie strumienia, a użytkownik chce dołączyć go do dodatkowego węzła lub modelu użytkowego, można wykonać następujące operacje:

- Wybierz komentarz, naciśnij klawisz F2, a następnie wybierz węzeł lub model użytkowy.
- Wybierz węzeł lub model użytkowy, naciśnij klawisz F2, a następnie wybierz komentarz.
- (Tylko przy użyciu myszy wyposażonej w trzy przyciski) Ustaw wskaźnik myszy na komentarzu, przytrzymaj środkowy przycisk, przeciągnij wskaźnik na węzeł lub model użytkowy i zwolnij przycisk myszy.

Edytowanie istniejącego komentarza

1. Wykonaj jedną z czynności:
   - Kliknij dwukrotnie pole tekstowe komentarza.
   - Zaznacz pole tekstowe i naciśnij klawisz Enter.
   - Kliknij prawym przyciskiem myszy pole tekstowe, aby wyświetlić jego menu, a następnie kliknij pozycję Edytuj.
2. Zmodyfikuj tekst komentarza. Podczas edycji można stosować standardowe skróty klawiaturowe systemu Windows, np. Ctrl+C, aby skopiować tekst. Inne opcje dostępne podczas edycji zawarto na liście w menu podręcznym komentarza.
3. Kliknij poza polem tekstowym, aby wyświetlić elementy sterujące rozmiarem, a następnie kliknij ponownie w celu zakończenia edycji komentarza.

Zmiana rozmiaru pola tekstowego komentarza

1. Wybierz komentarz, aby wyświetlić elementy sterujące rozmiarem.
2. Kliknij i przeciągnij element sterujący, aby zmienić rozmiar pola.
3. Kliknij poza polem tekstowym, aby zapisać zmiany.

Przenoszenie istniejącego komentarza

Aby przenieść komentarz bez dołączonych do niego obiektów (o ile istnieją), wykonaj jedną z operacji:

- Ustaw wskaźnik myszy na komentarzu, przytrzymaj lewy przycisk i przeciągnij komentarz do nowego położenia.
- Zaznacz komentarz, przytrzymaj klawisz Alt i przenieś komentarz za pomocą klawiszy strzałek.

Aby przenieść komentarz wraz z węzłami lub modelami użytkowymi, do których komentarz jest dołączony:

1. Zaznacz wszystkie obiekty, które chcesz przenieść.
2. Wykonaj jedną z czynności:

- Ustaw wskaźnik myszy na obiektach, przytrzymaj lewy przycisk i przeciągnij obiekty do nowego położenia.
- Zaznacz jeden z obiektów, przytrzymaj klawisz Alt i przenieś obiekty za pomocą klawiszy strzałek.

Odłączanie komentarza od węzła lub modelu użytkowego

1. Wybierz co najmniej jeden komentarz do odłączenia.
2. Wykonaj jedną z czynności:

- Naciśnij klawisz F3.
- Prawym przyciskiem myszy kliknij zaznaczony komentarz i w menu podręcznym wybierz pozycję Odłącz.

Usuwanie komentarza

1. Wybierz co najmniej jeden komentarz do usunięcia.
2. Wykonaj jedną z czynności:

- Naciśnij klawisz Delete.
- Prawym przyciskiem myszy kliknij zaznaczony komentarz i w menu podręcznym wybierz pozycję Usuń.

Jeśli komentarz był połączony z węzłem lub modelem użytkowym, linia połączenia również zostanie usunięta.

Jeśli pierwotnie komentarz był adnotacją strumienia lub Superwęzła przekształconą na komentarz swobodny, to zostanie on usunięty z obszaru roboczego, ale jego tekst zostanie zachowany na karcie Adnotacje strumienia lub Superwęzła.

Pokazywanie lub ukrywanie wszystkich komentarzy w strumieniu

1. Wykonaj jedną z czynności:

- W menu głównym kliknij opcje:

  Widok > **Komentarze**
- Kliknij przycisk **Pokaż/Ukryj komentarze** dostępny na pasku narzędzi.

Wyświetlanie listy komentarzy w strumieniu

Istnieje możliwość wyświetlenia wszystkich komentarzy opracowanych dla określonego strumienia lub Superwęzła.

Na tej liście można:

- zmienić kolejność komentarzy,
- edytować treść komentarza,
- zmienić kolor tła i pierwszego planu komentarza.

Lista komentarzy

Aby wyświetlić listę komentarzy utworzonych dla strumienia, wykonaj jedną z następujących procedur.

- W menu głównym kliknij opcje:

  Narzędzia > **Właściwości strumienia** > **Komentarze**

- Kliknij prawym przyciskiem myszy strumień w panelu Zarządzanie, a następnie kliknij opcję **Właściwości strumienia** i **Komentarze**.
- Kliknij prawym przyciskiem myszy tło strumienia w obszarze roboczym i kliknij opcję **Właściwości strumienia**, a następnie **Komentarze**.

Tekst. Tekst komentarza. Kliknij dwukrotnie tekst, aby zmienić pole w edytowalne pole tekstowe.

Łącza. Nazwa węzła, do którego dołączony jest komentarz. Jeśli ta zmienna jest pusta, komentarz dotyczy strumienia.

Przyciski pozycji. Umożliwiają one przemieszczanie wybranego komentarza w górę lub w dół na liście.

Kolory komentarzy. Aby zmienić kolor pierwszego planu albo kolor tła komentarza, wybierz komentarz, wybierz pole wyboru **Kolory niestandardowe**, a następnie wybierz kolor z listy **Tło** lub **Pierwszy plan** (lub z obu). Kliknij przycisk **Zastosuj**, a następnie kliknij tło strumienia, aby zobaczyć efekt zmiany. Kliknij przycisk **OK**, aby zapisać zmiany.

Konwertowanie adnotacji na komentarze

Adnotacje wprowadzone do strumieni lub superwęzłów mogą zostać przekształcone w komentarze.

W przypadku strumieni adnotacja jest przekształcana w komentarz niezależny (to jest, niepowiązany z żadnym z węzłów) w obszarze roboczym strumienia.

W przypadku przekształcania adnotacji superwęzła w komentarz ten nie jest dołączany do superwęzła w obszarze roboczym strumienia, lecz jest widoczny po powiększeniu superwęzła.

Aby przekształcić adnotację strumienia w komentarz

1. Kliknij opcję **Właściwości strumienia** w menu Narzędzia. (Można też kliknąć strumień prawym przyciskiem myszy w panelu Zarządzanie, a następnie kliknąć opcję **Właściwości strumienia**.)
2. Kliknij kartę **Adnotacje**.
3. Zaznacz pole wyboru **Pokaż adnotację jako komentarz**.
4. Kliknij przycisk **OK**.

Aby przekształcić adnotację superwęzła w komentarz

1. Kliknij dwukrotnie ikonę superwęzła w obszarze roboczym.
2. Kliknij kartę **Adnotacje**.
3. Zaznacz pole wyboru **Pokaż adnotację jako komentarz**.
4. Kliknij przycisk **OK**.

Adnotacje

Do węzłów, strumieni i modeli można dodawać adnotacje na szereg sposobów. Można dodawać adnotacje opisowe i nadawać im nazwy niestandardowe. Opcje te są przydatne zwłaszcza podczas tworzenia raportów strumieni dodawanych w panelu projektu. W przypadku węzłów i modeli użytkowych można także dodawać treść podpowiedzi, co pomoże w odróżnieniu podobnych węzłów w obszarze roboczym strumienia.

Dodawanie adnotacji

Przystąpienie do edycji węzła lub modelu użytkowego powoduje otwarcie okna dialogowego z kartami, zawierającego kartę Adnotacje służącą do ustawiania szeregu opcji adnotacji. Można także bezpośrednio otworzyć kartę Adnotacje.

1. Aby dodać adnotację do węzła lub modelu użytkowego, kliknij prawym przyciskiem myszy węzeł lub model użytkowy w obszarze roboczym strumienia, a następnie kliknij przycisk **Zmień nazwę i skomentuj**. Zostanie otwarte okno dialogowe edycji z widoczną kartą Adnotacje.

2. Aby dodawać adnotacje do strumienia, kliknij opcję **Właściwości strumienia** w menu Narzędzia. (Można też kliknąć strumień prawym przyciskiem myszy w panelu Zarządzanie, a następnie kliknąć opcję **Właściwości strumienia**.) Kliknij kartę Adnotacje.

Nazwa. Wybierz opcję **Niestandardowa**, aby zmodyfikować automatycznie wygenerowaną nazwę lub utworzyć unikalną nazwę dla węzła, która ma być wyświetlana w obszarze roboczym strumienia.

Tekst podpowiedzi. (Tylko dla węzłów i modeli użytkowych) Wprowadź tekst, który będzie używany jako podpowiedź w obszarze roboczym strumienia. Podpowiedzi są szczególnie przydatne w przypadku pracy z dużą liczbą podobnych węzłów.

Słowa kluczowe. Określ słowa kluczowe, które mają być używane w raportach projektu i przy wyszukiwaniu węzłów w strumieniu, a także do oznaczania obiektów składowanych w repozytorium (patrz “Informacje o programie IBM SPSS Collaboration and Deployment Services Repository ” na stronie 203). Można określić więcej niż jedno słowo kluczowe, rozdzielając słowa średnikami, na przykład: `income; crop type; claim value`. Białe znaki na początku i końcu słów kluczowych są obcinane, na przykład zapis `income ; crop type` jest równoważny `income;crop type`. (Białe spacje wewnątrz słów kluczowych nie są jednak usuwane. Na przykład, zapisy `crop type` z jedną spacją i `crop type` z dwiema spacjami różnią się od siebie).

W głównym obszarze tekstowym można wprowadzać dłuższe adnotacje dotyczące działania węzła lub decyzji podejmowanych w węźle. Na przykład, gdy udostępniamy strumienie i wykorzystujemy je w różnych projektach, warto zanotować pewne decyzje, np. odrzucenie zmiennej z wieloma wartościami pustymi przy użyciu węzła filtrowania. Adnotacja dodana do węzła jest składowana razem z nim. Można również umieścić adnotacje w raporcie z projektu utworzonym w panelu projektu. Więcej informacji można znaleźć w temacie “Wprowadzenie do projektów” na stronie 227.

Pokaż adnotację jako komentarz. (Tylko w przypadku adnotacji strumieni i Superwęzła) Zaznaczenie tego pola powoduje przekształcenie adnotacji w autonomiczny komentarz, który będzie widoczny w obszarze roboczym strumienia. Więcej informacji można znaleźć w temacie “Dodawanie komentarzy i adnotacji do węzłów i strumieni” na stronie 56.

Identyfikator. Unikalny identyfikator, za pośrednictwem którego można odwoływać się do węzła w skryptach lub mechanizmach automatyzacji. Ta wartość jest generowana automatycznie w momencie tworzenia węzła i później nie ulega zmianie. Należy także zwrócić uwagę, że w identyfikatorach węzłów nie są używane zera, aby nie były omyłkowo interpretowane jako litery „O”. Za pomocą przycisku kopiowania po prawej stronie można skopiować identyfikator do schowka, a potem wklejać go w skryptach lub w innych miejscach, w których będzie potrzebny.

Zapisywanie strumieni danych

Po utworzeniu strumienia można go zapisać z przeznaczeniem do ponownego użycia w przyszłości.

Aby zapisać strumień

1. W menu Plik kliknij pozycję **Zapisz strumień** lub **Zapisz strumień jako**.
2. W oknie dialogowym Zapisz przejdź do folderu, w którym zostanie zapisany plik strumienia.
3. W polu tekstowym Nazwa pliku wprowadź nazwę strumienia.
4. Aby dodać zapisany strumień do bieżącego projektu, wybierz opcję **Dodaj do projektu**.

Kliknięcie przycisku **Zapisz** powoduje zachowanie strumienia z rozszerzeniem `*.str` w określonym katalogu.

Automatyczne pliki kopii zapasowej. Po każdorazowym zapisaniu strumienia poprzednie zapisane wersje pliku są automatycznie zachowywane jako kopia zapasowa. Do ich nazw dodawany jest znak łącznika (przykładowo: `mójstrumień.str-`). Aby przywrócić wersję kopii zapasowej, wystarczy usunąć znak łącznika i ponownie otworzyć plik.

Zapisywanie stanów

Oprócz strumieni można także zapisywać **stany**, które opisują obecnie wyświetlany diagram strumienia i wszelkie utworzone modele użytkowe (wymienione na karcie Modele w panelu zarządzania).

Aby zapisać stan:

1. W menu Plik kliknij:

   Stan > **Zapisz stan lub Zapisz stan jako**
2. W oknie dialogowym Zapisz przejdź do folderu, w którym zostanie zapisany plik stanu.

Kliknięcie przycisku **Zapisz** powoduje zapisanie stanu z rozszerzeniem **.cst* w określonym katalogu.

Zapisywanie węzłów

Pojedynczy węzeł można również zapisać, klikając prawym przyciskiem myszy węzeł w obszarze roboczym strumienia, a następnie klikając opcję **Zapisz węzeł** w menu podręcznym. Rozszerzenie pliku, jakiego należy użyć, to **.nod*.

Zapisywanie wielu obiektów strumieni

Po wyjściu z programu IBM SPSS Modeler z wieloma niezapisanymi obiektami, takimi jak strumienie, projekty lub modele użytkowe użytkownik jest monitowany o ich zapisanie przed całkowitym zamknięciem oprogramowania. W przypadku decyzji o zapisaniu elementów zostanie wyświetlone okno dialogowe z opcjami zapisywania każdego obiektu.

1. Wystarczy po prostu zaznaczyć pola wyboru dla obiektów, które mają zostać zapisane.
2. Kliknij przycisk **OK**, aby zapisać każdy obiekt w żądanej lokalizacji.

Dla każdego obiektu zostanie wyświetlone standardowe okno dialogowe Zapisz. Po zakończeniu zapisywania aplikacja zostanie zamknięta.

Zapisywanie wyniku

Tabele, wykresy i raporty generowane przez węzły wynikowe IBM SPSS Modeler mogą być zapisywane w formacie obiektu wynikowego (**.cou*).

1. Podczas wyświetlania wyniku, który ma zostać zapisany, w menu okna wyników kliknij:

   File > **Save**
2. Podaj nazwę i lokalizację pliku wynikowego.
3. Opcjonalnie wybierz opcję **Dodaj plik do projektu** w oknie dialogowym Zapisz, aby uwzględnić plik w bieżącym projekcie. Więcej informacji można znaleźć w temacie “Wprowadzenie do projektów” na stronie 227.

Można też kliknąć prawym przyciskiem myszy dowolny obiekt wynikowy wymieniony w panelu zarządzania i wybrać z menu podręcznego pozycję **Zapisz**.

Informacje dotyczące szyfrowania i usuwania szyfrowania

Zapisując strumień, węzeł, projekt, plik wyjściowy lub model użytkowy, można zaszyfrować go, aby uniemożliwić jego nieuprawnione użycie. W tym celu można wybrać dodatkową opcję podczas zapisywania i dodać hasło do zapisywanego elementu. To szyfrowanie można ustawić dla dowolnego z zapisywanych elementów, dla których są dodawane dodatkowe zabezpieczenia; jest to szyfrowanie różniące się od szyfrowania SSL używanego w przypadku próby przekazywania plików między produktem IBM SPSS Modeler a produktem IBM SPSS Modeler Server.

Podczas próby otwarcia zaszyfrowanego elementu użytkownik jest monitowany o wprowadzenie hasła. Po poprawnym wprowadzeniu hasła szyfrowanie elementu jest automatycznie usuwane i otwieranie przebiega jak zwykle.

Aby zaszyfrować element

1. W oknie dialogowym Zapisz kliknij **Opcje** dla elementu do zaszyfrowania. Zostanie otwarte okno dialogowe Opcje szyfrowania.
2. Wybierz opcje **Zaszyfruj ten plik**.
3. Opcjonalnie, w celu zapewnienia większego bezpieczeństwa, wybierz opcję **Ukryj hasło**. Powoduje to ukrywanie wprowadzanych znaków i wyświetlanie zamiast nich serii kropek.
4. Wprowadź hasło *Ostrzeżenie*: Jeśli zapomnisz hasła, nie będzie możliwe otwarcie pliku lub modelu.
5. W przypadku zaznaczenia opcji **Ukryj hasło** ponownie wprowadź hasło, aby potwierdzić jego poprawność.
6. Kliknij przycisk **OK**, aby powrócić do okna dialogowego Zapisz.

Uwaga: W przypadku zapisania kopii elementu chronionego szyfrem nowy element jest automatycznie zapisywany w formacie zaszyfrowanym, z użyciem oryginalnego hasła, o ile użytkownik nie zmieni ustawień w oknie dialogowym Opcje szyfrowania.

Ładowanie plików

Użytkownik może ponownie załadować wiele obiektów zapisanych w oprogramowaniu IBM SPSS Modeler:

- Strumienie (.str)
- Stany (.cst)
- Modele (.gm)
- Palety modeli (.gen)
- Węzły (.nod)
- Dane wyjściowe (.cou)
- Projekty (.cpj)

Otwieranie nowych plików

Strumienie można ładować bezpośrednio z menu Plik.

- W menu Plik kliknij pozycję **Otwórz strumień**.

Wszystkie inne typy plików można otworzyć za pomocą elementów podmenu dostępnych w menu Plik. Przykładowo: aby załadować model w menu Plik, kliknij:

Modele > **Otwórz model lub Załaduj paletę modeli**

Otwieranie ostatnio używanych plików

Dzięki opcjom dostępnym w dolnej części menu Plik można szybko ładować ostatnio używane pliki.

Wybierz pozycje **Ostatnie strumienie**, **Ostatnie projekty** lub **Ostatnie stany**, aby rozwinąć listę ostatnio używanych plików.

Mapowanie strumieni danych

Korzystając z narzędzia do mapowania, można podłączyć nowe źródło danych do już istniejącego strumienia. Narzędzie do mapowania pozwala nie tylko na określenie połączenia, lecz także pomaga w określeniu sposobu, w jaki zmienne w nowym źródle zastąpią te w istniejącym strumieniu. Zamiast ponownego tworzenia całego strumienia danych dla nowego źródła danych, można po prostu nawiązać połączenie z istniejącym strumieniem.

Narzędzie do mapowania danych umożliwia połączenie dwu fragmentów strumieni, zapewniając jednocześnie prawidłowe dopasowanie wszystkich (a przynajmniej wszystkich istotnych) z tych nazw zmiennych. W szczególności mapowanie danych skutkuje utworzeniem nowego węzła Filtr, który jest dopasowywany do odpowiednich zmiennych przez zmianę ich nazw.

Istnieją dwa równoważne sposoby mapowania danych:

Wybierz węzeł zamiany. Ta metoda rozpoczyna się na węźle, który ma zostać zamieniony. Najpierw należy kliknąć prawym przyciskiem myszy węzeł do zamiany; następnie, korzystając z opcji **Mapowanie danych** > **Wybierz węzeł zamiany** z menu podręcznego, wybierz węzeł, na który ma zostać dokonana zamiana.

Mapuj do. Ta metoda rozpoczyna się od węzła, który ma zostać wprowadzony do strumienia. Najpierw należy kliknąć prawym przyciskiem myszy węzeł do wprowadzenia; następnie, korzystając z opcji **Mapowanie danych** > **Mapuj na** z menu podręcznego, wybierz węzeł, do którego ma on zostać dołączony. Ta metoda jest szczególnie przydatna w przypadku mapowania na węzeł końcowy. *Uwaga*: Nie jest możliwe mapowanie na węzły Łączenie oraz Dołączanie. Zamiast tego należy po prostu podłączyć strumień do węzła Łączenie w zwykły sposób.

Mapowanie danych jest ściśle powiązane z tworzeniem strumienia. W przypadku próby nawiązania połączenia z węzłem, który ma już połączenie, zostanie wyświetlona opcja zastąpienia połączenia lub zmapowania na ten węzeł.

Mapowanie danych na szablon

Aby zastąpić źródło danych w szablonie strumienia nowym węzłem źródłowym, który wprowadzi właściwe dane do programu IBM SPSS Modeler, należy użyć opcji **Wybierz węzeł zamiany** z menu podręcznego Mapowanie danych. Ta opcja jest dostępna w odniesieniu do wszystkich węzłów z wyjątkiem węzłów łączenia, agregacji i wszystkich węzłów końcowych. Zastosowanie narzędzia do mapowania danych przy wykonywaniu tej czynności gwarantuje prawidłowe zmapowanie zmiennych między operacjami istniejącego strumienia a nowym źródłem danych. Poniżej ogólnie przedstawiono kolejne kroki procesu mapowania danych.

Krok 1: Określenie niezbędnych zmiennych w pierwotnym węźle źródłowym. Aby operacje strumienia były wykonywane prawidłowo, należy wskazać, które zmienne są niezbędne. Więcej informacji można znaleźć w temacie “Określanie niezbędnych zmiennych” na stronie 65.

Krok 2: Dodanie nowego źródła danych do obszaru roboczego strumienia. Korzystając z jednego z węzłów źródłowych, należy wprowadzić nowe dane zastępcze.

Krok 3: Zastąpienie węzła źródłowego w szablonie. Korzystając z opcji Data Mapping w menu podręcznym węzła źródłowego szablonu, kliknij opcję **Wybierz węzeł zamiany**, a następnie wybierz nowy węzeł źródłowy.

Krok 4: Sprawdzenie zmapowanych zmiennych. W oknie dialogowym, które zostanie otwarte, sprawdź czy program prawidłowo mapuje zmienne między nowym źródłem danych a strumieniem. Wszelkie niezbędne zmienne, które nie są zmapowane, są wyświetlane na czerwono. Zmienne te są używane w operacjach strumienia i należy je zastąpić podobnymi zmiennymi z nowego źródła danych, aby dalsze operacje strumienia działały prawidłowo. Więcej informacji można znaleźć w temacie “Analizowanie zmapowanych zmiennych” na stronie 65.

Po zapewnieniu prawidłowego mapowania wszystkich niezbędnych zmiennych w oknie dialogowym stare źródło danych zostaje odłączone, a nowe źródło danych zostaje podłączone do strumienia przy użyciu węzła filtrowania o nazwie *Mapuj*. Ten węzeł filtrowania kieruje faktycznym mapowaniem zmiennych w strumieniu. W obszarze roboczym strumienia znajduje się także węzeł filtrowania *Usuń mapowanie*. Węzeł filtrowania *Usuń mapowanie* może posłużyć do cofnięcia mapowań pól, gdy zostanie dodany do strumienia. Po dodaniu do strumienia węzeł ten cofnie mapowanie pól, jednak konieczne będzie edytowanie wszelkich dalszych węzłów końcowych w celu ponownego wybrania zmiennych i nałożeń.

Mapowanie między strumieniami

Ta metoda mapowania, podobnie jak łączenie węzłów, nie wymaga wcześniejszego określenia zmiennych niezbędnych. W metodzie tej wystarczy wykonać połączenie między jednym strumieniem a drugim, korzystając z opcji **Mapuj do** w menu podręcznym Mapowanie danych. Ten sposób mapowania danych jest przydatny w przypadku mapowania do węzłów końcowych oraz kopiowania i wklejania między strumieniami. *Uwaga*: Korzystając z opcji **Mapuj do** nie można mapować do węzłów łączenia, dołączania i jakichkolwiek węzłów źródłowych.

Aby zmapować dane między strumieniami:

1. Prawym przyciskiem myszy kliknij węzeł, który chcesz połączyć z nowym strumieniem.
2. W menu kliknij:

   Mapowanie danych > **Mapuj do**
3. Za pomocą kursora wybierz węzeł docelowy w strumieniu docelowym.
4. W oknie dialogowym, które zostanie otwarte, upewnij się, że zmienne zostały prawidłowo dopasowane, a następnie kliknij przycisk **OK**.

Określanie niezbędnych zmiennych

Podczas mapowania istniejącego strumienia jego autor określa zwykle niezbędne zmienne. W ten sposób wskazuje, czy konkretne zmienne są używane w dalszych operacjach w strumieniu. Na przykład istniejący strumień może budować model korzystający ze zmiennej o nazwie *Odchodzenie*. W tym strumieniu *Odchodzenie* jest zmienną niezbędną, ponieważ nie można bez niej zbudować modelu. Podobnie, zmienne używane w węzłach manipulacji, takich jak węzeł wyliczeń, są niezbędne do wyliczenia nowej zmiennej. Wyraźne wskazanie takich zmiennych jako niezbędnych pomaga w prawidłowym zmapowaniu na nie właściwych zmiennych w nowym węźle źródłowym. Jeśli zmienne obowiązkowe nie będą zmapowane, zostanie wyświetlony komunikat o błędzie. Jeśli uznasz, że niektóre węzły manipulacji lub wyników są zbędne, możesz usunąć je ze strumienia, a odpowiednie zmienne usunąć z listy niezbędnych.

Aby określić niezbędne zmienne:

1. Prawym przyciskiem myszy kliknij węzeł źródłowy szablonu strumienia, który chcesz zastąpić.
2. W menu kliknij:

   Mapowanie danych > **Określ niezbędne zmienne**
3. Za pomocą selektora zmiennych można dodawać lub usuwać zmienne z listy. Aby otworzyć selektor zmiennych, prawym przyciskiem myszy kliknij ikonę po prawej stronie listy zmiennych.

Analizowanie zmapowanych zmiennych

Po wybraniu punktu, w którym jeden strumień danych lub źródło danych zostanie zmapowane na inny/-e, wyświetlane jest okno dialogowe umożliwiające wybór pól dla mapowania lub upewnienie się do co poprawności domyślnego mapowania systemowego. W przypadku ustawienia dla strumienia lub źródła danych istotnych zmiennych i występowania niedopasowań te niepasujące zmienne są wyświetlane w kolorze czerwonym. Wszelkie niezmapowane zmienne ze źródła danych przechodzą w niezmienionym stanie przez węzeł filtrowania, należy jednak pamiętać, że można mapować również pozostałe, nie tak istotne zmienne.

Oryginalne. Zawiera listę wszystkich zmiennych w szablonie lub w istniejącym strumieniu — wszystkich zmiennych w dół strumienia. Zmienne z nowego źródła danych zostaną zmapowane na te zmienne.

Mapowane. Zawiera listę zmiennych wybranych do mapowania na te zmienne szablonu. Są to zmienne, dla których może zaistnieć konieczność zmiany nazw, tak aby były one zgodne ze zmiennymi oryginalnymi w operacjach strumieniowych. Kliknij w komórce tabeli dla zmiennej, aby aktywować listę dostępnych zmiennych.

W przypadku wątpliwości co do tego, które zmienne wymagają odwzorowania, może okazać się użyteczne zbadanie danych strumienia przed przystąpieniem do mapowania. Na przykład można za pomocą karty Typy w węźle źródłowym przejrzeć podsumowanie danych źródłowych.

Porady i skróty

Przyspiesz swoją pracę i uprość wykonywanie operacji dzięki poniższym skrótom i poradom:

- **Przyśpiesz tworzenie strumieni za pomocą kliknięcia dwukrotnego.** Wystarczy kliknąć dwukrotnie węzeł na palecie, aby go dodać i połączyć z bieżącym strumieniem.
- **Za pomocą kombinacji klawiszy wybieraj wiele węzłów.** Naciśnij klawisze Ctrl+Q i Ctrl+W, aby zmienić stan zaznaczenia wszystkich węzłów poniżej bieżącego.

- **Skróty klawiaturowe umożliwiają łączenie i rozłączanie węzłów.** Po wybraniu węzła w obszarze roboczym naciśnij klawisz F2, aby rozpocząć połączenie. Za pomocą klawisza Tab przejdź do żądanego węzła i naciśnij klawisze Shift+spacja, aby zakończyć połączenie. Naciśnij klawisz F3, aby odłączyć wszystkie wejścia i wyjścia od wybranego węzła.
- **Dostosuj kartę Paleta węzłów, dodając ulubione węzły.** W menu Narzędzia kliknij pozycję **Zarządzaj paletami** i otwórz okno dialogowe umożliwiające dodawanie, usuwanie i przenoszenie węzłów widocznych na palecie.
- **Zmieniaj nazwy węzłów i dodawaj podpowiedzi.** W każdym oknie dialogowym węzła znajduje się karta Adnotacje umożliwiająca określenie niestandardowej nazwy węzła w obszarze roboczym oraz dodanie podpowiedzi ułatwiającej zorganizowanie strumienia. Za pomocą długich adnotacji można śledzić postęp, zapisywać szczegóły dotyczące procesu oraz oznaczać decyzje biznesowe do zrealizowania.
- **Automatycznie wstawiaj wartości do wyrażeń CLEM.** Dzięki Konstruktorowi wyrażeń dostępnemu z poziomu wielu okien dialogowych (np. w węzłach wyliczenia lub wypełnienia) można automatycznie wstawiać wartości zmiennych do wyrażenia CLEM. Kliknij przycisk wartości dostępny w Konstruktorze wyrażeń, aby wybrać żądaną wartość zmiennej na liście dostępnych wartości.

  Rysunek 13. Przycisk wartości
- **Szybko przeglądaj pliki.** Podczas wyszukiwania plików w oknie dialogowym Otwórz użyj listy Plik (kliknij żółty przycisk w kształcie rombu widoczny w górnej części okna dialogowego obok pola Szukaj w), aby przejść do wcześniej przeglądanych katalogów oraz do katalogów domyślnych oprogramowania IBM SPSS Modeler. Za pomocą przycisków „w przód” i „w tył” można przechodzić między przeglądanymi katalogami.
- **Zwiększ czytelność w oknie wyników.** Wyniki można zamykać i usuwać za pomocą czerwonego przycisku *X* widocznego w prawym górnym narożniku każdego okna wyniku. W ten sposób na karcie wyników w panelu zarządzania można zachować tylko rokujące lub interesujące wyniki.

Dostępna jest pełna lista skrótów klawiaturowych obsługiwanych w oprogramowaniu. Więcej informacji można znaleźć w temacie “Ułatwienia dostępu z użyciem klawiatury” na stronie 256.

Czy wiesz, że można...

- Za pomocą myszy przeciągać i wybierać grupy węzłów w obszarze roboczym strumienia.
- Kopiować i wklejać węzły między strumieniami.
- Otwierać zasoby pomocy z poziomu każdego okna dialogowego i okna wyniku.
- Uzyskiwać pomoc dotyczącą metodologii CRISP-DM (Cross-Industry Standard Process for Data Mining). (W menu Pomoc wystarczy kliknąć pozycję **Metodologia CRISP-DM**.)

Rozdział 6. Praca z danymi

Możesz obejrzeć dane strumienia, klikając prawym przyciskiem węzeł danych i wybierając opcję **Wyświetl dane**. W oknie, które zostanie otwarte:

- Karta **Wykres** umożliwia tworzenie zaawansowanych wizualizacji danych służących do eksplorowania ich z różnych perspektyw i wykrywania wzorców, powiązań oraz relacji.
- Karta **Arkusz** przedstawia dane w formacie tabelarycznym, przeznaczonym tylko do odczytu.
- Karta **Audyt danych** przedstawia częstości i statystyki kolumn.
- Na karcie **Panel kontrolny** można kliknąć opcję **Uruchom projekt układu** i utworzyć układ służący do wyświetlania wielu wykresów na jednej stronie. Układy można zapisywać jako szablony, a następnie przeciągać i upuszczać zapisane wykresy w żądanych miejscach w układzie.
- Karta **Preferencje** umożliwia skonfigurowanie preferencji interfejsu użytkownika, takich jak język oraz wygląd i sposób działania.

Ta funkcja domyślnie korzysta z portu numer 28900. Aby użyć innego portu, zmień wartość ustawienia konfiguracyjnego `data_view_port_number` w pliku `options.cfg`.

Uwaga: Funkcja **Wyświetl dane** obecnie nie jest dostępna w wersji Subscription produktu SPSS Modeler.

Budowanie wykresów

Karta **Wykresy**umożliwia tworzenie wykresów na podstawie wstępnie zdefiniowanych wykresów z galerii lub składanie wykresów z odrębnych części (np. osi i słupków). Wykres buduje się, wybierając typ galerii wykresów lub podstawowe elementy z dostępnych opcji typów wykresów.

Podczas budowania wykresu obszar roboczy wyświetla podgląd wykresu. Podgląd wykorzystuje rzeczywiste etykiety zmiennych i poziomy pomiaru, które są reprezentatywne dla rzeczywistych danych.

Preferowaną metodą dla nowych użytkowników jest korzystanie z galerii. Informacje na temat korzystania z galerii zawiera sekcja “Budowanie wykresu za pomocą galerii typów wykresów” na stronie 68.

Uruchamianie kreatora wykresów

- Prawym przyciskiem myszy kliknij węzeł danych, z którym chcesz pracować, a następnie wybierz opcję **Wyświetl dane**.

Układ kreatora wykresów i związane z nim terminy

Ekran początkowy
Po uruchomieniu kreatora wykresów prezentowane są opcje umożliwiające wybór typu wykresu lub wybór kolumn z aktywnego zbioru danych. W trakcie dodawania kolumn do wizualizacji opcje typów wykresów są aktualizowane i wyświetlają typy wykresów zalecane dla wybranych kolumn.

Obszar roboczy
Obszar roboczy to obszar okna dialogowego kreatora wykresów, w którym budowany jest wykres.

Typ wykresu
Przedstawia dostępne typy wykresów. Elementy graficzne to pozycje na wykresie, które przestawiają dane. Są to słupki, punkty, linie itd.

Panel szczegółów
Panel szczegółów zawiera elementy podstawowe wykresu.

Ustawienia wykresu
Zawiera opcje umożliwiające wybór zmiennych używanych do budowania wykresu, metody rozkładu, pól tytułu i podtytułu itd. Opcje dostępne na panelu szczegółów będą się różniły w

zależności od wybranego typu wykresu. Szczegółowe informacje na temat opcji dostępnych dla każdego wykresu zawiera “Typy wykresów” na stronie 68.

Actions
Udostępnia opcje pobierania plików konfiguracji wykresów, pobierania wykresów jako plików obrazów, resetowania wykresów i ustawiania globalnych preferencji wykresów.

Budowanie wykresu za pomocą galerii typów wykresów

Najprostszym sposobem na utworzenie wykresu jest skorzystanie z galerii typów wykresów. Poniżej znajdują się ogólne kroki budowania wykresu z galerii.

1. W sekcji **Typy wykresów** wybierz kategorię wykresów. W obszarze roboczym wykresu wyświetlany jest podgląd wybranego typu wykresu.
2. Jeśli w obszarze roboczym jest już wyświetlany wykres, nowy wykres zastąpi zbiór osi wykresu i elementy graficzne.
   a. W zależności od wybranego typu wykresu dostępne zmienne są wyświetlane pod wieloma różnymi nagłówkami na panelu szczegółów (na przykład **Kategoria** w przypadku wykresów słupkowych, **Oś X** i **Oś Y** w przypadku wykresów liniowych). Wybierz zmienne odpowiednie dla wybranego typu wykresu.

Typy wykresów

Galeria zawiera zbiór najczęściej używanych wykresów. Należą do nich:

wykresy 3-W
Proste. Więcej informacji można znaleźć w temacie “wykresy 3-W” na stronie 69.

Wykresy słupkowe
Proste, zestawione i tworzące skupienia. Więcej informacji można znaleźć w temacie “Wykresy słupkowe” na stronie 70.

Wykresy skrzynkowe
Proste i tworzące skupienia. Więcej informacji można znaleźć w temacie “Wykresy skrzynkowe” na stronie 73.

Bąbelkowy
Proste. Więcej informacji można znaleźć w temacie “Wykresy bąbelkowe” na stronie 73.

Wykresy świecowe
Proste. Więcej informacji można znaleźć w temacie “Wykresy świecowe” na stronie 74.

Upakowane koła
Proste. Więcej informacji można znaleźć w temacie “Wykresy Upakowane koła” na stronie 75.

Wykresy zmodyfikowane przez użytkownika
Więcej informacji można znaleźć w temacie “Wykresy niestandardowe” na stronie 76.

Wykresy z dwiema osiami Y
Proste. Więcej informacji można znaleźć w temacie “Wykresy z dwiema osiami Y” na stronie 76.

Wykresy słupków błędu
Proste. Więcej informacji można znaleźć w temacie “Wykresy słupków błędu” na stronie 77.

Wykresy ewaluacyjne
Proste. Więcej informacji można znaleźć w temacie “Wykresy ewaluacyjne” na stronie 78.

Mapa natężeń
Proste. Więcej informacji można znaleźć w temacie “Wykresy map natężeń” na stronie 80.

Histogramy
Proste, zestawione i wieloboki częstości. Więcej informacji można znaleźć w temacie “Wykresy histogramów” na stronie 82.

Wykresy liniowe
Proste. Więcej informacji można znaleźć w temacie “Wykresy liniowe” na stronie 83.

Wykresy map
Proste. Więcej informacji można znaleźć w temacie “Wykresy map” na stronie 84.

Wykresy krzywych matematycznych
Proste. Więcej informacji można znaleźć w temacie “Wykresy krzywych matematycznych” na stronie 87.

Wykresy złożone
Proste. Więcej informacji można znaleźć w temacie “Wykresy złożone” na stronie 88.

Wykresy dla wielu szeregów
Proste. Więcej informacji można znaleźć w temacie “Wykresy dla wielu szeregów” na stronie 89.

Wykresy równoległe
Proste. Więcej informacji można znaleźć w temacie “Wykresy równoległe” na stronie 90.

Wykresy kołowe
Proste. Więcej informacji można znaleźć w temacie “Wykresy kołowe” na stronie 90.

piramidy populacyjne
Proste. Więcej informacji można znaleźć w temacie “Piramidy populacyjne” na stronie 92.

Wykresy K-K
Proste. Więcej informacji można znaleźć w temacie “Wykresy K-K” na stronie 92.

Wykresy radarowe
Proste. Więcej informacji można znaleźć w temacie “Wykresy radarowe” na stronie 94.

Wykresy relacji
Proste. Więcej informacji można znaleźć w temacie “Wykresy relacji” na stronie 95.

Wykresy rozrzutu i wykresy punktowe
wykresy rozrzutu 1-W, proste, zgrupowania, nakładania i rozrzutu 3-W; podsumowujące wykresy punktowe, wykresy punktowe 1-W i wykresy linii rzutowania. Więcej informacji można znaleźć w temacie “Wykresy rozrzutu i wykresy punktowe” na stronie 95.

Wykresy macierzy wykresów rozrzutu
Proste. Więcej informacji można znaleźć w temacie “Wykresy macierzy wykresów rozrzutu” na stronie 97.

Wykresy pierścieniowe
Proste. Więcej informacji można znaleźć w temacie “Wykresy pierścieniowe” na stronie 98.

Sekwencyjny
Proste. Więcej informacji można znaleźć w temacie “Wykresy sekwencyjne” na stronie 99.

Rzeka tematyczna
Proste. Więcej informacji można znaleźć w temacie “Wykresy Rzeka tematyczna” na stronie 100.

wykresy drzewa
Proste. Więcej informacji można znaleźć w temacie “wykresy drzewa” na stronie 101.

mapy drzew
Proste i pierścieniowe. Więcej informacji można znaleźć w temacie “mapy drzew” na stronie 102.

Wykresy t-SNE
Proste. Więcej informacji można znaleźć w temacie “Wykresy t-SNE” na stronie 103.

Chmury słów
Proste. Więcej informacji można znaleźć w temacie “Wykresy chmury słów” na stronie 103.

wykresy 3-W

Wykresy 3-W są powszechnie używane do przedstawiania funkcji wielu zmiennych i zawierają zmienną osi Z, która jest funkcją zmiennych osi X i Y.

Tworzenie prostego wykresu 3-W

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.

Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu 3-W.

2. Wybierz **typ** wykresu z listy rozwijanej.
3. Wybierz zmienną **osi X** z listy rozwijanej.
4. Wybierz zmienną **osi Y** z listy rozwijanej.
5. Wybierz zmienną **osi Z** z listy rozwijanej.

Funkcje dodatkowe

Typ
Przedstawia typy wykresów, których można użyć do przedstawienia danych.

Oś X
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Oś Y
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *Y* wykresu.

Oś Z
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *Z* wykresu.

Podpowiedź
Przedstawia zmienne, których można użyć w celu wygenerowania podpowiedzi po ustawieniu kursora nad punktem danych.

Odwzorowanie kolorów
Przedstawia dostępne zmienne odwzorowania kolorów. Zmienne te wykorzystują progresję kolorów, na podstawie zakresu wartości w określonej kolumnie, w celu przedstawienia ich w punktach wykresu. Odwzorowania kolorów są znane także jako kartogram.

Odwzorowanie wielkości
Przedstawia dostępne zmienne odwzorowania wielkości. Zmienne te wykorzystują różne wielkości do przedstawiania ich w punktach wykresu.

Iloraz Z
Ustawia skalę wartości danych osi *Z* w stosunku do osi *X* i Y.

Obróć
Włącza lub wyłącza obrót wykresu.

Podpowiedzi punktów danych
Określa miejsce wyświetlania podpowiedzi do punktów danych (z prawej strony punktów danych, w prawym górnym rogu wykresu albo ukryte).

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy słupkowe

Wykresy słupkowe są użyteczne do podsumowywania zmiennych jakościowych. Na przykład wykresu słupkowego można użyć do przedstawienia liczby mężczyzn i kobiet, którzy brali udział w ankiecie lub do przedstawienia średniego wynagrodzenia mężczyzn i kobiet.

Tworzenie prostego wykresu słupkowego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.

Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu słupkowego.

2. Wybierz zmienną jakościową (nominalną lub porządkową) jako **kategorię** zmiennej. Można użyć zmiennej ilościowej, ale wyniki będą użyteczne tylko w rzadkich, specjalnych przypadkach. Wykres słupkowy wygląda najlepiej z ograniczoną liczbą odmiennych wartości. Utworzenie wykresu słupkowego z osią **kategorii** ilościowej spowoduje, że słupki będą bardzo wąskie, ponieważ każdy słupek jest rysowany przy określonej wartości i nie może nakładać się na inne wartości ciągłe.
3. Wybierz statystykę z listy **Podsumowanie**. Wynik dowolnej statystyki określa wysokość słupka. Jeśli żądana statystyka nie pojawia się na liście **Podsumowanie**, może wymagać zmiennej. Wybierz zmienną z listy **Wartości** i sprawdź, czy statystyka jest teraz dostępna. Mogą występować inne ograniczenia typu wykresu. Na przykład wykresy słupków błędu można obliczyć tylko dla określonych statystyk.

Funkcje dodatkowe

Category
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Kolejność na podstawie
Wybierz opcję sortowania dla kategorii w obrębie zmiennej.

Nazwa kategorii
Użyj etykiet kategorii do sortowania kategorii zmiennej. Są to etykiety, które pojawiają się na wykresie, zwykle jako znacznik lub etykiety legendy.

Wartość kategorii
Użyj wartości przechowywanych w zbiorze danych do sortowania kategorii zmiennej. Wartość kategorii identyfikuje kategorię w zbiorze danych. Często różni się ona od jej etykiety i nie musi być opisowa. Na przykład wartość może być liczbą (np. *1*), podczas gdy etykieta jest opisem kategorii (np. *Kobieta*).

Porządek kategorii
Wybierz porządek sortowania kategorii zmiennej

Jak w danych
Kategorie zmiennej są wyświetlane w kolejności w jakiej pojawiają się w zbiorze danych.

Rosnąco
Posortuj kategorie zmiennych w porządku rosnącym.

Malejąco
Posortuj kategorie zmiennych w porządku malejącym.

Podsumowanie
Wybierz statystyczną funkcję podsumowującą. Wynik statystyki określa pozycję elementów graficznych na osi *Y*. W przypadku wykresu 2-W statystyka jest obliczana dla każdej wartości na osi *X*. W przypadku wykresu 3-wymiarowego statystyka jest obliczana dla punktów przecięcia wartości na osi *X* i osi *Z*.

Istnieją dwa typy funkcji podsumowania statystycznego. Rozróżnienie ich jest ważne, ponieważ od tego zależy, czy konieczne jest określenie zmiennej **wartości**.

- **Funkcje, które nie wymagają zmiennej wartości.** Są to funkcje, które nie wymagają zmiennej. W tej kategorii są wszystkie statystyki liczebności i procentowe. Te statystyki są dostępne, jeśli nie zdefiniowano zmiennej **wartości**.
- **Funkcje, które wymagają zmiennej wartości.** Są to funkcje, które wymagają zmiennej **wartości**. Na przykład funkcja *Średnia* wymaga zmiennej, na podstawie której obliczana jest średnia. Te statystyki są dostępne, jeśli zdefiniowano zmienną **wartości**.

Wartość
To pole jest wyświetlane, gdy wybrana jest funkcja **Podsumowanie**, która wymaga wartości zmiennej. Wybierz zmienną, która będzie służyła jako wartość.

Podziel według
Wybierz zmienną jakościową, która tworzy tabelę wykresów, z komórką dla każdej kategorii w zmiennej Podziel według. Podobnie jak w przypadku grupowania, opcja podziel według powoduje dodanie dodatkowych wymiarów do wykresu, wyświetlając informacje dla każdej kategorii zmiennej.

Typ podziału
Po wybraniu zmiennej **Podziel według** można wybrać wyświetlenie wygenerowanych słupków kategorii w formie zestawionej lub zgrupowanej. Grupowanie i zestawianie dodaje wymiarowość w obrębie wykresu. Grupowanie dzieli jeden słupek na wiele słupków, a zestawianie tworzy segmenty w każdym słupku. Należy być ostrożnym, aby wybrać statystykę odpowiednią do zestawiania. Po dodaniu wartości do siebie (zestawieniu ich) wynik musi mieć sens. Na przykład dodawanie i zestawianie średnich (uśrednionych) wartości zwykle nie ma sensu.

Typ wykresu słupkowego
Wybierz typ wykresu słupkowego z dostępnych opcji.

- Oś X
- Oś Y
- Odwrotna oś X
- Odwrotna oś Y
- Kąt biegunowy, oś
- Promień biegunowy, oś
- Biegunowy-tęczowy

Pozycja etykiety
Wybierz pozycję etykiety z menu rozwijanego.

- brak
- na górze
- na dole
- po prawej stronie
- po lewej stronie
- wewnątrz
- wewnątrz, po lewej stronie
- wewnątrz, po prawej stronie
- wewnątrz, na górze
- wewnątrz, na dole
- wewnątrz, na górze, po lewej stronie
- wewnątrz, na dole, po lewej stronie
- wewnątrz, na górze, po prawej stronie
- wewnątrz, na dole, po prawej stronie

Pokaż linię odniesienia
Przełącznik umożliwia włączenie/wyłączenie wyświetlania linii odniesienia na wykresie. Dostępne opcje to **Minimum**, **Maksimum** i **Średnia**, które powodują wyświetlenie linii odniesienia odpowiednio dla wartości minimalnych, maksymalnych i średnich na wykresie.

Wprowadź wartość linii odniesienia
Gdy ustawienie **Pokaż linię odniesienia** jest włączone, umożliwia określenie wartości linii odniesienia. Kliknij opcję **Dodaj kolejną kolumnę**, aby określić dodatkowe wartości linii odniesienia.

Transpozycja
Gdy opcja ta jest włączona, osie *X* i *Y* są transponowane.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy skrzynkowe

Wykres skrzynkowy przedstawia pięć statystyk (minimum, pierwszy kwartyl, mediana, trzeci kwartyl i maksimum). Jest on użyteczny do wyświetlania rozkłady zmiennej ilościowej i wskazywania wartości odstających.

Tworzenie prostego wykresu skrzynkowego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.


   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu skrzynkowego.
2. Wybierz co najmniej dwie zmienne ilościowe jako zmienne **kolumnowe**.

Uwaga: Statystyką dla wykresu punktowego jest wykres skrzynkowy. Nie można tego zmienić.

Funkcje dodatkowe

Kolumny
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe kolumny.

Podziel według
Wybierz zmienną jakościową, która tworzy tabelę wykresów, z komórką dla każdej kategorii w zmiennej Podziel według. Podobnie jak w przypadku grupowania, opcja podziel według powoduje dodanie dodatkowych wymiarów do wykresu, wyświetlając informacje dla każdej kategorii zmiennej.

Porządek kategorii
Wybierz porządek sortowania kategorii zmiennej

Jak w danych
Kategorie zmiennej są wyświetlane w kolejności w jakiej pojawiają się w zbiorze danych.

Rosnąco
Posortuj kategorie zmiennych w porządku rosnącym.

Malejąco
Posortuj kategorie zmiennych w porządku malejącym.

Normalizuj dane
Po włączeniu tego ustawienia dane są przekształcane do rozkładu normalnego, który umożliwia porównanie danych z wielu zbiorów danych lub z wielu kolumn. To ustawienie tworzy 100% zestawianie dla liczebności i przekształca statystyki na procenty.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy bąbelkowe

Wykresy bąbelkowe przedstawiają kategorie w grupach jako niehierarchiczne okręgi upakowane. Wielkość każdego okręgu (bąbelka) jest proporcjonalna do jego wartości. Wykresy bąbelkowe są przydatne do porównywania relacji w danych.

Tworzenie prostego wykresu bąbelkowego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Bąbelkowy**.

   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu bąbelkowego.
2. Wybierz zmienną **Kolumny** z listy rozwijanej.

   Uwaga: Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe zmienne kolumn.

Funkcje dodatkowe

Kolor grupy
Włącz lub wyłącz grupowanie kolorów.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy świecowe

Wykresy świecowe to styl wykresów finansowych, które służą do opisywania zmian cen bezpieczeństwa, pochodnej lub waluty. Każdy element świecowy zwykle przedstawia jeden dzień. Wykres jednego miesiąca może przedstawiać 20 dni handlowych jako 20 elementów świecowych. Wykresy świecowe są najczęściej używane w analizie schematów kapitału i cen waluty i są podobne do wykresów skrzynkowych.

Zbiór danych używany do utworzenia wykresu świecowego musi zawierać wartości otwarcia, wysokie, niskie i zamknięcia dla każdego okresu, który ma zostać wyświetlony.

Tworzenie prostego wykresu świecowego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.

   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu świecowego.
2. Wybierz zmienną jako zmienną **osi X**.
3. Wybierz zmienną jako zmienną **wysoką**.
4. Wybierz zmienną jako zmienną **niską**.

Funkcje dodatkowe

Oś X
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Wysoka
Przedstawia zmienne ze zbioru danych, które są dostępne dla wartości wysokiej ceny.

Podsumowanie zmiennej wysokiej
Wybierz funkcję podsumowania statystycznego dla wybranej zmiennej wysokiej.

Niska
Przedstawia zmienne ze zbioru danych, które są dostępne dla wartości niskiej ceny.

Otwarcie
Przedstawia zmienne ze zbioru danych, które są dostępne dla wartości ceny otwarcia.

Zamknięcie
Przedstawia zmienne ze zbioru danych, które są dostępne dla wartości ceny zamknięcia.

Objętość
Przedstawia zmienne ze zbioru danych, które są dostępne dla słupków objętości wykresu.

Porządek kategorii
Wybierz porządek sortowania kategorii zmiennej

Jak w danych
Kategorie zmiennej są wyświetlane w kolejności w jakiej pojawiają się w zbiorze danych.

Rosnąco
Posortuj kategorie zmiennych w porządku rosnącym.

Malejąco
Posortuj kategorie zmiennych w porządku malejącym.

Świeca
Przełącza dane wykresu między widokiem świec i widokiem linii.

Średnia ruchoma
Udostępnia opcje wyświetlania średniej ruchomej na wykresie. Dostępne opcje to 5, 10, 20, 30, 60 i 120.

Przełącz kolor
Przełącznik umożliwia przełączanie kolorów wzrostu i spadku.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

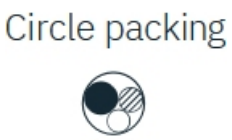
Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy Upakowane koła

Wykres upakowanych kół przedstawia dane hierarchiczne jako zestaw zagnieżdżonych powierzchni, pozwalając na wizualizację dużych ilości danych o strukturze hierarchicznej. Jest podobny do mapy drzewa, ale składa się z kół, a nie z prostokątów. Wykresy Upakowane koła przedstawiają dane hierarchiczne poprzez zawieranie (zagnieżdżanie) kół.

Tworzenie prostego wykresu Upakowane koła

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres sekwencyjny**.



   Obszar roboczy zostaje zaktualizowany i wyświetla szablon upakowanych kół.
2. Wybierz zmienną **Kolumny** z listy rozwijanej.

   Uwaga: Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe zmienne kolumn.

Funkcje dodatkowe

Kolor grupy
Włącz lub wyłącz grupowanie kolorów.

Podsumowanie
Wybierz funkcję podsumowania statystycznego. Jest to metoda tworzenia podsumowania każdej kategorii.

Istnieją dwa typy funkcji podsumowania statystycznego. Rozróżnienie ich jest ważne, ponieważ od tego zależy, czy konieczne jest określenie zmiennej **wartości**.

- **Funkcje, które nie wymagają zmiennej wartości.** Są to funkcje, które nie wymagają zmiennej. W tej kategorii są wszystkie statystyki liczebności i procentowe. Te statystyki są dostępne, jeśli nie zdefiniowano zmiennej **wartości**.
- **Funkcje, które wymagają zmiennej wartości.** Są to funkcje, które wymagają zmiennej **wartości**. Na przykład funkcja *Średnia* wymaga zmiennej, na podstawie której obliczana jest średnia. Te statystyki są dostępne, jeśli zdefiniowano zmienną **wartości**.

Wartość
To pole jest wyświetlane, gdy wybrana jest funkcja **Podsumowanie**, która wymaga wartości zmiennej. Wybierz zmienną, która będzie służyła jako wartość.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy niestandardowe

Opcja wykresów niestandardowych umożliwia wklejenie/edytowanie kodu JSON w celu utworzenia żądanego wykresu.

Tworzenie wykresu niestandardowego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.


2. Wklej kod JSON, który zawiera specyfikacje wykresu, do udostępnionego pola na panelu szczegółów.
3. Kliknij opcję **Wygeneruj wykres**.

Wykresy z dwiema osiami Y

Wykres z dwiema osiami *Y* umożliwia podsumowanie lub wykreślenie dwóch zmiennych w osi *y* o różnych dziedzinach. Na przykład można wykreślić liczbę obserwacji na jednej osi i średnią pensję na drugiej osi. Ten wykres może zawierać jednocześnie różne elementy graficzne, dlatego wykres z dwiema osiami *Y* zawiera wiele omówionych wcześniej typów wykresów. Może na przykład prezentować liczebności jako linię oraz średnią dla każdej kategorii jako słupek.

Tworzenie prostego wykresu z dwiema osiami Y

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Dwie osie Y**.


   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu z dwiema osiami Y.
2. Wybierz zmienną jako zmienną **osi X**.
3. Wybierz zmienną dla pierwszej **osi Y**, a następnie wybierz typ wykresu, który będzie reprezentować zmienną (**Słupkowy**, **Liniowy** lub **Wykres rozrzutu**).
4. Wybierz zmienną dla drugiej **osi Y**, a następnie wybierz typ wykresu, który będzie reprezentować zmienną (**Słupkowy**, **Liniowy** lub **Wykres rozrzutu**).

Uwaga: Za pomocą klawiszy strzałek w górę i w dół można zmienić kolejność osi *y*.

Funkcje dodatkowe

Oś X
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Oś Y
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *y* wykresu.

Podsumowanie
Po włączeniu tej opcji wyświetlane są opcje służące do wyboru metody podsumowywania każdej kategorii.

Podsumowanie lewej osi Y
Określa metodę podsumowania dla osi Y, która jest wyświetlana po lewej stronie wykresu. Dostępne opcje to: **Suma**, **Średnia**, **Maksimum** i **Minimum**.

Podsumowanie prawej osi Y
Określa metodę podsumowania dla osi Y, która jest wyświetlana po prawej stronie wykresu. Dostępne opcje to: **Suma**, **Średnia**, **Maksimum** i **Minimum**.

Normalizuj dane
Włączenie tej opcji powoduje przekształcenie danych do rozkładu normalnego. Umożliwia to łatwe porównywanie danych z wielu zestawów danych lub kolumn.

Linie drugiej osi
Gdy ta opcja jest włączona, wyświetlana jest linia drugiej osi wykresu.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy słupków błędu

Wykresy słupków błędu przedstawiają zmienność danych i wskazują błąd (lub niepewność) w raportowanym pomiarze. Słupki błędów pomagają określić, czy różnice są statystycznie istotne. Słupki błędów mogą również sugerować dobroć dopasowania dla danej funkcji.

Tworzenie prostego wykresu słupków błędu

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.
   Error bar

   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu słupków błędu.
2. Wybierz zmienną skali jako zmienną **kategorii**. Jest to zmienna, której dane są prezentowane na osi *X*.
3. Wybierz zmienną skali jako zmienną **osi Y**. Jest to zmienna, której dane są prezentowane na osi *Y*.

Funkcje dodatkowe

Category
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Oś Y
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *Y* wykresu.

Porządek kategorii
Wybierz porządek sortowania kategorii zmiennej

Jak w danych
Kategorie zmiennej są wyświetlane w kolejności w jakiej pojawiają się w zbiorze danych.

Rosnąco
Posortuj kategorie zmiennych w porządku rosnącym.

Malejąco
Posortuj kategorie zmiennych w porządku malejącym.

Podziel według
Wybierz zmienną jakościową, która tworzy tabelę wykresów, z komórką dla każdej kategorii w zmiennej Podziel według. Podobnie jak w przypadku grupowania, opcja podziel według powoduje dodanie dodatkowych wymiarów do wykresu, wyświetlając informacje dla każdej kategorii zmiennej.

Linia odniesienia
Po włączeniu wyświetla linię odniesienia. Linia odniesienia koreluje z wybraną **metodą statystyczną**.

Słupki błędów
Po włączeniu linie, które przedstawiają przedział błędu są wyświetlane na wykresie.

Miara
Wybierz typ miary reprezentowany przez słupki błędów:

Przedziały ufności
Ustawia przedziały ufności dla wybranych zmiennych. Wartość domyślna to 0,95 (95%), co odzwierciedla pole **Wartość reprezentacyjna**.

Błąd standardowy
Mierzy błąd standardowy wybranych zmiennych.

Odchylenie standardowe
Mierzy odchylenie standardowe wybranych zmiennych.

Poziom ufności
Ta wartość przedstawia przedziały ufności dla wybranych **miar**. Wartość domyślna to 0,95 (95%).

Metoda statystyczna
Wybierz metodę opisywania tendencji centralnej:

Średnia
Suma proporcji podzielona przez ich całkowitą liczbę.

Mediana
Wartość, przy której liczba proporcji mniejszych i większych od niej jest równa.

Tryb widoku
Wybierz sposób wyświetlania wyboru **metody statystycznej** (słupek, linia lub koło).

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy ewaluacyjne

Wykresy ewaluacyjne są podobne do histogramów lub wykresów zbiorów. Wykresy ewaluacyjne przedstawiają skuteczność modeli w przewidywaniu konkretnych wyników. Ich działanie polega na sortowaniu rekordów na podstawie wartości przewidywanej i ufności przewidywania, dzieleniu rekordów na grupy o równej wielkości (**kwantyle**), a następnie wykreśleniu wartości kryterium dla każdego kwantyla, od najwyższego do najniższego. Modele wielokrotne prezentowane są jako osobne linie na wykresie.

Wyniki uzyskuje się poprzez zdefiniowanie konkretnej wartości lub zakresu wartości jako **trafienia**. Trafienia zwykle oznaczają sukces (np. sprzedaż klientowi) lub zdarzenie będące przedmiotem zainteresowania (np. konkretną diagnozę medyczną).

Flaga
Zmienne wynikowe nie wymagają dodatkowych objaśnień: trafienia odpowiadają wartościom *prawda*.

Nominalny
W przypadku nominalnych zmiennych wynikowych trafieniem jest pierwsza wartość w zbiorze.

Ilościowy
W przypadku ciągłych zmiennych wynikowych trafieniami są wartości większe od połowy zakresu wartości zmiennej.

Wykresy ewaluacyjne mogą być także skumulowane, tak aby każdy punkt był równy wartości odpowiedniego kwantyla plus wartości wszystkich wyższych kwantyli. Wykresy skumulowane zwykle lepiej obrazują ogólną wydajność modeli, natomiast wykresy nieskumulowane często lepiej nadają się do ujawniania konkretnych obszarów problemowych w modelach.

Tworzenie prostego wykresu ewaluacyjnego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres ewaluacyjny**.
   Evaluation
   

   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu ewaluacyjnego.
2. Ustaw zmienne **Zmienna docelowa**, **Zmienna przewidywana** i **Zmienna ufności**. Zmienna przewidywana może być dowolną zmienną typu flaga lub nominalna z dwiema lub większą wartością. Jest to zmienna używana jako wartość przewidywana. Zmienna ufności określa zmienną używaną do ustalania ufności predykcji.

   Uwaga: Typ zmiennej określonej w polu **Przewiduj zmienną** musi zgadzać się z typem zmiennej wybranej jako **Zmienna przewidywana**.
3. Określ warunek niestandardowy, którego spełnienie oznacza **Trafienie zdefiniowane przez użytkownika**. Opcja ta jest przydatna do definiowania wyniku zamiast wyliczania go na podstawie typu zmiennej przewidywanej i kolejności wartości.

   Należy określić wyrażenie CLEM definiujące warunek trafienia. Przykładowo, warunek `@TARGET = "YES"` jest poprawnym warunkiem określającym, że wartość *Yes* (Tak) dla zmiennej przewidywanej zostanie zliczona jako trafienie podczas ewaluacji. Określony warunek zostanie użyty dla wszystkich zmiennych przewidywanych.

Funkcje dodatkowe

Wykres kumulacyjny
Gdy ta opcja jest włączona, tworzony jest wykres kumulacyjny. Wartości na wykresach skumulowanych są wykreślane dla każdego kwantyla plus wszystkie wyższe kwantyle.

Tryb widoku
Ustawienia określają, które wykresy będą widoczne w trybie podglądu i w wynikach.

Tryb klasyczny
Gdy ta opcja jest wybrana, w trybie podglądu i w wynikach widoczne są wykresy Strojenie klasyfikacji wyniku, Punkt odcięcia, Słupkowy macierzowy, ROC, Korzyści, ROI i Zyski.

Tryb pojedynczy
Gdy ta opcja jest wybrana, w trybie podglądu i w wynikach widoczny jest tylko wykres Strojenie klasyfikacji wyniku.

Tryb pełny
Gdy ta opcja jest wybrana, w trybie podglądu i w wynikach widoczne są wykresy Strojenie klasyfikacji wyniku, Punkt odcięcia, Słupkowy macierzowy, ROC, Korzyści, ROI, Zyski, GINI, Przyrost i Odpowiedź.

Wykresy ewaluacyjne

Punkt odcięcia
Wykres punktu odcięcia przedstawia wartości przewidywane i rzeczywiste wybranych zmiennych dla danej wartości odcięcia.

Słupkowy macierzowy
Słupkowe wykresy macierzowe są dobrym sposobem na określenie, czy między wieloma zmiennymi istnieją korelacje liniowe.

ROC
Wykres ROC (Receiver Operating Characteristic — charakterystyka działania odbiornika) ocenia wydajność schematów klasyfikacji, w których istnieje jedna zmienna z dwiema kategoriami klasyfikującymi obiekt.

Korzyści
Korzyści zdefiniowane są jako proporcja łącznej liczby trafień w każdym kwantylu. Korzyści oblicza się według wzoru (liczba trafień w kwantylu / łączna liczba trafień) × 100%.

ROI
Zwrot z inwestycji (ROI — Return on investment) jest podobny do zysku w tym, że obliczany jest na podstawie przychodów i kosztów. Zwrot z inwestycji jest porównaniem zysków z kosztami z danego kwantyla. Zwrot z inwestycji oblicza się według wzoru (zyski z kwantyla / koszty z kwantyla) × 100%.

Profit
Zysk równy jest przychodowi w każdym rekordzie pomniejszonemu o koszt w tym rekordzie. Zysk z kwantyla jest po prostu sumą zysków z wszystkich rekordów w tym kwantylu. Przyjmuje się, że przychody mają zastosowanie tylko do trafień, ale koszty — do wszystkich rekordów. Zyski i koszty mogą być stałe lub zdefiniowane przez zmienne w danych. Zyski oblicza się według wzoru (suma przychodów z rekordów w kwantylu – suma kosztów z rekordów w kwantylu).

Kołmogorow-Smirnow
Porównuje obserwowaną funkcję skumulowanego rozkładu dla zmiennej z określonym teoretycznym rozkładem normalnym, jednostajnym, wykładniczym lub Poissona.

GINI
Wskaźnik GINI jest miarą rozproszenia statystycznego i służy do przedstawiania rozkładu przychodów lub zamożności. Jest to jedna z najczęściej używanych miar nierówności.

Wzrost
Przyrost porównuje odsetek rekordów w każdym kwantylu będących trafieniami z łącznym odsetkiem trafień w danych uczących. Obliczany jest według wzoru (trafienia w kwantylu / rekordy w kwantylu) / (łącznie trafień / łącznie rekordów).

Response
Odpowiedź jest to po prostu odsetek rekordów w kwantylu będących trafieniami. Odpowiedź obliczana jest według wzoru (trafienia w kwantylu / rekordy w kwantylu) × 100%.

Ustawienia wykresu ewaluacyjnego
Następujące ustawienia mają zastosowanie tylko do wykresów zysków i ROI.

Koszty
Określ stały koszt związany z każdym rekordem.

Przychód
Określ stały przychód związany z każdym rekordem, który reprezentuje trafienie.

Grubość linii
Jeśli rekordy w danych reprezentują więcej niż jedną jednostkę, można użyć wag częstości, aby skorygować wyniki. Określ stałą wagę związaną z każdym rekordem.

Wykresy map natężeń

Wykresy map natężeń przedstawiają dane, gdzie pojedyncze wartości zawarte w macierzy są oznaczane kolorami.

Tworzenie prostego wykresu mapy natężeń

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.

   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu mapy natężeń.
2. Wybierz zmienną jako zmienną **kolumnową**. Każda kategoria zmiennej jest przedstawiana jako pojedyncza kolumna wykresu.
3. Wybierz zmienną jako zmienną **wierszową**. Każda kategoria zmiennej jest przedstawiana jako pojedynczy wiersz wykresu.

Funkcje dodatkowe

Kolumna
Przedstawia zmienne ze zbioru danych, które są dostępne dla kolumn wykresu. Każda kategoria zmiennej jest przedstawiana jako pojedyncza kolumna wykresu.

Wiersz
Przedstawia zmienne ze zbioru danych, które są dostępne dla wierszy wykresu. Każda kategoria zmiennej jest przedstawiana jako pojedynczy wiersz wykresu.

Porządek kategorii
Wybierz porządek sortowania kategorii zmiennej

Jak w danych
Kategorie zmiennej są wyświetlane w kolejności w jakiej pojawiają się w zbiorze danych.

Rosnąco
Posortuj kategorie zmiennych w porządku rosnącym.

Malejąco
Posortuj kategorie zmiennych w porządku malejącym.

Podsumowanie
Wybierz statystyczną funkcję podsumowującą. Wynik statystyki określa pozycję elementów graficznych na osi *Y*. W przypadku wykresu 2-W statystyka jest obliczana dla każdej wartości na osi *X*. W przypadku wykresu 3-wymiarowego statystyka jest obliczana dla punktów przecięcia wartości na osi *X* i osi *Z*.

Istnieją dwa typy funkcji podsumowania statystycznego. Rozróżnienie ich jest ważne, ponieważ od tego zależy, czy konieczne jest określenie zmiennej **wartości**.

- **Funkcje, które nie wymagają zmiennej wartości.** Są to funkcje, które nie wymagają zmiennej. W tej kategorii są wszystkie statystyki liczebności i procentowe. Te statystyki są dostępne, jeśli nie zdefiniowano zmiennej **wartości**.
- **Funkcje, które wymagają zmiennej wartości.** Są to funkcje, które wymagają zmiennej **wartości**. Na przykład funkcja *Średnia* wymaga zmiennej, na podstawie której obliczana jest średnia. Te statystyki są dostępne, jeśli zdefiniowano zmienną **wartości**.

Wartość
To pole jest wyświetlane, gdy wybrana jest funkcja **Podsumowanie**, która wymaga wartości zmiennej. Wybierz zmienną, która będzie służyła jako wartość.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy histogramów

Histogram wygląda podobnie do wykresu słupkowego, ale zamiast porównywania kategorii lub wyszukiwania trendów w czasie, każdy słupek reprezentuje sposób rozkładu danych w pojedynczej kategorii. Każdy słupek reprezentuje ciągły przedział danych lub liczbę częstości dla określonego punktu danych.

Histogramy są użyteczne do przedstawiania rozkładu pojedynczej zmiennej ilościowej. Dane są dzielone i podsumowywane za pomocą statystyki liczebności lub procentowej. Odmianą histogramu jest wielokąt częstości, który jest podobny do typowego histogramu, ale zamiast słupkowego elementu graficznego jest używany obszarowy element graficzny.

Inną odmianą histogramu jest piramida populacyjna. Jej nazwa pochodzi od jej najczęstszego zastosowania: podsumowywania danych populacji. Podczas używania jej z danymi populacji jest ona podzielona według płci, generując przylegające do siebie, poziome histogramy danych wiekowych. W przypadku krajów z młodą populacją kształt wygenerowanego wykresu przypomina piramidę.

Tworzenie wykresu histogramu

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.

   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu histogramu.
2. Wybierz zmienną ilościową jako zmienną **osi X**.

   Uwaga: Statystyka dla histogramu to Histogram lub Histogram Procentowo. Te statystyki dzielą dane i obliczają liczebność dla każdego przedziału.

Funkcje dodatkowe

Oś X
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Podziel według
Wybierz zmienną jakościową, która tworzy tabelę wykresów, z komórką dla każdej kategorii w zmiennej Podziel według. Podobnie jak w przypadku grupowania, opcja podziel według powoduje dodanie dodatkowych wymiarów do wykresu, wyświetlając informacje dla każdej kategorii zmiennej.

Pokaż krzywą kde
Gdy opcja ta jest włączona, na wykresie wyświetlana jest krzywa estymatora jądrowego gęstości.

Pokaż krzywą rozkładu
Gdy opcja ta jest włączona, na wykresie wyświetlana jest krzywa dopasowania rozkładu.

Rozkład
Na liście rozwijanej dostępne są następujące opcje rozkładu.

Rozkład z dopasowaniem automatycznym
Jest to ustawienie domyślne.

Beta
Zwraca wartość o rozkładzie Beta z określoną wartością parametrów kształtu.

Wykładniczy
Zwraca wartość o rozkładzie wykładniczym.

Gamma
Zwraca wartość o rozkładzie Gamma z określoną wartością parametrów kształtu i skali.

Logarytmiczno-normalny
Zwraca wartość o rozkładzie logarytmiczno-normalnym z określoną wartością parametrów.

Normalny
Zwraca liczbę o rozkładzie normalnym z określoną wartością parametrów: średnia oraz odchylenie standardowe.

Trójkątny
Zwraca wartość o rozkładzie trójkątnym z określoną wartością parametrów.

Jednostajny
Zwraca wartość o rozkładzie jednostajnym między wartością minimalną a wartością maksymalną.

Weibulla
Zwraca wartość o rozkładzie Weibulla z określoną wartością parametrów.

Szerokość przedziału
Suwak steruje rozmiarem przedziału, który jest używany do podziału danych na grupy.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy liniowe

Wykres liniowy wykreśla szeregi punktów danych na wykresie i łączy je linią. Wykres liniowy jest szczególnie użyteczny podczas wyświetlania nieznacznie różniących się linii trendów lub przecinających się linii danych. Wykresu liniowego można użyć do podsumowania zmiennych jakościowych, co przebiega podobnie jak w przypadku wykresu słupkowego (patrz “Wykresy słupkowe” na stronie 70). Wykresy liniowe są również użyteczne dla danych szeregu czasowego.

Tworzenie prostego wykresu liniowego szeregów czasowych

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.


   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu liniowego.
2. Wybierz zmienną daty jako zmienną **osi X**.
3. Wybierz zmienną skali jako zmienną **osi Y**. Jest to zmienna, której wartości były zapisywane w czasie.

Funkcje dodatkowe

Oś X
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Oś Y
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *Y* wykresu.

Podziel według
Wybierz zmienną jakościową, która tworzy tabelę wykresów, z komórką dla każdej kategorii w zmiennej Podziel według. Podobnie jak w przypadku grupowania, opcja podziel według powoduje dodanie dodatkowych wymiarów do wykresu, wyświetlając informacje dla każdej kategorii zmiennej.

Obszar
Gdy opcja ta jest włączona, obszar pod linią jest wypełniony innym kolorem.

Wygładź
Gdy ta opcja jest włączona, na wykresie wyświetlana jest gładka krzywa.

Pokaż punkt danych
Gdy ta opcja jest włączona, na wykresie wyświetlany jest punkt danych.

Reorganizuj
Przełącznik umożliwia reorganizację danych na podstawie wartości osi X i Y.

Pokaż linię odniesienia
Gdy ta opcja jest włączona, na wykresie wyświetlana jest linia odniesienia oparta na określonych wartościach osi X i osi Y.

Wprowadź wartość linii odniesienia na osi X
Gdy ustawienie **Pokaż linię odniesienia** jest włączone, umożliwia określenie wartości linii odniesienia dla osi X. Kliknij opcję **Dodaj kolejną kolumnę**, aby określić dodatkowe wartości linii odniesienia.

Wprowadź wartość linii odniesienia na osi Y
Gdy ustawienie **Pokaż linię odniesienia** jest włączone, umożliwia określenie wartości linii odniesienia dla osi Y. Kliknij opcję **Dodaj kolejną kolumnę**, aby określić dodatkowe wartości linii odniesienia.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy map

Wykresy map są powszechnie używane w celu porównywania wartości i przedstawiania kategorii wśród regionów geograficznych. Są one najlepiej wykorzystywane, gdy dane zawierają informacje geograficzne (kraje, regiony, stany, miasta, kody pocztowe itd.)

Tworzenie prostego wykresu mapy

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.

   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu mapy.
2. Wybierz usługę mapy, która będzie używana do udostępniania obrazów map z listy rozwijanej **Usługi mapy**. Ta lista zawiera opcje, które obejmują określone regiony globalne.
3. Wybierz **Typ** wykresu mapy z menu rozwijanego. W zależności od wybranego typu wykresu dostępne są następujące opcje:

   Długość geograficzna
   Wybierz zmienną, która ma służyć jako wartość długości geograficznej, z listy rozwijanej.

   Szerokość geograficzna
   Wybierz zmienną, która ma służyć jako wartość szerokości geograficznej, z listy rozwijanej.

   Grupa
   Wybierz zmienną, która grupuje lokalizacje punktów danych, z listy rozwijanej.

   Category
   Wybierz zmienną kolumnową, która ma zostać zwizualizowana.

   Podsumowanie
   Wybierz funkcję podsumowania statystycznego. Jest to metoda tworzenia podsumowania każdej kategorii.

   Istnieją dwa typy funkcji podsumowania statystycznego. Rozróżnienie ich jest ważne, ponieważ od tego zależy, czy konieczne jest określenie zmiennej **wartości**.

   - **Funkcje, które nie wymagają zmiennej wartości.** Są to funkcje, które nie wymagają zmiennej. W tej kategorii są wszystkie statystyki liczebności i procentowe. Te statystyki są dostępne, jeśli nie zdefiniowano zmiennej **wartości**.
   - **Funkcje, które wymagają zmiennej wartości.** Są to funkcje, które wymagają zmiennej **wartości**. Na przykład funkcja *Średnia* wymaga zmiennej, na podstawie której obliczana jest średnia. Te statystyki są dostępne, jeśli zdefiniowano zmienną **wartości**.

Wartość
To pole jest wyświetlane, gdy wybrana jest funkcja **Podsumowanie**, która wymaga wartości zmiennej. Wybierz zmienną, która będzie służyła jako wartość.

Funkcje dodatkowe

Usługa map
Przedstawia usługi, które są dostępne w celu udostępniania obrazów map. Patrz "Opcje usługi mapy" na stronie 85.

Typ
Przedstawia typy wykresów, których można użyć do przedstawienia danych.

Długość geograficzna
Przedstawia zmienne, których można użyć jako wartości długości geograficznych.

Szerokość geograficzna
Przedstawia zmienne, których można użyć jako wartości szerokości geograficznych.

Grupa
Zawiera listę zmiennych, których można użyć do pogrupowania lokalizacji punktów danych.

Category
Zawiera listę zmiennych kolumnowych.

Podsumowanie
Wybierz funkcję podsumowania statystycznego. Jest to metoda tworzenia podsumowania każdej kategorii.

Istnieją dwa typy funkcji podsumowania statystycznego. Rozróżnienie ich jest ważne, ponieważ od tego zależy, czy konieczne jest określenie zmiennej **wartości**.

- **Funkcje, które nie wymagają zmiennej wartości.** Są to funkcje, które nie wymagają zmiennej. W tej kategorii są wszystkie statystyki liczebności i procentowe. Te statystyki są dostępne, jeśli nie zdefiniowano zmiennej **wartości**.
- **Funkcje, które wymagają zmiennej wartości.** Są to funkcje, które wymagają zmiennej **wartości**. Na przykład funkcja *Średnia* wymaga zmiennej, na podstawie której obliczana jest średnia. Te statystyki są dostępne, jeśli zdefiniowano zmienną **wartości**.

Wartość
To pole jest wyświetlane, gdy wybrana jest funkcja **Podsumowanie**, która wymaga wartości zmiennej. Wybierz zmienną, która będzie służyła jako wartość.

Podpowiedź
Przedstawia zmienne, których można użyć w celu wygenerowania podpowiedzi po ustawieniu kursora nad punktem danych.

Odwzorowanie wielkości
Przedstawia dostępne zmienne odwzorowania wielkości. Zmienne te wykorzystują różne wielkości do przedstawiania ich w punktach wykresu.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Opcje usługi mapy

Na panelu Szczegóły wykresów map można wybrać usługę mapy, która będzie używana do udostępniania obrazów map.

Konfigurowanie opcji usługi mapy

1. Zatrzymaj serwer IBM SPSS Modeler Server.
2. Jeśli mają być stosowane lokalne pliki map w formacie geojson, należy umieścić je w katalogu <Katalog_instalacyjny_programu_Modeler>/dataview/conf/public/mapfiles. Jeśli ten katalog nie istnieje, należy go utworzyć.
3. Otwórz plik <Katalog_instalacyjny_programu_Modeler>/dataview/conf/application.conf w edytorze tekstu i dodaj następujące opcje do sekcji map.resources. W tym przykładzie dodawane są dwie nowe mapy o nazwach World - Local Map i World - Online Map.

```
map.resources {
  "mapservices" :
    [

      {
        "type"      : "geojson",
        "name"      : "World (Bulit-in)",
        "location"  : "built_in_world",
        "id"        : "built_in_world"
      },
      {
        "type"      : "geojson",
        "name"      : "Blank (Built-in)",
        "location"  : "built_in_blank",
        "id"        : "built_in_blank"
      },
      {
        "type"      : "geojson",
        "name"      : "World - Local Map",
        "location"  : "/mapfiles/world.json",
        "useProxy"  : "false",
        "id"        : "world_local"
      },
      {
        "type"      : "geojson",
        "name"      : "World - Online Map",
        "location"  : "http://map.example.com/map/world.json",
        "useProxy"  : "true",
        "id"        : "world_online"
      }    ]
}
```

Dla opcji mapy dostępne są następujące parametry:

Tabela 4. Parametry usługi mapy

Opcja	Wartość	Typ	Wymagane	Opis
type	geojson	Łańcuch	Tak	Obecnie obsługiwana jest tylko jedna wartość tej opcji: geojson.
name	Nazwa mapy	Łańcuch	Tak	Wpisz nazwę mapy. Ta nazwa jest wyświetlana na liście rozwijanej **Usługa mapy** na panelu **Szczegóły** dla wykresów map.

Tabela 4. Parametry usługi mapy (kontynuacja)				
Opcja	**Wartość**	**Typ**	**Wymagane**	**Opis**
`location`	Lokalizacja mapy	Łańcuch	Tak	Wprowadź adres URL lub położenie pliku lokalnego mapy. Adres URL musi zaczynać się od `http://` lub `https://`. W przypadku pliku lokalnego musi zaczynać się od ścieżki `/mapfiles/`.
`useProxy`	`true` albo `false`	Łańcuch	Tak	Należy użyć opcji `true`, jeśli położeniem mapy jest adres URL. Należy użyć opcji `false`, jeśli położeniem mapy jest plik lokalny.
`id`	Id. mapy	Łańcuch	Tak	Użyj unikalnego identyfikatora dla każdej mapy.

Uwaga:

- Nie usuwaj domyślnych opcji **World (Built-in)** ani **Blank (Built-in)**.
- Należy pamiętać, aby używać przecinka między blokami treści, jak pokazano to w przykładzie, i upewnić się, że nowa sekcja `map.resources` jest poprawną treścią w formacie JSON.
- IBM SPSS Modeler Server (lub autonomiczny klient SPSS Modeler bez serwera) musi mieć dostęp do położenia mapy.

4. Uruchom serwer IBM SPSS Modeler Server.

Podczas pracy z wykresami map na panelu **Szczegóły** pojawi się nowa usługa mapy, która została dodana.

Wykresy krzywych matematycznych

Wykres krzywej matematycznej wykreśla matematyczne krzywe równań w oparciu o wyrażenia wprowadzone przez użytkownika.

Tworzenie prostego wykresu krzywej matematycznej

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Krzywa matematyczna**.
   Math curve

   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu krzywej matematycznej.
2. Wprowadź pierwszą wartość na osi X w polu **Początek wartości X**.
3. Wprowadź ostatnią wartość na osi X w polu **Koniec wartości X**.
4. Wprowadź równanie, które wykreśli krzywą na wykresie, w polu równania. Kliknij opcję Dodaj kolejną kolumnę, aby dodać dodatkowe równania.

Każde równanie traktuje x jako zmienną niezależną. Dostępne są następujące równania:

- +, -, *, /, % i ^
- abs(x)
- ceil(x)
- floor(x)
- log(x)
- max(a,b,c...)
- min(a,b,c...)
- random()
- round(x)
- sqrt(x)
- sin
- cos
- exp
- tan
- atan
- atan2
- asin
- acos

Funkcje dodatkowe

Początek wartości X
Pierwsza wartość na osi X.

Koniec wartości X
Ostatnia wartość na osi X.

wyrażenia
Równania wprowadzone przez użytkownika, które wykreślają krzywą wykresu.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy złożone

Wykresy złożone umożliwiają tworzenie wielu wykresów. Wykresy mogą być tego samego lub różnego typu, a ponadto mogą zawierać różne zmienne z tego samego zestawu danych.

Tworzenie prostego wykresu złożonego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres złożony**.
   Multi-chart
   
   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu złożonego.
2. Kliknij opcję **Dodaj następny wykres podrzędny**, aby dodać wykres podrzędny.
3. Wybierz typ wykresu z listy rozwijanej **Typ**.

4. W zależności od wybranego typu wykresu dostępne są następujące opcje:

   Wykresy słupkowe i kołowe
   Wybierz zmienną **Kategoria** z listy rozwijanej.

   Wykresy liniowe i wykresy rozrzutu
   Wybierz zmienną **osi X** z listy rozwijanej.

   Wybierz zmienną **osi Y** z listy rozwijanej.
5. Kliknij opcję **Dodaj następny wykres podrzędny**, aby uwzględnić dodatkowe typy wykresów.

Funkcje dodatkowe

Tytuł
Tytuł wykresu podrzędnego.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy dla wielu szeregów

Wykresy dla wielu szeregów są podobne do wykresów liniowych, ale na osi *Y* można wykreślić wiele zmiennych.

Tworzenie prostego wykresu dla wielu szeregów

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.
   Multi-series
   

   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu dla wielu szeregów.
2. Wybierz zmienną jako zmienną **osi X**.
3. Wybierz co najmniej dwie zmienne jako zmienne **osi Y**.

Funkcje dodatkowe

Oś X
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Oś Y
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *Y* wykresu.

Wybierz typ wykresu (liniowy, słupkowy lub rozrzutu) z listy rozwijanej.

Kliknij opcję **Dodaj kolejną kolumnę**, aby uwzględnić więcej kolumn na wykresie.

Styl szeregu
Udostępnia opcje definiowania orientacji osi Y. Dostępne opcje:

- Domyślny
- Odrębne osie Y
- Pokaż pomocniczą oś Y
- Podwójne osie Y

Normalizuj dane
Po włączeniu tego ustawienia dane są przekształcane do rozkładu normalnego, który umożliwia porównanie danych z wielu zbiorów danych lub z wielu kolumn. To ustawienie tworzy 100% zestawianie dla liczebności i przekształca statystyki na procenty.

Pomocnicza oś Y
Przedstawia zmienne ze zbioru danych, które są dostępne dla pomocniczej osi *Y* wykresu.

Wybierz typ wykresu (liniowy, słupkowy lub rozrzutu) z listy rozwijanej.

Kliknij opcję **Dodaj kolejną kolumnę**, aby uwzględnić więcej kolumn na wykresie.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy równoległe

Wykresy równoległe są użyteczne podczas wizualizacji geometrii wysokowymiarowych oraz podczas analizowania danych z wieloma zmiennymi. Wykresy równoległe przypominają wykresy liniowe dla danych szeregu czasowego, ale osi nie odpowiadają punktom w czasie (nie ma porządku naturalnego).

Tworzenie prostego wykresu równoległego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.


   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu równoległego.
2. Wybierz co najmniej dwie zmienne jako zmienne **kolumnowe**. Każda kolumna przedstawia pionową, równoległą oś na wykresie.

Uwaga: Porządek kolumn jest ważny dla odnalezienia właściwości. W przypadku typowej analizy danych może być konieczna wielokrotna zmiana porządku kolumn.

Funkcje dodatkowe

Kolumny
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *Y* wykresu.

Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe kolumny.

Odwzorowanie kolorów
Przedstawia dostępne zmienne odwzorowania kolorów. Zmienne te wykorzystują progresję kolorów, na podstawie zakresu wartości w określonej kolumnie, w celu przedstawienia ich w punktach wykresu. Odwzorowania kolorów są znane także jako kartogram.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy kołowe

Wykres kołowy jest przydatny do porównywania proporcji. Na przykład, wykresu kołowego można użyć do przedstawienia, że do określonej klasy włączono większą liczbę kobiet.

Tworzenie prostego wykresu kołowego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.


Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu kołowego.

2. Wybierz zmienną jakościową (nominalną lub porządkową) z listy **Kategoria**. Kategorie w tej zmiennej określają liczbę wycinków wykresu kołowego.
3. Wybierz statystyczną funkcję podsumowującą. W przypadku wykresów kołowych zwykle chcesz otrzymać statystykę opartą na liczebności lub sumę. Wynik statystyki określa rozmiar każdego wycinka.

Funkcje dodatkowe

Category
Wybierz zmienną jakościową (nominalną lub porządkową), która określa liczbę wycinków wykresu kołowego.

Podsumowanie
Wybierz statystyczną funkcję podsumowującą. W przypadku wykresów kołowych zwykle chcesz otrzymać statystykę opartą na liczebności lub sumę. Wynik statystyki określa rozmiar każdego wycinka.

Istnieją dwa typy funkcji podsumowania statystycznego. Rozróżnienie ich jest ważne, ponieważ od tego zależy, czy konieczne jest określenie zmiennej **wartości**.

- **Funkcje, które nie wymagają zmiennej wartości.** Są to funkcje, które nie wymagają zmiennej. W tej kategorii są wszystkie statystyki liczebności i procentowe. Te statystyki są dostępne, jeśli nie zdefiniowano zmiennej **wartości**.
- **Funkcje, które wymagają zmiennej wartości.** Są to funkcje, które wymagają zmiennej **wartości**. Na przykład funkcja *Średnia* wymaga zmiennej, na podstawie której obliczana jest średnia. Te statystyki są dostępne, jeśli zdefiniowano zmienną **wartości**.

Wartość
To pole jest wyświetlane, jeśli wybrano funkcję **Podsumowanie**, która wymaga zmiennej ilościowej. Wybierz zmienną, która ma pełnić rolę zmiennej ilościowej.

Typ wykresu kołowego
Dostępne są następujące style:

Normalny
Segmenty wykresu kołowego są wyświetlane jako normalne wycinki.

Pierścień
Segmenty wykresu kołowego są wyświetlane jako pierścienie. Ten styl jest również nazywany wykresem pierścieniowym.

Róża
W przeciwieństwie do normalnego wykresu kołowego, który używa wspólnego promienia, rozmiary segmentów wykresu różnią się w zależności od ich wartości.

Obszar róży
W przeciwieństwie do normalnego wykresu kołowego, który wykorzystuje wspólny promień, wielkości segmentu kołowego różnią się w zależności od ich obszaru.

Pierścień róży
W przeciwieństwie do normalnego wykresu kołowego, który wykorzystuje wspólny promień, wielkości segmentu kołowego różnią się w zależności od ich wartości, a segmenty są wyświetlane w postaci pierścienia.

Połowa róży
Podobnie jak w przypadku opcji Róża, ale wykres znajduje się na połowie koła.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Piramidy populacyjne

Piramidy populacyjne (znane również jako „piramidy wieku i płci”) to wykresy powszechnie używane do prezentowania i analizowania informacji na temat populacji w oparciu o wiek i płeć.

Tworzenie prostego wykresu piramidy populacyjnej

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Piramida populacyjna**.

Obszar roboczy zostaje zaktualizowany i wyświetla szablon piramidy populacyjnej.

2. Wybierz zmienną **osi Y** z listy rozwijanej.
3. Wybierz zmienną **Podziel według** z listy rozwijanej.

Funkcje dodatkowe

Oś Y
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *Y* wykresu.

Podziel według
Wybierz zmienną jakościową, która tworzy tabelę wykresów, z komórką dla każdej kategorii w zmiennej Podziel według. Podobnie jak w przypadku grupowania, opcja podziel według powoduje dodanie dodatkowych wymiarów do wykresu, wyświetlając informacje dla każdej kategorii zmiennej.

Szerokość przedziału
Suwak steruje rozmiarem przedziału, który jest używany do podziału danych na grupy.

Pokaż krzywą rozkładu
Gdy opcja ta jest włączona, na wykresie wyświetlana jest krzywa dopasowania rozkładu.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy K-K

Wykresy K-K (kwantyl-kwantyl) porównują dwa rozkłady prawdopodobieństw, wykreślając ich kwantyle względem siebie nawzajem. Wykres K-K służy do porównywania kształtów rozkładów, udostępniając graficzny widok przedstawiający podobieństwo właściwości, takich jak lokalizacja, skala i skośność, w dwóch rozkładach.

Tworzenie prostego wykresu K-K

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.

Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu K-K.

2. Wybierz zmienną jako zmienną **osi X**.

Funkcje dodatkowe

Oś X
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Rozkład
Lista rozwijana zawiera wszystkie dostępne metody rozkładu.

Rozkład z dopasowaniem automatycznym
Jest to ustawienie domyślne.

Beta
Zwraca wartość o rozkładzie Beta z określoną wartością parametrów kształtu.

Wykładniczy
Zwraca wartość o rozkładzie wykładniczym.

Gamma
Zwraca wartość o rozkładzie Gamma z określoną wartością parametrów kształtu i skali.

Logarytmiczno-normalny
Zwraca wartość o rozkładzie logarytmiczno-normalnym z określoną wartością parametrów.

Normalny
Zwraca liczbę o rozkładzie normalnym z określoną wartością parametrów: średnia oraz odchylenie standardowe.

Jednostajny
Zwraca wartość o rozkładzie jednostajnym między wartością minimalną a wartością maksymalną.

t Studenta
Zwraca wartość o rozkładzie t Studenta z określonymi stopniami swobody.

Typ wykresu
Wybierz wykres Q-Q (kwantyl-kwantyl) albo P-P (procent-procent).

Dopasowanie automatyczne
Gdy opcja ta jest włączona, parametry danych dla wybranego **rozkładu** są szacowane automatycznie. Gdy ta właściwość jest wyłączona, wyświetlane są wartości rozkładu: **Kształt** i **Skala**.

Kształt1
Ustawia wartość kształt1 dla rozkładu **Beta**. To ustawienie jest dostępne tylko wtedy, gdy opcja **Dopasowanie automatyczne** nie jest włączona, a jako **Rozkład** wybrano opcję **Beta**.

Kształt2
Ustawia wartość kształt2 dla rozkładu **Beta**. To ustawienie jest dostępne tylko wtedy, gdy opcja **Dopasowanie automatyczne** nie jest włączona, a jako **Rozkład** wybrano opcję **Beta**.

Kształt
Ustawia wartość kształtu dla wybranego rozkładu. To ustawienie jest dostępne tylko wtedy, gdy opcja **Dopasowanie automatyczne** nie jest włączona, a jako **Rozkład** wybrano opcję **Gamma** lub **Logarytmiczno-normalny**.

Skala
Ustawia wartość skali dla wybranego rozkładu. To ustawienie jest dostępne tylko wtedy, gdy opcja **Dopasowanie automatyczne** nie jest włączona, a jako **Rozkład** wybrano opcję **Wykładniczy**, **Gamma** lub **Logarytmiczno-normalny**.

Średnia
Ustawia wartość średnią dla rozkładu **normalnego**. To ustawienie jest dostępne tylko wtedy, gdy opcja **Dopasowanie automatyczne** nie jest włączona, a jako **Rozkład** wybrano opcję **Normalny**.

Odch. standard.
Ustawia wartość odchylenia standardowego dla rozkładu **normalnego**. To ustawienie jest dostępne tylko wtedy, gdy jako **Rozkład** wybrano opcję **Normalny**.

Min.
Ustawia wartość minimalną dla rozkładu **jednostajnego**. To ustawienie jest dostępne tylko wtedy, gdy opcja **Dopasowanie automatyczne** nie jest włączona, a jako **Rozkład** wybrano opcję **Jednostajny**.

Maks.
Ustawia wartość maksymalną dla rozkładu **jednostajnego**. To ustawienie jest dostępne tylko wtedy, gdy opcja **Dopasowanie automatyczne** nie jest włączona, a jako **Rozkład** wybrano opcję **Jednostajny**.

Stopnie swobody (df)
Ustawia wartość stopni swobody dla rozkładu **t Studenta**. To ustawienie jest dostępne tylko wtedy, gdy opcja **Dopasowanie automatyczne** nie jest włączona, a jako **Rozkład** wybrano opcję **t Studenta**.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy radarowe

Wykresy radarowe porównują wiele zmiennych ilościowych i ułatwiają dostrzeżenie, które zmienne mają podobne wartości, a które zmienne mają wartości odstające. Wykresy radarowe składają się z szeregu ramion, przy czym każde z nich reprezentuje jedną zmienną. Wykresy radarowe są również przydatne do określenia, które zmienne w zestawie danych uzyskują wysokie i niskie oceny.

Tworzenie prostego wykresu radarowego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Radarowy**.
   Radar
   
   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu radarowego.
2. Wybierz zmienną **Kolumny** z listy rozwijanej.

   Uwaga: Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe kolumny. Należy określić co najmniej trzy zmienne kolumn.

Funkcje dodatkowe

Category
Wybierz zmienną jakościową (nominalną lub porządkową). Jeśli jako wartości kategorii zostanie wybrana wartość **Brak**, wszystkie wartości zostaną wyświetlone osobno i nie zostanie zastosowana żadna metoda podsumowania.

Podsumowanie
Gdy wybrana jest zmienna jakościowa, dostępne są następujące statystyki podsumowujące:

Liczebność
Łączna liczba obserwacji.

Suma
Suma wartości.

Średnia
Średnia arytmetyczna; suma podzielona przez liczbę obserwacji.

Maksimum
Największa (najwyższa) wartość.

Minimum
Najmniejsza (najniższa) wartość.

Układ radaru
Określa układ obrazu będącego tłem dla wykresu radarowego:

Okrągły
Po wybraniu tej opcji wykres radarowy jest wyświetlany nad układem kołowym.

Wielokąt
Po wybraniu tej opcji wykres radarowy jest wyświetlany nad układem wielokątnym.

Podziel według
Wybierz zmienną jakościową, która tworzy tabelę wykresów, z komórką dla każdej kategorii w zmiennej Podziel według. Podobnie jak w przypadku grupowania, opcja podziel według powoduje dodanie dodatkowych wymiarów do wykresu, wyświetlając informacje dla każdej kategorii zmiennej.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy relacji

Wykres relacji jest użyteczny do określania w jaki sposób zmienne wiążą się ze sobą.

Tworzenie prostego wykresu relacji

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.


   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu relacji.
2. Wybierz co najmniej dwie zmienne jako zmienne **kolumnowe**.

   Uwaga: Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe zmienne kolumn.

Funkcje dodatkowe

Kolumny
Przedstawia dostępne zmienne ze zbioru danych.

Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe kolumny.

Styl linii
Umożliwia określenie styli linii między powiązanymi punktami danych.

Krzywa
Po wybraniu między powiązanymi punktami danych rysowane są krzywe.

Prosta
Po wybraniu między powiązanymi punktami danych rysowane są linie proste.

Etykieta wartości progowej
Wyświetla etykiety dla punktów danych, których wartości przekraczają zdefiniowaną wartość.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy rozrzutu i wykresy punktowe

Istnieje kilka szerokich kategorii wykresów tworzonych z punktowym elementem graficznym:

- **Wykresy rozrzutu.** Są użyteczne do wykreślania danych z wieloma zmiennymi. Pomagają określić potencjalne relacje między zmiennymi ilościowymi. W prostym wykresie rozrzutu do wykreślenia dwóch zmiennych używany jest układ współrzędnych 2-W. Na wykresie rozrzutu 3-W do wykreślenia trzech zmiennych używany jest układ współrzędnych 3-W. Gdy trzeba wykreślić więcej zmiennych, można spróbować nakładać wykresy rozrzutu i macierze wykresu rozproszenia (SPLOM). Nakładany wykres rozrzutu umożliwia nałożenie par zmiennych *x*-*y*, z których każda ma inny kolor lub kształt. Funkcja

SPLOM umożliwia utworzenie macierzy wykresów rozrzutu 2-W, w których każda zmienna jest kreślona w funkcji każdej pozostałej zmiennej w SPLOM.

- **Wykresy punktowe.** Podobnie jak histogramy, są one użyteczne do przedstawiania rozkładu pojedynczej zmiennej ilościowej. Dane są dzielone, ale, zamiast jednej wartości dla każdego przedziału (jak liczebność), wyświetlane i zestawiane są wszystkie punkty w każdym przedziale. Wykresy te nazywane są czasami wykresami gęstości.
- **Podsumowujące wykresy punktowe.** Są one takie jak wykresy słupkowe, ale punkty są wyświetlane w miejscu, w którym znajdowałby się szczyt słupka. Ze względy na to, że podsumowujący wykres punktowy jest tak podobny do wykresu słupkowego, aby uzyskać informacje na temat jego tworzenia, należy zapoznać się z sekcją "Wykresy słupkowe" na stronie 70.
- **Wykresy linii rzutowania.** Jest to specjalny typ podsumowującego wykresu punktowego. Punkty są grupowane, a przez punkty w każdej kategorii rysowana jest linia. Wykres linii rzutowania jest przydatny do porównywania statystyk wśród zmiennych jakościowych.

Tworzenie prostego wykresu rozrzutu

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.

   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu rozrzutu.
2. Wybierz zmienną ilościową jako zmienną **osi X**.
3. Wybierz zmienną skali jako zmienną **osi Y**. Określanie statystyki nie jest konieczne, gdyż wykresy rozrzutu zwykle przedstawiają wartości surowe.

Funkcje dodatkowe

Oś X
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *X* wykresu.

Oś Y
Przedstawia zmienne ze zbioru danych, które są dostępne dla osi *Y* wykresu.

Odwzorowanie kolorów
Przedstawia dostępne zmienne odwzorowania kolorów. Zmienne te wykorzystują progresję kolorów, na podstawie zakresu wartości w określonej kolumnie, w celu przedstawienia ich w punktach wykresu. Odwzorowania kolorów są znane także jako kartogram.

Odwzorowanie wielkości
Przedstawia dostępne zmienne odwzorowania wielkości. Zmienne te wykorzystują różne wielkości do przedstawiania ich w punktach wykresu.

Odwzorowanie kształtów
Przedstawia dostępne zmienne odwzorowania kształtów. Zmienne te wykorzystują różne kształty do przedstawiania ich w punktach wykresu.

Linia dopasowania
W linii dopasowania punkty danych są dopasowywane do linii, która zwykle nie przechodzi przez wszystkie punkty danych. Linia dopasowania przedstawia trend danych. Niektóre linie dopasowania opierają się na regresji. Inne opierają się na metodzie ważonych najmniejszych kwadratów. Wybierz opcję linii dopasowania z listy rozwijanej.

Gradient na wykresie bąbelkowym
Przełącznik umożliwia włączenie/wyłączenie wyświetlania gradientu kolorów efektów 3-W na wykresach bąbelkowych. To ustawienie jest wyłączone, gdy wybrana jest zmienna **Odwzorowanie kolorów**.

Minimalny rozmiar bąbelka
Ustawia minimalny rozmiar bąbelka. Wprowadź wartość z zakresu od 5 do 20.

Maksymalny rozmiar bąbelka
Ustawia maksymalny rozmiar bąbelka. Wprowadź wartość z zakresu od 20 do 80.

Pokaż linię odniesienia
Gdy ta opcja jest włączona, na wykresie wyświetlana jest linia odniesienia oparta na określonych wartościach osi X i osi Y.

Wprowadź wartość linii odniesienia na osi X
Gdy ustawienie **Pokaż linię odniesienia** jest włączone, umożliwia określenie wartości linii odniesienia dla osi X. Kliknij opcję **Dodaj kolejną kolumnę**, aby określić dodatkowe wartości linii odniesienia.

Wprowadź wartość linii odniesienia na osi Y
Gdy ustawienie **Pokaż linię odniesienia** jest włączone, umożliwia określenie wartości linii odniesienia dla osi Y. Kliknij opcję **Dodaj kolejną kolumnę**, aby określić dodatkowe wartości linii odniesienia.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy macierzy wykresów rozrzutu

Macierze wykresów rozrzutu są dobrym sposobem na określenie, czy między wieloma zmiennymi istnieją korelacje liniowe.

Tworzenie wykresu macierzy wykresu rozrzutu

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres macierzy wykresu rozrzutu**.


   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu macierzy wykresu rozrzutu.
2. Wybierz co najmniej dwie skalowane zmienne **kolumnowe**.

   Uwaga: Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe zmienne kolumn.

Każda wybrana zmienna jest wykreślana względem każdej innej zmiennej w celu utworzenia macierzy pojedynczych wykresów rozrzutu.

Funkcje dodatkowe

Kolumny
Wybierz co najmniej dwie zmienne macierzowe. Zmienne te muszą być liczbowe (ale nie mogą mieć formatu daty).

Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe kolumny.

Odwzorowanie kolorów
Przedstawia dostępne zmienne odwzorowania kolorów. Zmienne te wykorzystują progresję kolorów, na podstawie zakresu wartości w określonej kolumnie, w celu przedstawienia ich w punktach wykresu. Odwzorowania kolorów są znane także jako kartogram.

Korelacja
Po włączeniu dla wybranych zmiennych wyświetlane są informacje o korelacji liniowej (silna, średnia, słaba).

Pokaż krzywą kde
Gdy opcja ta jest włączona, na wykresie wyświetlana jest krzywa estymatora jądrowego gęstości.

Gradient na wykresie bąbelkowym
Przełącznik umożliwia włączenie/wyłączenie wyświetlania gradientu kolorów efektów 3-W na wykresach bąbelkowych. To ustawienie jest wyłączone, gdy wybrana jest zmienna **Odwzorowanie kolorów**.

Minimalny rozmiar bąbelka
Ustawia minimalny rozmiar bąbelka. Wprowadź wartość z zakresu od 5 do 20.

Maksymalny rozmiar bąbelka
Ustawia maksymalny rozmiar bąbelka. Wprowadź wartość z zakresu od 20 do 80.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy pierścieniowe

Wykres pierścieniowy służy do wizualizacji danych o strukturze hierarchicznej. Wykres pierścieniowy składa się z wewnętrznego okręgu otoczonego pierścieniami o głębszych poziomach hierarchii. Kąt każdego segmentu jest proporcjonalny do wartości lub podzielony na równe części pod jego wewnętrznym segmentem. Segmenty wykresu są oznaczone kolorami na podstawie kategorii lub poziomu hierarchicznego, do którego należą.

Tworzenie prostego wykresu pierścieniowego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Pierścieniowy**.
   Sunburst
   
   Obszar roboczy zostanie zaktualizowany w celu wyświetlenia szablonu wykresu pierścieniowego.
2. Wybierz zmienną jakościową (nominalną lub porządkową) z listy **Kolumny**. Kategorie w tej zmiennej określają liczbę segmentów na wykresie.
3. Kliknij opcję **Dodaj kolejną kolumnę** i wybierz inną zmienną jakościową (nominalną lub porządkową) z listy**Kolumny**. Kategorie w tej zmiennej określają liczbę segmentów w drugim pierścieniu wykresu i reprezentują poziom hierarchiczny.
4. Wybierz funkcję podsumowania statystycznego dla elementu graficznego (statystyka bazująca na liczebności lub suma). Wynik statystyki określa rozmiar każdego segmentu. W przypadku zaznaczenia opcji **Suma** wybierz zmienną ilościową z listy **Wartość**, aby odzwierciedlić wartość w zbiorze danych do podsumowania.
5. Wybierz opcję **Układ pierścienia** (**Tradycyjny** lub **Rozbieżny**).

Funkcje dodatkowe

Kolumny
Wybierz zmienną jakościową (nominalną lub porządkową), która określa liczbę segmentów na wykresie.

Podsumowanie
Wybierz funkcję podsumowania statystycznego dla elementu graficznego (statystyka bazująca na liczebności lub suma). Wynik statystyki określa rozmiar każdego wycinka.

Istnieją dwa typy funkcji podsumowania statystycznego. Rozróżnienie ich jest ważne, ponieważ od tego zależy, czy konieczne jest określenie zmiennej **wartości**.

- **Funkcje, które nie wymagają zmiennej wartości.** Są to funkcje, które nie wymagają zmiennej **wartości**. W tej kategorii są wszystkie statystyki liczebności i procentowe. Te statystyki są dostępne, jeśli nie zdefiniowano zmiennej **wartości**.
- **Funkcje, które wymagają zmiennej wartości.** Są to funkcje, które wymagają zmiennej **wartości**. Na przykład funkcja *Suma* wymaga zmiennej, dla której obliczane jest podsumowanie.

Wartość
To pole jest wyświetlane, jeśli wybrano funkcję **Podsumowanie**, która wymaga zmiennej ilościowej. Wybierz zmienną, która ma pełnić rolę zmiennej ilościowej.

Układ pierścienia
Dostępne opcje to: **Tradycyjny** i **Rozbieżny**.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

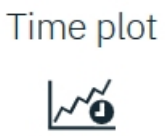
Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy sekwencyjne

Wykres sekwencyjny przedstawia punkty danych po kolejnych odstępach czasu. Wykreślany szereg czasowy musi zawierać wartości liczbowe. Zakłada się ich występowanie w przedziale czasu, w którym okresy są równomierne. Wykresy sekwencyjne pozwalają na wstępną analizę charakterystyk szeregów czasowych pod względem podstawowych statystyk i testów, a tym samym dostarczają przydatnych informacji na temat danych jeszcze przed rozpoczęciem modelowania. Wykresy sekwencyjne uwzględniają takie metody analizy, jak dekompozycja, rozszerzony test Dickeya-Fullera (ADF), korelacje (ACF/PACF) i analiza spektralna.

Tworzenie prostego wykresu sekwencyjnego

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres sekwencyjny**.



Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu sekwencyjnego.

2. Z listy rozwijanej **Wartości** wybierz wartość dla osi Y.

Funkcje dodatkowe

Data
Wybierz datę z listy rozwijanej (jeśli ma zastosowanie). Każda obserwacja jest zasadniczo oddzielona od pozostałych tym samym odstępem czasu. Jeśli zostanie wybrana zmienna daty, pojawi się opcja ponownego próbkowania. Można jej użyć do agregacji zmiennych wartości w celu dopasowania ich do określonego przedziału czasu.

Algorytm wykresu sekwencyjnego
Algorytm wykresu sekwencyjnego używany do analizy danych szeregów czasowych:

- **Dekompozycja.** Zdekomponuj szereg czasowy na trzy składniki (cykl trendu, sezonowy i nieregularny). Dekompozycja jest wykonywana w sposób addytywny.
- **Test ADF.** Rozszerzony test Dickeya-Fullera (ADF) testuje hipotezę zerową, zgodnie z którą w szeregu występuje pierwiastek jednostkowy, a szereg nie jest stacjonarny. Jeśli test odrzuci hipotezę zerową, oznacza to, że szereg jest stacjonarny lub może być przedstawiony jako stacjonarny przy użyciu modelu różnic.
- **ACF/PACF.** Korelacje szeregów.
- **Analiza spektralna.** Narzędzie analityczne w dziedzinie częstotliwości. Najwyższa częstotliwość jest oznaczona rombem.

Zmień położenie wykresu
Odwraca położenia wykresów w układzie współrzędnych prostokątnych oraz biegunowych.

Pokaż punkt zwrotu
Uwidacznia lub ukrywa punkty zwrotu na wykresach. Na podstawie wyników dekompozycji szeregu czasowego badamy, czy szereg wykazuje jeden ogólny trend, czy też ma kilka punktów zwrotu, które zmieniają kierunek trendu.

Pokaż wartość odstającą
Umożliwia wyświetlenie lub ukrycie wartości odstających. Wartości odstające szeregów czasowych są analizowane na podstawie składnika nieregularnego dekompozycji szeregu czasowego.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

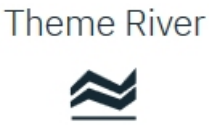
Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy Rzeka tematyczna

Rzeka tematyczna to specjalny wykres przepływowy, który obrazuje zmiany zachodzące z czasem.

Tworzenie prostego wykresu rzeki tematycznej

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Rzeka tematyczna**.



   Obszar roboczy zostaje zaktualizowany i wyświetla szablon rzeki tematycznej.
2. Wybierz zmienną **Oś X**.
3. Wybierz zmienną **kategorii**.

   Ograniczenie: Jeśli zmienna kategorii ma więcej niż 50 odrębnych kategorii, jako zdarzenia na wykresie zostanie użytych tylko 50 najwyższych kategorii.

Funkcje dodatkowe

Kolejność na podstawie
W zależności od wybranej wartości osi x możliwe jest określenie, czy kolejność kategorii zależy od nazwy kategorii czy od wartości kategorii.

Porządek kategorii
W zależności od wybranej wartości osi x możliwe jest określenie, czy kolejność kategorii jest rosnąca, malejąca, czy zależna od kolejności odczytu z zestawu danych.

Podsumowanie
Wybierz funkcję podsumowania statystycznego. Jest to metoda tworzenia podsumowania każdej kategorii.

Istnieją dwa typy funkcji podsumowania statystycznego. Rozróżnienie ich jest ważne, ponieważ od tego zależy, czy konieczne jest określenie zmiennej **wartości**.

- **Funkcje, które nie wymagają zmiennej wartości.** Są to funkcje, które nie wymagają zmiennej. W tej kategorii są wszystkie statystyki liczebności i procentowe. Te statystyki są dostępne, jeśli nie zdefiniowano zmiennej **wartości**.
- **Funkcje, które wymagają zmiennej wartości.** Są to funkcje, które wymagają zmiennej **wartości**. Na przykład funkcja *Średnia* wymaga zmiennej, na podstawie której obliczana jest średnia. Te statystyki są dostępne, jeśli zdefiniowano zmienną **wartości**.

Wartość
To pole jest wyświetlane, gdy wybrana jest funkcja **Podsumowanie**, która wymaga wartości zmiennej. Wybierz zmienną, która będzie służyła jako wartość.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

wykresy drzewa

Wykresy drzewa reprezentują hierarchię w strukturze podobnej do drzewa. Struktura wykresu drzewa składa się z węzła głównego (nie ma węzła nadrzędnego), linii łączących nazywanych gałęziami (reprezentują relacje i połączenia między elementami) oraz węzłów-liści (nie mają węzłów podrzędnych).

Tworzenie prostego wykresu drzewa

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Typy wykresów**.
   Tree
   

   Obszar roboczy zostanie zaktualizowany w celu wyświetlenia szablonu wykresu drzewa.
2. Wybierz zmienną **Kolumny** z listy rozwijanej.

   Uwaga: Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe zmienne kolumn.

Funkcje dodatkowe

Kolumny
Przedstawia zmienne ze zbioru danych, które są dostępne do przypisania do kolumn wykresu.

Podsumowanie
Wybierz statystyczną funkcję podsumowującą. Wynik statystyki określa położenie elementów graficznych.

Istnieją dwa typy funkcji podsumowania statystycznego. Rozróżnienie ich jest ważne, ponieważ od tego zależy, czy konieczne jest określenie zmiennej **wartości**.

- **Funkcje, które nie wymagają zmiennej wartości.** Są to funkcje, które nie wymagają wartości zmiennej. W tej kategorii są wszystkie statystyki liczebności i procentowe. Te statystyki są dostępne, jeśli nie zdefiniowano zmiennej **wartości**.
- **Funkcje, które wymagają zmiennej wartości.** Są to funkcje, które wymagają zmiennej **wartości**. Na przykład funkcja *Średnia* wymaga zmiennej, na podstawie której obliczana jest średnia. Te statystyki są dostępne, jeśli zdefiniowano zmienną **wartości**.

Wartość
To pole jest wyświetlane, gdy wybrana jest funkcja **Podsumowanie**, która wymaga wartości zmiennej. Wybierz zmienną, która ma służyć jako baza wartości podsumowania.

Układ drzewa

Od lewej do prawej
Węzeł główny jest wyświetlany po lewej stronie, a węzły-liście są wyświetlane po prawej stronie.

Od prawej do lewej
Węzeł główny jest wyświetlany po prawej stronie, a węzły-liście są wyświetlane po lewej stronie.

Od góry do dołu
Węzeł główny jest wyświetlany na górze, a węzły-liście są wyświetlane na dole.

Od dołu do góry
Węzeł główny jest wyświetlany na dole, a węzły-liście są wyświetlane na górze.

Promieniowy
Węzeł główny jest wyświetlany w środku, a węzły-liście promieniują od węzła głównego.

Głębokość liści
Ustawia liczbę poziomów, na których tworzone są węzły-liście.

Pokaż etykiety liści
Gdy ta opcja jest włączona, wyświetlane są etykiety poszczególnych węzłów-liści.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

mapy drzew

Mapy drzew są alternatywną metodą wizualizowania hierarchicznej struktury drzewiastej, a także prezentowanie ilości dla każdej kategorii. Mapy drzew są szczególnie przydatne do wykrywania wzorców i schematów w danych. Gałęzie drzewa są reprezentowane przez prostokąty, a każda podgałąź jest reprezentowana przez mniejsze prostokąty.

Tworzenie prostej mapy drzewa

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Mapa drzewa**.


   Obszar roboczy zostaje zaktualizowany i wyświetla szablon mapy drzewa.
2. Wybierz zmienną **Kolumny** z listy rozwijanej.

   Uwaga: Kliknij opcję **Dodaj kolejną kolumnę**, aby dodać dodatkowe zmienne kolumn.

Funkcje dodatkowe

Kolumny
Przedstawia zmienne ze zbioru danych, które są dostępne do przypisania do kolumn wykresu.

Podsumowanie
Wybierz statystyczną funkcję podsumowującą. Wynik statystyki określa położenie elementów graficznych.

Istnieją dwa typy funkcji podsumowania statystycznego. Rozróżnienie ich jest ważne, ponieważ od tego zależy, czy konieczne jest określenie zmiennej **wartości**.

- **Funkcje, które nie wymagają zmiennej wartości.** Są to funkcje, które nie wymagają wartości zmiennej. W tej kategorii są wszystkie statystyki liczebności i procentowe. Te statystyki są dostępne, jeśli nie zdefiniowano zmiennej **wartości**.
- **Funkcje, które wymagają zmiennej wartości.** Są to funkcje, które wymagają zmiennej **wartości**. Na przykład funkcja *Średnia* wymaga zmiennej, na podstawie której obliczana jest średnia. Te statystyki są dostępne, jeśli zdefiniowano zmienną **wartości**.

Wartość
To pole jest wyświetlane, gdy wybrana jest funkcja **Podsumowanie**, która wymaga wartości zmiennej. Wybierz zmienną, która ma służyć jako baza wartości podsumowania.

Głębokość liści
Ustawia liczbę poziomów, na których tworzone są węzły-liście.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy t-SNE

Stochastyczna metoda porządkowania sąsiadów w oparciu o rozkład t (t-SNE) to algorytm uczenia maszynowego dla wizualizacji. Wykresy t-SNE modelują każdy obiekt wielowymiarowy poprzez dwu lub trójwymiarowy w taki sposób, że obiekty podobne są modelowane poprzez sąsiednie punkty, a obiekty niepodobne poprzez odległe punkty z wysokim prawdopodobieństwem.

Tworzenie prostego wykresu t-SNE

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.


   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu t-SNE.
2. Ustaw wartości **Zmieszanie**, **Współczynnik uczenia** i **Maksymalna liczba iteracji**.
3. Opcjonalnie wybierz zmienną **odwzorowania kolorów**.

Funkcje dodatkowe

Zmieszanie
Ustawia liczbę, która ustanawia przypuszczenie dotyczące liczby bliskich sąsiadów dla każdego punktu danych. Ma to na celu zrównoważenie lokalnych i globalnych aspektów danych.

Współczynnik uczenia
Ta wartość wpływa na szybkość uczenia poprzez określanie zmiany rozmiaru wagi przy każdej iteracji.

Maksymalna liczba iteracji
Maksymalna liczba wykonywanych iteracji.

Odwzorowanie kolorów
Przedstawia dostępne zmienne odwzorowania kolorów. Zmienne te wykorzystują progresję kolorów, na podstawie zakresu wartości w określonej kolumnie, w celu przedstawienia ich w punktach wykresu. Odwzorowania kolorów są znane także jako kartogram.

Tytuł główny
Tytuł wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Wykresy chmury słów

Wykresy chmury słów przedstawiają dane jako słowa, przy czym rozmiar i umieszczenie pojedynczego słowa jest określany przez sposób jego ważenia.

Tworzenie prostego wykresu chmury słów

1. W sekcji **Typy wykresów** kreatora wykresów kliknij ikonę **Wykres K-K**.


   Obszar roboczy zostaje zaktualizowany i wyświetla szablon wykresu chmury słów.
2. Wybierz zmienną jako zmienną **źródłową**. Każda kategoria zmiennej jest przedstawiana na wykresie na podstawie jej wartości wagi.

3. Wybierz wartość **kształtu** dla wykresu. Generowane dane wykresu są prezentowane w wybranym kształcie.

Funkcje dodatkowe

Źródło
Przedstawia zmienne ze zbioru danych, które są dostępne jako źródło wykresu. Każda kategoria zmiennej jest przedstawiana na wykresie na podstawie jej wartości wagi.

Kształt
Przedstawia dostępne kształty wykresów. Generowane dane wykresu są prezentowane w wybranym kształcie.

Tytuł główny
Tytuł wykresu.

Podtytuł
Podtytuł wykresu wyświetlany bezpośrednio pod tytułem wykresu.

Przypis
Przypis wykresu, który jest wyświetlany pod wykresem.

Kokpit

Istnieje możliwość utworzenia układu kokpitu wykresów, który będzie zawierał kilka wykresów jednocześnie. Układy można zapisać jako szablony, a następnie przeciągać i upuszczać zapisane wykresy w celu rozmieszczenia ich w układzie.

1. Po kliknięciu węzła danych prawym przyciskiem myszy i wybraniu opcji **Wyświetl dane** kliknij element sterujący Kokpit w sekcji **Działania**.

Rysunek 14. Element sterujący Kokpit

Kokpit zawiera i udostępnia następujące ustawienia.

Szablon
W tej sekcji można tworzyć nowe szablony układu lub wybierać spośród predefiniowanych bądź zapisanych szablonów układu.

Wybierz szablon układu
Dokonaj wyboru spośród dostępnych szablonów układów lub rozpocznij z nowym szablonem.

Actions
Kliknij ikonę **Edytuj układ**, aby edytować wybrany szablon układu. Możesz też zaimportować plik kokpitu. Po zakończeniu zapisz szablon.

Po zakończeniu kliknij ikonę **Wyjdź z trybu edytowania układu**.

Zawartość
Ta sekcja umożliwia przeciąganie i upuszczanie elementów do układu kokpitu.

Wybierz zapisany wykres
W tej sekcji znajduje się lista zapisanych wykresów. Przeciągaj i upuszczaj wykresy w żądane miejsca w układzie kokpitu.

Wybierz obiekt
Można także przeciągnąć tekst HTML lub obrazy do układu kokpitu.

Preferencje wizualizacji globalnej

Ustawienia domyślne dla tytułów, suwaka rozstępu, linii siatki i śledzenia myszy można zastąpić. Można również określić inny kolor szablonu schematu.

1. Po kliknięciu węzła danych prawym przyciskiem myszy i wybraniu opcji **Wyświetl dane** kliknij element sterujący globalnymi preferencjami dotyczącymi wizualizacji w sekcji **Czynności**.

Rysunek 15. Przycisk preferencji wizualizacji globalnej

Okno dialogowe preferencji wizualizacji globalnej wyświetla i zawiera następujące ustawienia.

Tytuły
Ta sekcja zawiera ustawienia globalne tytułów wykresów.

Tytuły globalne
Włącza lub wyłącza globalne tytuły dla wszystkich wykresów.

Globalny tytuł główny
Włącza lub wyłącza widok globalnych, głównych tytułów wykresów. Gdy opcja ta jest włączona, wprowadzony tutaj tytuł wykresu najwyższego poziomu jest stosowany do wszystkich wykresów, skutecznie nadpisując poszczególne ustawienie **Tytuł główny** każdego wykresu.

Podtytuł globalny
Włącza lub wyłącza widok globalnych, głównych podtytułów wykresów. Gdy opcja ta jest włączona, wprowadzony tutaj podtytuł wykresu najwyższego poziomu jest stosowany do wszystkich wykresów, skutecznie nadpisując poszczególne ustawienie **Podtytuł** każdego wykresu.

Tytuły domyślne
Włącza lub wyłącza domyślne tytuły dla wszystkich wykresów.

Narzędzia
Ta sekcja zawiera opcje, które umożliwiają sterowanie zachowaniem wykresu.

Suwak rozstępu
Włącza lub wyłącza suwak rozstępu dla każdego wykresu. Gdy opcja ta jest włączona, można kontrolować ilość wyświetlanych danych wykresu za pomocą suwaka rozstępu, który znajduje się pod każdym wykresem.

Linie siatki
Kontroluje widok linii siatki osi X (pionowa) i osi Y (pozioma).

Lokalizator myszy
Gdy opcja ta jest włączona, lokalizacja kursora myszy, względem danych wykresu, jest śledzona i wyświetlana po umieszczeniu go w dowolnym miejscu nad wykresem.

Zestaw narzędzi
Włącza lub wyłącza zestaw narzędzi dla każdego wykresu. W zależności od typu wykresu zestaw narzędzi po prawej stronie ekranu zawiera takie narzędzia, jak powiększanie, zapisywanie w formie obrazu, odtwarzanie, wybór danych i czyszczenie zaznaczenia.

ARIA
Gdy ta opcja jest włączona, treści i aplikacje internetowe są bardziej przystępne dla użytkowników niepełnosprawnych.

Odfiltruj puste
Włącza lub wyłącza filtrowanie pustych danych wykresu.

Oś X w punkcie zero
Gdy ta opcja jest włączona, oś X leży w pozycji wyjściowej drugiej osi wykresu. Gdy ta opcja jest wyłączona, oś X zawsze rozpoczyna się od 0.

Oś Y w punkcie zero
Gdy ta opcja jest włączona, oś Y leży w pozycji wyjściowej drugiej osi wykresu. Gdy ta opcja jest wyłączona, oś Y zawsze rozpoczyna się od 0.

Pokaż etykietę osi X
Włącza lub wyłącza etykietę osi X.

Pokaż etykietę osi Y
Włącza lub wyłącza etykietę osi Y.

Pokaż linię osi X
Włącza lub wyłącza linię osi X.

Pokaż linię osi Y
Włącza lub wyłącza linię osi Y.

Kompozycja
Wybierz szablon, aby zmienić kolory używane na wykresach, które mają zmienną grupującą lub zestawiającą. Wszelkie atrybuty elementów zdefiniowane w wybranym pliku szablonu nadpiszą ustawienia domyślnego szablonu dla tych atrybutów elementów.

2. Kliknij przycisk **Zastosuj**, aby zapisać ustawienia lub przycisk **Anuluj**, aby odrzucić zmiany.

Rozdział 7. Praca z wynikami

Po uruchomieniu niektórych strumieni wyniki zostaną udostępnione w obszarze Okno raportu, na karcie **Model** lub **Zaawansowane** w węzłach modelu użytkowego. W oknie raportów można łatwo przejść do potrzebnego wyniku. Okno to umożliwia także operowanie wynikiem oraz utworzenie dokumentu zawierającego tylko wybrane wyniki. Niektóre wyniki graficzne wykorzystują także obszar Okno raportu.

Obszar Okno raportu jest używany także do obsługi następujących wyników tworzonych przez produkt IBM SPSS Modeler:

- Modele użytkowe TCM
- Modele użytkowe STP
- Wartościowe informacje z modelu użytkowego Dwustopniowa-AS
- Model użytkowy GSAR
- Węzeł wykresu z wizualizacją na mapie

Okno raportu

Wyniki są wyświetlane w obszarze Okno raportu. Obszaru Okno raportu można używać do:

- Przeglądanie wyników
- Wyświetlanie lub ukrywanie wybranych tabel i wykresów
- Zmiana kolejności wyświetlania wyników poprzez przemieszczanie wybranych elementów
- Przenoszenie elementów między obszarem Okno raportu a innymi aplikacjami.

Obszar Okno raportu jest podzielony na dwa okna:

- Lewe okno zawiera widok struktury zawartości raportu.
- Prawe okno zawiera tabele, wykresy i wyniki tekstowe działania procedur statystycznych.

Po kliknięciu wybranego elementu w strukturze następuje przejście bezpośrednio do odpowiedniej tabeli lub wykresu. Możliwe jest kliknięcie i przeciągnięcie prawej krawędzi okna struktury, aby zmienić jego szerokość.

Wyświetlanie i ukrywanie wyników

Obszar Okno raportu pozwala na selektywne wyświetlanie i ukrywanie poszczególnych tabel lub wyników całej procedury. Możliwość taka jest użyteczna, gdy zachodzi konieczność skrócenia ilości widocznych elementów wynikowych w oknie raportu.

Ukrywanie tabeli lub wykresu bez ich usuwania

1. Kliknij dwukrotnie ikonę książki w oknie struktury Edytora raportów.

   lub

2. Kliknij element, aby go zaznaczyć.
3. Z menu wybierz:

   Widok > **Ukryj**

   lub

4. Kliknij ikonę zamkniętej książki (Ukryj) na pasku narzędzi Struktura.

Aktywna stanie się ikona otwartej książki (Pokaż), informując, że dany element jest aktualnie ukryty.

Ukrywanie wszystkich wyników działania procedury

1. Kliknij pole znajdujące się z lewej strony nazwy procedury w oknie struktury raportu.

Spowoduje to ukrycie wszystkich wyników danej procedury i zwinięcie widoku struktury.

Przenoszenie, usuwanie i kopiowanie wyników

Możliwa jest zmiana ułożenia wyników przez kopiowanie, przenoszenie lub usuwanie elementów albo ich grup.

Przenoszenie wyników w oknie raportu

1. Zaznacz elementy w strukturze lub w oknie raportu.
2. Przeciągnij zaznaczone elementy i upuść je w innym miejscu.

Element, który został wycięty w kroku 1, zostanie przeniesiony do wybranego miejsca.

Usuwanie wyników w oknie raportów

1. Zaznacz elementy w strukturze lub w oknie raportu.
2. Naciśnij klawisz **Delete**.

   lub
3. Z menu wybierz:

   Edycja > **Usuń**

Zmiana początkowego wyrównania

Domyślnie, wszystkie wyniki są początkowo wyrównane do lewej strony. Aby zmienić początkowe wyrównanie nowych elementów wyników:

1. Z menu wybierz:

   Edycja > **Opcje**
2. Kliknij kartę **Raporty**.
3. W grupie Stan raportu wynikowego zaznacz typ elementu (np. tabela przestawna, wykres, wyniki tekstowe).
4. Wybierz żądaną opcję wyrównania.

Zmiana wyrównania wyników

1. W oknie struktury lub oknie raportu zaznacz elementy, które mają zostać wyrównane.
2. Z menu wybierz:

   Format > **Wyrównanie do lewej**

   lub

   Format > **Wyśrodkowane**

   lub

   Format > **Wyrównanie do prawej**

Struktura obszaru Okno raportu

Okno struktury zawiera spis zawartości dokumentu w obszarze Okno raportu Okna struktury można używać do poruszania się pomiędzy wynikami procedur i do sterowania sposobem ich wyświetlania. Większość czynności podejmowanych w oknie struktury ma wpływ na zawartość okna raportu.

- Wybranie elementu w oknie struktury powoduje wyświetlenie odpowiedniego elementu w oknie raportu.
- Przeniesienie elementu w oknie struktury powoduje przeniesienie odpowiedniego elementu w oknie raportu.
- Zwinięcie widoku struktury powoduje ukrycie rezultatów dla wszystkich elementów znajdujących się w zwiniętych poziomach.

Sterowanie sposobem wyświetlania struktury

Aby sterować wyświetlaniem struktury, można wykonać następujące czynności:

- Rozwijać i zwijać widok struktury.
- Zmieniać poziom struktury dla wybranych elementów.
- Zmieniać rozmiar elementów w widoku struktury.
- Zmieniać czcionki używane w widoku struktury.

Zwijanie i rozwijanie widoku struktury

1. Kliknij pole znajdujące się z lewej strony elementu struktury, który ma zostać zwinięty lub rozwinięty.

   lub
2. Kliknij element w strukturze.
3. Z menu wybierz:

   Widok > **Zwiń**

   lub

Widok > **Rozwiń**

Zmiana poziomu struktury

1. Kliknij element w oknie struktury.
2. Z menu wybierz:

   Edycja > **Struktura** > **Podnieś poziom**

   lub

   Edycja > **Struktura** > **Obniż poziom**

Zmiana rozmiaru elementów w strukturze

1. Z menu wybierz:

   Widok > **Rozmiar struktury**
2. Wybierz rozmiar struktury (**Mała**, **Średnia** lub **Duża**).

Zmiana czcionek w strukturze

1. Z menu wybierz:

   Widok > **Czcionka struktury...**
2. Wybierz czcionkę.

Edytowanie lub dodawanie elementów w obszarze Okno raportu

Okno raportu pozwala na dodawanie elementów, takich jak tytuły, nowy tekst, wykresy lub dane z innych aplikacji.

Dodawanie tytułu lub elementu tekstowego

Do obszaru Okno raportu można dodawać elementy tekstowe, które nie są połączone z tabelą ani z wykresem.

1. Kliknij tabelę, wykres lub inny obiekt, który będzie poprzedzać tytuł lub tekst.
2. Z menu wybierz:

   Wstaw > **Nowy tytuł**

   lub

   Wstaw > **Nowy tekst**
3. Kliknij dwukrotnie nowy obiekt.
4. Wprowadź tekst.

Dodawanie pliku tekstowego

1. W dowolnym oknie Edytora raportów kliknij tabelę, wykres lub inny obiekt, który będzie poprzedzać tekst.
2. Z menu wybierz:

   Wstaw > **Plik tekstowy...**
3. Wybierz plik tekstowy.

Aby edytować tekst, należy go dwukrotnie kliknąć.

Wklejanie obiektów do obszaru Okno raportu

Do obszaru Okno raportu można wklejać obiekty z innych aplikacji. Można użyć polecenia **Wklej poniżej** lub **Wklej specjalnie**. Oba rodzaje wklejania umożliwiają wstawienie nowego obiektu za aktualnie zaznaczonym obiektem w Oknie raportu. Polecenia **Wklej specjalnie** należy używać, gdy istnieje konieczność wybrania formatu wklejanego obiektu.

Znajdowanie i zastępowanie informacji w oknie raportu

1. Aby znaleźć lub zastąpić informacje w programie Viewer, z menu wybierz kolejno następujące pozycje:

   Edycja > **Znajdź**

   lub

   Edycja > **Zastąp**

Za pomocą funkcji znajdowania i zastępowania można wykonać następujące operacje:

- Przeszukiwanie całego dokumentu lub tylko wybranych elementów.
- Wyszukiwanie w górę lub w dół z wybranego miejsca.
- Wyszukiwanie w obu oknach lub ograniczenie wyszukiwania do okna raportu lub struktury.
- Wyszukiwanie elementów ukrytych. Należą do nich wszystkie elementy ukryte w oknie raportu (np. domyślnie ukryte tabele Uwagi) oraz ukryte wiersze i kolumny w tabelach przestawnych.
- Ograniczanie kryteriów wyszukiwania do dopasowań uwzględniających wielkość liter.
- Ograniczanie kryteriów wyszukiwania w tabelach przestawnych do dopasowań całej zawartości komórki.
- Ograniczanie kryteriów wyszukiwania w tabelach przestawnych tylko do znaczników przypisów. Ta opcja nie jest dostępna, jeśli zaznaczenie w Edytorze raportów obejmuje coś innego niż tabele przestawne.

Elementy ukryte i warstwy tabeli przestawnej

- Warstwy znajdujące się pod obecnie widoczną warstwą w wielowymiarowej tabeli przestawnej nie są uważane za ukryte i są przeszukiwane nawet wtedy, gdy wyszukiwanie nie obejmuje elementów ukrytych.
- Elementy ukryte to elementy ukryte w oknie raportu (elementy oznaczone ikoną zamkniętej książki w oknie struktury lub ukryte w zwiniętych blokach okna struktury) oraz wiersze i kolumny w tabelach przestawnych ukryte domyślnie (na przykład puste wiersze i kolumny są ukryte domyślnie) lub wybrane wiersze i kolumny ukryte ręcznie podczas edycji tabeli. Elementy ukryte są przeszukiwane tylko wtedy, gdy zostanie zaznaczona opcja **Dołącz elementy ukryte**.
- W obu przypadkach element ukryty lub niewidoczny zawierający szukany tekst lub wartość jest wyświetlany po znalezieniu, ale następnie jest przywracany jego pierwotny stan.

Znajdywanie zakresu wartości w tabelach przestawnych

Aby w tabelach przestawnych znaleźć wartości zawierające się w określonym zakresie wartości:

1. Aktywuj tabelę przestawną lub zaznacz w Edytorze raportów przynajmniej jedną tabelę przestawną. Upewnij się, że wybrane są tylko tabele przestawne. Jeśli wybrane zostały jakiekolwiek inne obiekty opcja Przedział nie będzie dostępna.
2. Z menu wybierz:

   Edycja > **Znajdź**
3. Kliknij kartę **Przedział**.
4. Wybierz typ przedziału: Między, Większy od lub równy albo Mniejszy od lub równy.
5. Wybierz wartość lub wartości definiujące przedział.

- Jeśli którakolwiek z wartości zawiera znaki nieliczbowe, obydwie wartości będą traktowane jako łańcuchy znaków.
- Jeśli obydwie wartości są liczbami, wyszukane zostaną tylko wartości liczbowe.
- Nie można użyć karty Przedziału do zastępowania wartości.

Ta opcja nie jest dostępna dla tabel starszej wersji. Więcej informacji można znaleźć w temacie “Starsze wersje tabel” na stronie 135.

Kopiowanie wyniku do innych aplikacji

Obiekty wynikowe mogą być kopiowane i wklejane do innych aplikacji, takich jak arkusze kalkulacyjne lub edytory tekstu. Wyniki można wklejać w różnych formatach. Zależnie od rodzaju aplikacji docelowe i wybranych obiektów wyniku, niektóre lub wszystkie formaty mogą nie być dostępne:

Plik meta. Format pliku meta WMF/EMF. Te formaty są dostępne tylko w systemach operacyjnych Windows.

RTF (rich text format). Wiele wybranych obiektów, wynik tekstowy oraz tabele przestawne można kopiować i wklejać w formacie RTF. W przypadku tabel przestawnych w większości przypadków powoduje to wklejenie tabeli jako tabeli, która może być edytowana w innej aplikacji. Tabele przestawne, które są za szerokie w stosunku do szerokości dokumentu zostaną zwinięte lub zmniejszone, aby pasowały do szerokości dokumentu, lub pozostaną bez zmian, zależnie od ustawień opcji tabeli przestawnej. Więcej informacji można znaleźć w temacie “Opcje tabel przestawnych” na stronie 137.

Uwaga: Program Microsoft Word może nie wyświetlić prawidłowo szczególnie szerokich tabel.

Obraz. Formaty obrazów JPG oraz PNG.

BIFF. Tabele przestawne oraz wynik tekstowy mogą być wklejane do arkusza w formacie BIFF. Liczby w tabelach przestawnych zachowują precyzję liczbową. Ten format jest dostępny tylko w systemach operacyjnych Windows.

Tekst. Tabele przestawne i wynik tekstowy można kopiować i wklejać jako tekst. Opcja ta może być użyteczna podczas korzystania z aplikacji, takich jak poczta elektroniczna, które mogą przyjmować i przekazywać jedynie dane tekstowe.

Obiekt graficzny pakietu Microsoft Office. Wykresy obsługujące ten format można kopiować do aplikacji Microsoft Office i w nich edytować jako rodzime wykresy Microsoft Office. Ze względu na różnice między wykresami programu SPSS Statistics/SPSS Modeler a wykresami Microsoft Office, niektóre opcje wykresów programu SPSS Statistics/SPSS Modeler nie są zachowywane w skopiowanej wersji. Nie jest możliwe kopiowanie więcej niż jednego zaznaczonego wykresu jako obiektów graficznych pakietu Microsoft Office.

Jeśli aplikacja docelowa obsługuje wiele dostępnych formatów, może zawierać pozycję menu Wklej specjalnie, która powoduje wyświetlenie listy dostępnych formatów lub pozwala na samodzielny wybór formatu najwłaściwszego w danej sytuacji.

Uwaga: Podczas kopiowania i wklejania wykresów skrzynkowych i histogramów wymagana jest wersja 16 (lub wyższa) pakietu Microsoft Office.

Kopiowanie i wklejanie wielu obiektów wynikowych

Podczas wklejania obiektów wielowynikowych do innych aplikacji występują poniższe ograniczenia:

- **Format RTF**. W większości aplikacji tabele przestawne są wklejane jako tabele, które można edytować w tej aplikacji. Wykresy, drzewa i widoki modeli są wklejane jako obrazy.
- **Format pliku meta oraz formaty obrazów**. Wszystkie wybrane obiekty wyników są wklejane do innych aplikacji jako pojedynczy obiekt.
- **Format BIFF**. Wykresy, drzewa i widoki modelu są wykluczone.

Kopiuj specjalne

W przypadku kopiowania i wklejania dużych ilości wyników, szczególnie bardzo dużych tabel przestawnych, można zwiększyć prędkość przetwarzania przy pomocy opcji **Edytuj** > **Kopiuj specjalne**, aby ograniczyć liczbę formatów kopiowanych do schowka.

Można także zapisać wybrane formaty jako domyślny zestaw formatów do kopiowania do schowka. Ustawienie to utrzymuje się przez wszystkie sesje.

Kopiuj jako

Można kliknąć wybrany obiekt prawym przyciskiem myszy w przeglądarce wyników i wybrać opcję**Edytuj** > **Kopiuj jako**, aby skopiować obiekt w jednym z popularnych formatów (na przykład **Wszystkie**, **Obraz** lub **Obiekt graficzny pakietu Microsoft Office**). Wybranie opcji **Edytuj** > **Kopiuj** jest równoznaczne z wybraniem opcji **Wszystkie**. Jeśli opcja **Kopiuj jako** jest wyszarzona lub niedostępna dla wybranego obiektu, to ten format kopiowania jest w jego przypadku niedostępny.

Wyniki interaktywne

Obiekty wyników interaktywnych zawierają wiele powiązanych obiektów wynikowych. Wybór jednego obiektu może zmienić zawartość wyświetloną lub wyróżnioną w innym obiekcie. Na przykład wybranie wiersza w tabeli może wyróżnić obszar na mapie lub wyświetlić wykres innej kategorii.

Obiekty wyników interaktywnych nie obsługują opcji edycji, takich jak zmiana tekstu, koloru, czcionek lub obramowania tabeli. Pojedyncze obiekty można kopiować z obiektu interaktywnego do okna raportów. Tabele skopiowane z wyników interaktywnych mogą być edytowane w edytorze tabel przestawnych.

Kopiowanie obiektów z wyników interaktywnych

Polecenie **Plik**> **Kopiuj do okna raportów** kopiuje poszczególne obiekty wynikowe do **okna raportów**.

- Dostępne opcje zależą od zawartości wyników interaktywnych.

- Polecenia **Wykres** i **Mapa** tworzą obiekty wykresu.
- Polecenie **Tabela** tworzy tabelę przestawną, którą można edytować w edytorze tabel przestawnych.
- Polecenie **Migawka** tworzy obraz bieżącego widoku.
- Polecenie **Model** tworzy kopię bieżącego obiektu wyników interaktywnych.

Polecenie **Edycja**>**Kopiuj obiekt** kopiuje poszczególne obiekty wyniku do schowka.

- Wklejenie skopiowanego obiektu do Edytora raportów jest równoważne z wybraniem polecenia **Plik**>**Kopiuj do okna raportów**.
- Wklejenie obiektu do innej aplikacji wkleja obiekt jako obraz.

Powiększenie i przesunięcie

W przypadku map można użyć polecenia **Widok**>**Powiększenie** w celu powiększenia widoku mapy. Przy powiększonej mapie można użyć polecenia **Widok**>**Przesuń** do przesunięcia widoku.

Ustawienia drukowania

Polecenie **Plik**>**Ustawienia drukowania** steruje sposobem drukowania obiektów.

- **Drukuj tylko widoczne widoki**. Drukuje tylko te widoki, które są aktualnie wyświetlone. Jest to opcja domyślna.
- **Drukuj wszystkie widoki**. Drukuje wszystkie widoki zawarte w wynikach interaktywnych.
- Wybrana opcja określa także domyślną czynność przy eksporcie obiektów wynikowych.

Eksportowanie raportu wynikowego

Okno dialogowe Eksportuj wynik umożliwia zapisanie wyników okna raportu w formatach HTML, tekstowym, Word/RTF, Excel, PowerPoint (wymaga programu PowerPoint 97 lub nowszej wersji) oraz w formacie PDF. Wykresy również można eksportować do różnych formatów graficznych.

Uwaga: Eksport do formatu PowerPoint jest dostępny tylko w systemach operacyjnych Windows.

Eksportowanie wyników

1. Aktywuj Okno raportu (kliknij w dowolnym miejscu okna).
2. Kliknij przycisk **Eksportuj** znajdujący się na pasku narzędzi lub kliknij prawym przyciskiem myszy okno wyników i wybierz opcję **Eksportuj**.
3. Wprowadź nazwę pliku (lub prefiks dla wykresów) i wybierz format eksportu.

Eksport obiektów. Można wyeksportować wszystkie obiekty występujące w Edytorze raportów, wszystkie obiekty widoczne lub tylko obiekty zaznaczone.

Typ dokumentu. Dostępne opcje:

- **Word/RTF (*.doc).** Tabele przestawne są eksportowane jako tabele programu Word, wraz z wszystkimi atrybutami formatowania, np. obramowaniami komórek, stylami czcionek i kolorami tła. Wyniki tekstowe są eksportowane w postaci tekstu sformatowanego RTF. Wykresy, diagramy drzewa oraz widoki modelu są zapisywane w formacie PNG. Uwaga: Program Microsoft Word może nie wyświetlić prawidłowo szczególnie szerokich tabel.
- **Excel 97-2004 (*.xls)/Excel 2007 i nowszy (*.xlsx).** Wiersze tabel przestawnych, kolumny i komórki tabel przestawnych są eksportowane jako wiersze, kolumny i komórki programu Excel, wraz z wszystkimi atrybutami formatowania, np. obramowaniami komórek, stylami czcionek, kolorami tła itd. Wyniki tekstowe są eksportowane wraz z wszystkimi atrybutami czcionek. Każdy wiersz wyniku tekstowego stanowi wiersz w pliku Excel, przy czym zawartość całego wiersza znajduje się w pojedynczej komórce. Wykresy, diagramy drzewa oraz widoki modelu są zapisywane w formacie PNG. Wyniki mogą zostać wyeksportowane jako pliki programu *Excel 97-2004* lub *Excel 2007 i nowszych*.

- **HTML (*.htm).** Tabele przestawne są eksportowane jako tabele HTML. Wyniki tekstowe są eksportowane w postaci wstępnie sformatowanego kodu HTML. Wykresy, diagramy drzew i widoki modelu są osadzone w dokumencie w wybranym formacie graficznym. Do wyświetlania wyników eksportowanych w formacie HTML wymagana jest przeglądarka kompatybilna ze standardem HTML 5.
- **Portable Document Format (*.pdf).** Wszystkie wyniki eksportowane są w takiej postaci, w jakiej widoczne są w funkcji Podgląd wydruku, wraz ze wszystkimi atrybutami czcionek.
- **Tekst - zwykły/UTF8/UTF16 (*.txt).** Do wynikowych formatów tekstowych należą: tekst zwykły, UTF-8 i UTF-16. Tabele przestawne mogą być eksportowane z separatorami w postaci znaków tabulacji lub spacji. Wszystkie wyniki tekstowe są eksportowane w formacie z separatorami w postaci spacji. W przypadku wykresów, diagramów drzewa oraz widoków modelu w pliku tekstowym wstawiana jest linia dla każdej grafiki, określająca nazwę pliku obrazu.
- **Żaden (tylko grafika).** Dostępne formaty eksportowanych plików: EPS, JPEG, TIFF, PNG i BMP. W systemach operacyjnych Windows dostępny jest także format EMF (enhanced metafile).

Otwórz folder zawierający. Otwiera folder zawierający pliki utworzone podczas eksportowania.

Opcje formatu HTML

Eksport w formacie HTML wymaga przeglądarki kompatybilnej ze standardem HTML 5.

W przypadku eksportowania wyników w formacie HTML dostępne są następujące opcje:

Warstwy tabel przestawnych. Domyślnie do dołączania i wykluczania warstw tabel przestawnych służą właściwości tabeli każdej tabeli przestawnej. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich warstw z wyjątkiem widocznej. Więcej informacji można znaleźć w temacie “Właściwości tabeli: drukowanie” na stronie 130.

Eksportuj tabele warstwowe jako interaktywne. Tabele warstwowe są wyświetlane w taki sam sposób, jak w Edytorze raportów i możliwa jest interaktywna zmiana wyświetlanej warstwy w przeglądarce. Jeśli ta opcja nie zostanie zaznaczona, każda warstwa tabeli jest wyświetlana w osobnej tabeli.

Tabele jako HTML. Steruje on informacją o stylu dołączoną dla eksportowanych tabel przestawnych.

- **Wyeksportuj ze stylami oraz stałymi szerokościami kolumn.** Wszystkie informacje o stylu tabeli przestawnej (style czcionek, kolory tła itp.). i szerokości kolumn zostają zachowane.
- **Eksportuj bez stylizacji.** Tabele przestawne zostają przekształcone na domyślne tabele w formacie HTML. Nie zostają zachowane żadne atrybuty stylu. Szerokość kolumny jest określana automatycznie.

Załącz przypisy i podpisy. Służy do dołączania lub wykluczania przypisów i nagłówków tabeli przestawnej.

Widoki modelu. Domyślnie włączenie lub wyłączenie widoków modelu jest kontrolowane przez właściwości każdego modelu. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich widoków z wyjątkiem widocznego. (Uwaga: wszystkie widoki modelu, w tym tabele, są eksportowane jako grafika).

Uwaga: W przypadku dokumentów HTML można również określać format pliku obrazu dla eksportowanych wykresów. Więcej informacji można znaleźć w temacie “Opcje formatu graficznego” na stronie 119.

Określanie opcji eksportu w formacie HTML

1. Jako format eksportu wybierz opcję **HTML**.
2. Kliknij przycisk **Zmień opcje**.

Opcje raportu WWW

Raport WWW jest interaktywnym dokumentem, który jest kompatybilny z większością przeglądarek. Wiele interaktywnych funkcji tabel przestawnych dostępnych w Edytorze raportów jest również dostępnych w raportach WWW.

Tytuł raportu. Tytuł, który jest wyświetlany w nagłówku raportu. Domyślnie używana jest nazwa pliku. Możliwe jest określenie tytułu użytkownika używanego zamiast nazwy pliku.

Format. Istnieją dwie opcje dla formatu raportu:

- **Raport WWW SPSS (HTML 5)**. Ten format wymaga przeglądarki kompatybilnej ze standardem HTML 5.
- **Raport aktywny Cognos (mht)**. Ten format wymaga przeglądarki kompatybilnej z plikami w formacie MHT lub aplikacji Cognos Active Report.

Wykluczanie obiektów. Możliwe jest wykluczenie wybranych typów obiektów z raportu:

- **Tekst**. Obiekty tekstowe, które nie są dziennikami. Ta opcja obejmuje obiekty tekstowe, które zawierają informacje o aktywnym zbiorze danych.
- **Dzienniki**. Obiekty tekstowe, które zawierają listing uruchomionych komend. Obiekty dzienników zawierają komunikaty o błędach i ostrzeżeniach napotkanych przy uruchamianiu komend, które nie wygenerowały żadnego wyniku w oknie raportów.
- **Tabele uwag**. Wynik procedur statystycznych i tworzenia wykresów z tabelą uwag. Ta tabela zawiera informacje o używanym zbiorze danych, brakujących wartościach i składni komend użytej do uruchomienia procedury.
- **Komunikaty ostrzegawcze i komunikaty o błędach**. Komunikaty ostrzegawcze i komunikaty o błędach z procedur statystycznych i tworzenia wykresów.

Zmień styl tabel i wykresów w celu dostosowania do raportów WWW. Ta opcja stosuje standardowy styl raportu WWW do wszystkich tabel i wykresów. To ustawienie zastępuje wszystkie czcionki, kolory i inne style w wynikach wyświetlanych w oknie raportów. Nie można zmodyfikować standardowego stylu raportu WWW.

Połączenie z serwerem WWW. Można podać adres URL serwerów aplikacji z uruchomionym produktem IBM SPSS Statistics Web Report Application Server. Serwer aplikacji WWW dysponuje funkcjami obsługi tabel przestawnych, edycji wykresów i zapisywania zmodyfikowanych raportów WWW.

- Zaznacz pole **Użyj** przy każdym serwerze aplikacji, który ma być wyszczególniony w raporcie WWW.
- Jeśli w raporcie WWW zostanie podany adres URL, raport połączy się z serwerem aplikacji, co pozwoli na zaoferowanie dodatkowych opcji edycji.
- Przy wielu adresach URL raport WWW będzie łączyć się z serwerami w kolejności podania adresów.

Produkt IBM SPSS Statistics Web Report Application Server można pobrać z serwisu http://www.ibm.com/developerworks/spssdevcentral.

Opcje formatu Word

W przypadku eksportowania wyników w formacie programu Word dostępne są następujące opcje:

Zabezpiecz punkty przerwania. Jeśli zdefiniowano punkty przerwania, odpowiednie ustawienia zostaną zachowane w tabelach programu Word.

Załącz przypisy i podpisy. Służy do dołączania lub wykluczania przypisów i nagłówków tabeli przestawnej.

Widoki modelu. Domyślnie włączenie lub wyłączenie widoków modelu jest kontrolowane przez właściwości każdego modelu. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich widoków z wyjątkiem widocznego. (Uwaga: wszystkie widoki modelu, w tym tabele, są eksportowane jako grafika).

Ustawienia strony dla eksportu. Umożliwia to otwarcie okna dialogowego, gdzie można zdefiniować wielkość strony i marginesy eksportowanego dokumentu. Szerokość dokumentu, używana do określania zawijania i pomniejszania, jest to szerokość strony minus lewy i prawy margines.

Określanie opcji eksportu w formacie Word

1. Wybierz opcję **Word/RTF** jako format eksportu.
2. Kliknij przycisk **Zmień opcje**.

Opcje formatu Excel

W przypadku eksportowania wyników w formacie programu Excel dostępne są następujące opcje:

Utwórz arkusz lub skoroszyt albo zmodyfikuj istniejący arkusz. Domyślnie jest tworzony nowy skoroszyt. Jeśli plik o podanej nazwie już istnieje, zostanie on zastąpiony. W przypadku wybrania opcji utworzenia arkusza i jeśli arkusz o określonej nazwie już istnieje w podanym pliku, zostanie on zastąpiony. W przypadku wybrania opcji modyfikacji istniejącego arkusza należy również podać nazwę arkusza. (Jest to opcjonalne w przypadku tworzenia arkusza). Nazwy arkuszy nie mogą przekraczać 31 znaków i nie mogą zawierać ukośników ani ukośników odwrotnych, nawiasów kwadratowych, znaków zapytania i gwiazdek.

Przy eksportowaniu do programu Excel 97-2004 w przypadku zmodyfikowania istniejącego arkusza wykresy widoki modelu i diagramy drzewa nie zostaną uwzględnione w eksportowanych wynikach.

Położenie w arkuszu. Określa położenie wyeksportowanych wyników na arkuszu. Domyślnie wyeksportowane wyniki zostaną dodane za ostatnią kolumną mającą jakąkolwiek zawartość, poczynając od pierwszego wiersza, bez modyfikowania istniejących danych. Jest to dobry sposób dodawania nowych kolumn do istniejącego arkusza. Dodanie wyeksportowanych wyników za ostatnim wierszem jest dobrym sposobem dodania nowych wierszy do istniejącego arkusza. Dodanie wyeksportowanych wyników, poczynając od konkretnego położenia, spowoduje zastąpienie istniejącej zawartości obszaru po dodaniu wyeksportowanych wyników.

Warstwy tabel przestawnych. Domyślnie do dołączania i wykluczania warstw tabel przestawnych służą właściwości tabeli każdej tabeli przestawnej. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich warstw z wyjątkiem widocznej. Więcej informacji można znaleźć w temacie “Właściwości tabeli: drukowanie” na stronie 130.

Załącz przypisy i podpisy. Służy do dołączania lub wykluczania przypisów i nagłówków tabeli przestawnej.

Widoki modelu. Domyślnie włączenie lub wyłączenie widoków modelu jest kontrolowane przez właściwości każdego modelu. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich widoków z wyjątkiem widocznego. (Uwaga: wszystkie widoki modelu, w tym tabele, są eksportowane jako grafika).

Określanie opcji eksportu w formacie programu Excel

1. Jako format eksportu wybierz opcję **Excel**.
2. Kliknij przycisk **Zmień opcje**.

Opcje formatu PowerPoint

Dla formatu PowerPoint dostępne są następujące opcje:

Warstwy tabel przestawnych. Domyślnie do dołączania i wykluczania warstw tabel przestawnych służą właściwości tabeli każdej tabeli przestawnej. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich warstw z wyjątkiem widocznej. Więcej informacji można znaleźć w temacie “Właściwości tabeli: drukowanie” na stronie 130.

Szerokie tabele przestawne. Steruje sposobem postępowania w przypadku tabel, które są zbyt szerokie dla zdefiniowanej szerokości dokumentu. Domyślnie tabela jest zawijana tak, aby się zmieściła. Tabela jest dzielona na części, a etykiety wierszy są powtarzane dla każdej części tabeli. Można też pomniejszyć szerokie tabele lub nie wprowadzać w nich zmian, umożliwiając ich wychodzenie poza zdefiniowaną szerokość dokumentu.

Załącz przypisy i podpisy. Służy do dołączania lub wykluczania przypisów i nagłówków tabeli przestawnej.

Użyj pozycji ze struktury raportu jako tytułów slajdów. Dołącza tytuł do każdego slajdu tworzonego podczas eksportu. Każdy slajd zawiera jeden element wyeksportowany z Edytora raportów. Tytuł jest tworzony z pozycji struktury odpowiedniego elementu w panelu struktury Edytora raportów.

Widoki modelu. Domyślnie włączenie lub wyłączenie widoków modelu jest kontrolowane przez właściwości każdego modelu. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich widoków z wyjątkiem widocznego. (Uwaga: wszystkie widoki modelu, w tym tabele, są eksportowane jako grafika).

Ustawienia strony dla eksportu. Umożliwia to otwarcie okna dialogowego, gdzie można zdefiniować wielkość strony i marginesy eksportowanego dokumentu. Szerokość dokumentu, używana do określania zawijania i pomniejszania, jest to szerokość strony minus lewy i prawy margines.

Określanie opcji eksportu w formacie PowerPoint

1. Jako format eksportu wybierz opcję **PowerPoint**.
2. Kliknij przycisk **Zmień opcje**.

Uwaga: Eksport do formatu PowerPoint jest dostępny tylko w systemach operacyjnych Windows.

Opcje PDF

Dla formatu PDF dostępne są następujące opcje:

Osadź zakładki. Opcja pozwalająca na dołączanie zakładek do dokumentu PDF, które odpowiadają pozycjom w Edytorze raportów. Podobnie jak okno Edytora raportów, zakładki pozwalają na łatwiejszą nawigację po dokumentach z dużą liczbą obiektów wynikowych.

Osadź czcionki. Osadzenie czcionek zapewnia identyczny wygląd dokumentu PDF na wszystkich komputerach. Jeśli ta funkcja nie zostanie użyta, a niektóre użyte czcionki nie będą dostępne na komputerze wyświetlającym (lub drukującym) dokument PDF, funkcja zastępowania czcionek może spowodować obniżenie jakości końcowej.

Warstwy tabel przestawnych. Domyślnie do dołączania i wykluczania warstw tabel przestawnych służą właściwości tabeli każdej tabeli przestawnej. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich warstw z wyjątkiem widocznej. Więcej informacji można znaleźć w temacie “Właściwości tabeli: drukowanie” na stronie 130.

Widoki modelu. Domyślnie włączenie lub wyłączenie widoków modelu jest kontrolowane przez właściwości każdego modelu. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich widoków z wyjątkiem widocznego. (Uwaga: wszystkie widoki modelu, w tym tabele, są eksportowane jako grafika).

Określanie opcji eksportu formatu PDF

1. Jako format eksportu wybierz opcję **Portable Document Format**.
2. Kliknij przycisk **Zmień opcje**.

Inne ustawienia wpływające na wynikowy dokument PDF

Ustawienia strony/Atrybuty strony. Rozmiar strony, jej orientacja, marginesy, zawartość i wyświetlanie nagłówków i stopek oraz rozmiar wykresów drukowanych w dokumentach PDF są uzależnione od opcji ustawień strony i atrybutów strony.

Właściwości tabeli/Szablony TableLooks. Skalowaniem szerokich i/lub długich tabel oraz drukowaniem warstw tabel można sterować za pomocą właściwości każdej z tabel. Właściwości te można także zapisywać w szablonach TableLooks.

Domyślna/bieżąca drukarka. Rozdzielczość (DPI) dokumentu PDF to bieżące ustawienie rozdzielczości dla domyślnej lub bieżącej drukarki (co można zmienić za pomocą ustawień strony). Maksymalna rozdzielczość wynosi 1200 DPI. Jeśli ustawienie drukarki jest wyższe, rozdzielczość dokumentu PDF będzie równa 1200 DPI.

Uwaga: Jakość dokumentów w wysokiej rozdzielczości drukowanych na drukarkach o niższej rozdzielczości może być niska.

Opcje formatu tekstowego

Ustawianie opcji zapisywania wyników w formacie tekstowym

Dla eksportu tekstu dostępne są następujące opcje:

Format tabeli przestawnej. Tabele przestawne mogą być eksportowane z separatorami w postaci znaków tabulacji lub spacji. W przypadku plików z separatorami w postaci spacji można także określić:

- **Szerokość kolumny.** Użycie funkcji **Autodopasowanie** nie powoduje zawijania zawartości kolumn. Szerokość każdej kolumny będzie równa szerokości najszerszej etykiety lub wartości w danej kolumnie. Opcja **Użytkownika** pozwala na wprowadzenie maksymalnej szerokości kolumny, która następnie zostanie zastosowana w całej tabeli. Wartości szersze zostaną zawinięte do drugiego wiersza.
- **Znak obramowania wierszy/kolumn.** Zmiana znaku służącego do tworzenia obramowań wierszy i kolumn. Aby wyłączyć wyświetlanie obramowań wierszy i kolumn, wprowadź dla wszystkich wartości spację.

Warstwy tabel przestawnych. Domyślnie do dołączania i wykluczania warstw tabel przestawnych służą właściwości tabeli każdej tabeli przestawnej. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich warstw z wyjątkiem widocznej. Więcej informacji można znaleźć w temacie “Właściwości tabeli: drukowanie” na stronie 130.

Załącz przypisy i podpisy. Służy do dołączania lub wykluczania przypisów i nagłówków tabeli przestawnej.

Widoki modelu. Domyślnie włączenie lub wyłączenie widoków modelu jest kontrolowane przez właściwości każdego modelu. To ustawienie można zastąpić i spowodować dołączenie lub wykluczenie wszystkich widoków z wyjątkiem widocznego. (Uwaga: wszystkie widoki modelu, w tym tabele, są eksportowane jako grafika).

Określanie opcji eksportowania tekstu

1. Jako format eksportu wybierz opcję **Tekst**.
2. Kliknij przycisk **Zmień opcje**.

Opcja eksportu samych obrazów

Do eksportu samych obrazów przeznaczone są następujące opcje:

1. Uaktywnij kartę wyników (kliknij w dowolnym miejscu karty).
2. Wybierz elementy wykresu, które mają zostać wyeksportowane, kliknij przycisk pionowego wielokropka i wybierz opcję **Eksportuj** z menu. Można również kliknąć przycisk **Eksportuj wynik** znajdujący się na pasku narzędzi. Wyświetlane jest okno dialogowe Eksportuj raport wynikowy.

   Ustawienia dokumentu

   Typ
   Wybierz **Tylko obrazy** jako **Typ** dokumentu.

   Format obrazów
   Wybierz format eksportowanych obrazów.

   PNG
   Gdy ta opcja jest wybrana, obrazy są eksportowane w formacie Portable Network Graphics (*.png). Ustawienie to udostępnia opcje **Skala (%)** i **Głębia kolorów**.

   JPG
   Gdy ta opcja jest wybrana, obrazy są eksportowane w formacie Joint Photographic Experts Group (*.jpg). Ustawienie to udostępnia opcje **Skala (%)** i **Konwertuj na skalę szarości**.

   BMP
   Gdy ta opcja jest wybrana, obrazy są eksportowane w formacie mapy bitowej (*.bmp). Ustawienie to udostępnia opcje **Skala (%)** i **Kompresuj obraz, aby ograniczyć rozmiar pliku**.

Otwórz folder zawierający
Opcjonalnie wybierz opcję otwierania folderu docelowego po zakończeniu eksportu.

Eksport obiektów

Obiekty do uwzględnienia
Wybierz, które obiekty wynikowe mają być eksportowane: **Wszystkie**, **Widoczne** lub **Wybrane** (ustawienie domyślne).

3. Po określeniu ustawień dokumentu kliknij przycisk **Eksportuj**.

Opcje formatu graficznego

Dla dokumentów HTML i tekstowych oraz tylko dla eksportu wykresów można wybrać format graficzny, a dla każdego formatu graficznego można konfigurować różne ustawienia opcjonalne.

Aby wybrać format graficzny i opcje eksportowanych wykresów:

1. Jako typ dokumentu wybierz **HTML**, **Tekst** lub **Żaden (tylko grafika)**.
2. Wybierz format pliku graficznego z listy rozwijanej.
3. Kliknij przycisk **Zmień opcje**, aby zmienić opcje wybranego formatu pliku graficznego.

Opcje eksportu wykresów w formacie JPEG

Wielkość obrazu (%)
Wartość procentowa oryginalnego rozmiaru wykresu, maks. 200%.

Konwertuj na skalę szarości
Przekształca kolory na odcienie szarości.

Opcje eksportu wykresów w formacie BMP

Wielkość obrazu (%)
Wartość procentowa oryginalnego rozmiaru wykresu, maks. 200%.

Kompresuj obraz, aby ograniczyć rozmiar pliku
Metoda kompresji bezstratnej służąca do tworzenia mniejszych plików bez zmiany jakości obrazu.

Opcje eksportu wykresów w formacie PNG

Wielkość obrazu (%)
Wartość procentowa oryginalnego rozmiaru wykresu, maks. 200%.

Głębia kolorów
Określa liczbę kolorów w eksportowanym wykresie. Minimalna liczba kolorów w wykresie zapisanym z dowolną głębią kolorów jest równa liczbie kolorów faktycznie wykorzystanych, a liczba maksymalna jest określona głębią kolorów. Jeśli na przykład wykres zawiera trzy kolory — czerwony, biały i czarny i zostanie zapisany z 16 kolorami, po zapisaniu nadal będzie zawierał trzy kolory.

Uwaga: Jeśli liczba kolorów wykresu przewyższa liczbę kolorów określoną daną głębokością, do przedstawienia kolorów brakujących zostanie zastosowana symulacja kolorów.

Opcje eksportu wykresów w formacie EMF i TIFF

Rozmiar obrazu. Wartość procentowa oryginalnego rozmiaru wykresu, maks. 200%.

Uwaga: Format EMF (enhanced metafile) jest dostępny tylko w systemach operacyjnych Windows.

Opcje eksportu wykresów w formacie EPS

Rozmiar obrazu. Rozmiar można określić jako wartość procentową oryginalnego rozmiaru obrazu (maks. 200%) lub podać jego szerokość w pikselach (wysokość zostanie ustalona na podstawie wartości szerokości i współczynnika proporcji). Wyeksportowany obraz ma zawsze identyczne proporcje jak oryginał.

Załącz podgląd obrazu w formacie TIFF. Zapisanie z obrazem EPS podglądu w formacie TIFF, który można wyświetlić w aplikacjach, które nie mogą wyświetlać obrazów EPS na ekranie.

Czcionki. Kontrola sposobu postępowania z czcionkami w obrazach EPS.

- **Użyj referencji czcionek.** Jeśli czcionki wykorzystane w wykresie są dostępne na urządzeniu wynikowym, zostają wykorzystane właśnie te czcionki. W przeciwnym wypadku urządzenie wynikowe wykorzystuje czcionki alternatywne.
- **Zamień czcionki na krzywe.** Powoduje przekształcenie czcionek na dane krzywych PostScript. W aplikacjach umożliwiających edycję grafiki EPS sam tekst nie może być edytowany wówczas jako tekst. Ta opcja jest użyteczna, jeżeli czcionki wykorzystane w wykresie nie są dostępne na urządzeniu wynikowym.

Drukowanie z Edytora raportów

Zawartość okna Edytor raportów można drukować na dwa sposoby:

Cały widoczny raport. Wybranie tej opcji powoduje wydrukowanie tylko elementów aktualnie widocznych w oknie raportu. Elementy ukryte (elementy oznaczone ikoną zamkniętej książki w oknie struktury lub ukryte w zwiniętych warstwach struktury) nie są drukowane.

Zaznaczone. Powoduje wydrukowanie wyłącznie elementów aktualnie zaznaczonych w oknie struktury i/lub w oknie raportu.

Drukowanie raportu i wykresów

1. Uaktywnij okno Edytora raportów (kliknij w dowolnym miejscu okna).
2. Z menu wybierz:

   Plik > **Drukuj...**
3. Wybierz odpowiednie ustawienia drukowania.
4. Kliknij przycisk **OK**, aby rozpocząć drukowanie.

Podgląd wydruku

Za pomocą podglądu wydruku można zobaczyć, jak będzie wyglądać każda wydrukowana strona dokumentów Edytora raportów. Przed rozpoczęciem drukowania dokumentu z Edytora raportów warto zazwyczaj obejrzeć go za pomocą podglądu wydruku, gdyż umożliwia to wykrycie elementów, które mogą nie być widoczne w oknie Edytora raportu, włączając w to:

- Podziały stron
- Ukryte warstwy tabel przestawnych
- Podziały w szerokich tabelach
- Nagłówki i stopki umieszczane na każdej stronie

Jeśli w Edytorze raportów zaznaczone są pewne wyniki, to na podglądzie wydruku widoczne będą tylko one. Aby wyświetlić podgląd całości raportu, należy zadbać o to, aby żadne elementy nie były zaznaczone.

Atrybuty strony: nagłówki i stopki

Nagłówki i stopki są to informacje drukowane w górnej i dolnej części każdej strony. Nagłówki i stopki mogą zawierać dowolny tekst. Ponadto, za pomocą paska narzędzi znajdującego się w środkowej części okna można wstawić elementy, takie jak:

- Data i czas
- Numery stron
- Nazwa pliku Edytora raportów
- Etykiety nagłówków struktury

- Tytuły i podtytuły stron
- Opcja **Ustaw jako domyślne** powoduje użycie określonych tutaj ustawień jako ustawień domyślnych dla nowych dokumentów programu Viewer. (Uwaga: spowoduje to, że bieżące ustawienia na zakładkach Nagłówek/Stopka i Opcje zostaną ustawieniami domyślnymi).
- Etykiety nagłówków struktury zawierają treść nagłówka pierwszego, drugiego, trzeciego i/lub czwartego poziomu struktury dla pierwszego elementu na każdej stronie.
- Dodatkowe opcje pozwalają drukować tytuły i podtytuły stron. Są one tworzone za pomocą komendy Nowy tytuł strony w menu Wstaw programu Viewer lub za pomocą komend `TITLE` i `SUBTITLE`. W przypadku, gdy żadne tytuły ani podtytuły nie zostały określone, ustawienie to jest ignorowane.

  Uwaga: Parametry czcionek dla nowych tytułów i podtytułów stron są kontrolowane z poziomu karty Raporty w oknie dialogowym Opcje (dostępne przez wybranie elementu Opcje w menu Edycja). Parametry czcionek dla istniejących tytułów i podtytułów stron można zmienić, edytując tytuły w Edytorze raportów.

Pozycja Podgląd wydruku w menu Plik pozwala zobaczyć, jak nagłówki i stopki będą wyglądać na wydrukowanej stronie.

Sposób wstawiania nagłówków i stopek strony

1. Uaktywnij okno Edytora raportów (kliknij w dowolnym miejscu okna).
2. Z menu wybierz kolejno następujące pozycje:

   Plik > **Nagłówek i stopka...**
3. Wprowadź nagłówek i/lub stopkę, które będą znajdować się na każdej stronie.

Atrybuty strony w opcjach wydruku

To okno dialogowe umożliwia sterowanie rozmiarem wykresów na wydruku, odstępami między drukowanymi elementami oraz numerowaniem stron.

- **Rozmiar drukowanych wykresów.** Umożliwia określenie rozmiaru drukowanych wykresów względem zdefiniowanego rozmiaru strony. Ustawienie to nie ma wpływu na współczynnik proporcji wykresu (stosunek szerokości do wysokości). Rozmiar wykresu na wydruku jest ograniczany przez jego szerokość i wysokość. Kiedy wykres osiągnie szerokość na tyle dużą, że jego boczne krawędzie zetkną się z bocznymi krawędziami strony, nie będzie już możliwe dalsze powiększanie wykresu w celu wypełnienia pozostającej wysokości strony.
- **Odstęp między pozycjami.** Umożliwia określenie odstępów między elementami na wydruku. Każda tabela przestawna, wykres lub obiekt tekstowy są oddzielnymi elementami. Ustawienie to nie wpływa na sposób wyświetlania elementów w Edytorze raportów.
- **Numeracja stron rozpoczyna się od.** Powoduje nadawanie stronom kolejnych numerów, rozpoczynając od określonego numeru.
- **Ustaw jako domyślne.** Ta opcja powoduje użycie określonych tutaj ustawień jako ustawień domyślnych dla nowych dokumentów okna raportów. (Wybranie tej opcji spowoduje wskazanie bieżących ustawień nagłówka i stopki oraz opcji jako domyślnych).

Zmiana rozmiaru wykresu na wydruku, numerowania stron i odstępu między drukowanymi elementami

1. Uaktywnij okno Edytora raportów (kliknij w dowolnym miejscu okna).
2. Z menu wybierz:

   Plik > **Atrybuty strony...**
3. Kliknij kartę **Opcje**.
4. Zmień ustawienia i kliknij przycisk **OK**.

Zapisywanie wyników

Zawartość okna raportów można zapisać.

- **Obiekt wynikowy (*.cou)**. Ten format pozwala na zapisanie całego kontenera wynikowego, włącznie z wykresami, kartami, adnotacjami itd. Pliki w tym formacie można otwierać i przeglądać w programie IBM SPSS Modeler, dodawać do projektów oraz publikować i śledzić za pomocą repozytorium IBM SPSS Collaboration and Deployment Services Repository. Ten format nie jest zgodny z produktem IBM SPSS Statistics.
- **Pliki okna raportu (*.spv)**. Format używany do wyświetlania plików obszarze **Okno raportu** . Przy zapisywaniu danych do tego formatu z modelu użytkowego w produkcie IBM SPSS Modeler, zapisywana jest wyłącznie zawartość okna raportów z karty **Model**.

Aby kontrolować opcje zapisywania raportów WWW lub zapisać wyniki w innych formatach (na przykład: tekst, Word, Excel), użyj opcji **Eksportuj** w menu **Plik**.

Zapisywanie dokumentu okna raportu

1. Z menu Edytora raportów wybierz kolejno następujące pozycje:

   File > **Save**

2. Wpisz nazwę dokumentu i kliknij przycisk **Zapisz**.

   Opcjonalnie można wykonać następujące czynności:

   Blokowanie plików w celu uniemożliwienia edycji w produkcie IBM SPSS Smartreader
   Jeśli dokument Edytora raportów zostanie zablokowany, możesz manipulować tabelami obrotowymi (zamieniać wiersze i kolumny, wyświetlane warstwy itp.), ale nie możesz edytować wyników ani zapisywać żadnych zmian w dokumencie w programie IBM SPSS Smartreader (osobnym produkcie służącym do pracy z dokumentami Edytora raportów). To ustawienie nie ma wpływu na dokumenty okna raportów otwierane w produkcie IBM SPSS Statistics lub IBM SPSS Modeler.

   Szyfrowanie plików za pomocą hasła
   Możesz ochronić tajne informacje przechowywane w dokumencie Edytora raportów szyfrując dokument używając hasła. Po zaszyfrowaniu dokument będzie mógł zostać otwarty tylko po podaniu hasła. Użytkownicy IBM SPSS Smartreader będą także musieli podać hasło, aby otworzyć plik.

   Szyfrowanie dokumentu Edytora raportów:

   a. W oknie dialogowym Zapisz wynik jako zaznacz **Zaszyfruj plik używając hasła**.
   b. Kliknij przycisk **Zapisz**.
   c. W oknie dialogowym Zaszyfruj plik podaj hasło i powtórz je w polu Potwierdź hasło. Długość haseł jest ograniczona do 10 znaków, w których ważna jest wielkość liter.

   Ostrzeżenie: Jeśli hasła zaginą, nie będą one mogły zostać odzyskane. Jeśli hasło zaginie, nie można otworzyć pliku.

   Tworzenie mocnych haseł

   - Użyj przynajmniej ośmiu znaków.
   - W swoim haśle użyj liczb, symboli, a nawet znaków interpunkcyjnych.
   - Unikaj sekwencji liczb lub znaków takich, jak `"123"` i `"abc"`, unikaj też powtórzeń, jak na przykład: `"111aaa"`.
   - Nie twórz haseł używających prywatnych informacji takich, jak data urodzin czy przezwiska.
   - Co jakiś czas zmieniaj hasło.

   Uwaga: Zapisywanie zaszyfrowanych plików w IBM SPSS Collaboration and Deployment Services Repository nie jest obsługiwane.

Modyfikowanie zaszyfrowanych plików

- Jeśli otworzysz zaszyfrowany plik, dokonasz w nim modyfikacji i wybierzesz Plik > Zapisz, zmodyfikowany plik zostanie zapisany z użyciem tego samego hasła.
- Możesz zmienić hasło zaszyfrowanego pliku otwierając go, powtarzając kroki dla szyfrowania i określając inne hasło w oknie dialogowym Zaszyfruj plik.
- Możesz zapisać niezaszyfrowaną wersję zaszyfrowanego pliku otwierając go, wybierając Plik > Zapisz jako i usuwając zaznaczenie dla **Zaszyfruj plik używając hasła** w oknie dialogowym Zapisz wynik jako.

Uwaga: Zaszyfrowanych plików danych i dokumentów wynikowych nie można otworzyć za pomocą produktu IBM SPSS Statistics w wersji starszej niż 21. Zaszyfrowanych plików komend nie można otworzyć za pomocą produktu w wersji starszej niż 22.

Składowanie wymaganych informacji o modelu z dokumentem wynikowym
Ta opcja ma zastosowanie wyłącznie przy pozycjach okna modelu w dokumencie wynikowym wymagających dodatkowych informacji do włączenia niektórych funkcji interaktywnych. Kliknij polecenie **Więcej informacji**, aby wyświetlić listę tych pozycji okna modelu i funkcji interaktywnych, które wymagają dodatkowych informacji. Składowanie tych informacji z dokumentem wynikowym może znacząco zwiększyć rozmiar dokumentu. Jeśli informacje te nie mają być zapisywane, pozycje wynikowe można nadal otworzyć, lecz bez określonych właściwości interaktywnych.

Tabele przestawne

Tabele przestawne

Wiele wyników przedstawia się w tabelach, które można interaktywnie przestawiać. Oznacza to, że można zmieniać układ wierszy, kolumn i warstw.

Manipulowanie tabelą przestawną

Możliwości manipulacji tabelami przestawnymi obejmują:

- Transponowanie wierszy i kolumn
- Przenoszenie wierszy i kolumn
- Tworzenie wielowymiarowych warstw
- Grupowanie i rozdzielanie grup wierszy i kolumn
- Wyświetlanie i ukrywanie wierszy, kolumn i innych informacji
- Obracanie etykiet wierszy i kolumn
- Wyszukiwanie definicji terminów

Aktywowanie tabeli przestawnej

Aby można było manipulować tabelą przestawną lub modyfikować ją, należy ją najpierw **aktywować**. Aby aktywować tabelę:

1. Dwukrotnie kliknij tabelę.

   lub

2. Kliknij tabelę prawym przyciskiem myszy i z menu kontekstowego wybierz pozycję **Edytuj zawartość**.
3. Z podmenu wybierz opcję **W Edytorze raportów** lub **W osobnym oknie**.

Przestawianie tabeli

Tabela ma trzy wymiary: wiersze, kolumny i warstwy. Wymiar może zawierać wiele (lub nie zawierać żadnych) elementów. Sposób organizacji tabeli można zmieniać, przenosząc elementy między wymiarami

lub w ramach wymiarów. Aby przenieść element, wystarczy go przeciągnąć i upuścić w wybranym miejscu.

Zmiana porządku wyświetlania elementów w wymiarze

Aby zmienić porządek wyświetlania elementów w wymiarze tabeli (wierszu, kolumnie lub warstwie):

1. Jeśli panel przestawiania nie jest jeszcze aktywny, z menu tabeli przestawnej wybierz kolejno następujące pozycje:

   Przestaw > **Panel przestawiania**
2. Przeciągnij i upuść elementy w wymiarze na panelu przestawiania.

Przenoszenie wierszy i kolumn w elemencie wymiaru

1. W samej tabeli (nie na panelach przestawiania) kliknij etykietę wiersza lub kolumny, które chcesz przenieść.
2. Kliknij i przeciągnij etykietę na nowe miejsce.

Transponowanie wierszy i kolumn

Jeśli chcesz tylko zamienić wiersze i kolumny miejscami, zamiast używać paneli przestawiania, można wykorzystać prostszą opcję:

1. Z menu wybierz:

   Przestaw > **Transponuj wiersze i kolumny**

Efekt będzie taki sam jak po przeciągnięciu wszystkich elementów wierszy do wymiaru kolumn i wszystkich elementów kolumn do wymiaru wierszowego.

Grupowanie wierszy i kolumn

1. Zaznacz etykiety wierszy lub kolumn, które zamierzasz zgrupować (Aby zaznaczyć kilka etykiet naraz, kliknij i przeciągnij lub kliknij, przytrzymując klawisz Shift).
2. Z menu wybierz:

   Edytuj > **Grupa**

Automatycznie zostanie wstawiona etykieta grupy. Aby zmienić tekst etykiety, kliknij go dwukrotnie.

Uwaga: Aby dodać wiersze lub kolumny do istniejącej grupy, należy najpierw rozdzielić znajdujące się w niej obecnie pozycje. Następnie można utworzyć nową grupę, która zawiera dodatkowe pozycje.

Rozdzielanie grup wierszy lub kolumn

W wyniku rozdzielenia grupy automatycznie zostanie usunięta jej etykieta.

Obracanie nagłówków wierszy lub kolumn

Etykiety można obracać tak, aby etykiety kolumn umieszczonych najbardziej na wewnątrz i etykiety wierszy umieszczonych najbardziej na zewnątrz były wyświetlane w poziomie lub w pionie.

1. Z menu wybierz:

   Format > **Obróć nagłówki kolumn**

   lub

   Format > **Obróć nagłówki wierszy**

Można obrócić tylko wewnętrzne etykiety (nagłówki) kolumn i tylko zewnętrzne etykiety wierszy.

Sortowanie wierszy

Aby posortować wiersze tabeli przestawnej:

1. Uaktywnij tabelę.
2. Zaznacz dowolną komórkę w kolumnie, według której chcesz sortować. Aby posortować tylko zaznaczoną grupę wierszy, zaznacz dwie lub więcej sąsiadujących ze sobą komórek w kolumnie, według której chcesz sortować.
3. Z menu wybierz kolejno następujące pozycje:

   Edycja > **Sortuj wiersze**
4. Z menu podrzędnego wybierz **Ascending** lub **Descending**.

- Jeśli wymiar wierszowy zawiera grupy, sortowanie obejmie tylko grupę, która zawiera zaznaczenie.
- Nie możesz sortować przekraczając granice grup.
- Nie możesz sortować tabel z więcej niż jednym elementem w wymiarze wierszowym.

Wstawianie wierszy i kolumn

Aby wstawić wiersz lub kolumnę w tabeli przestawnej:

1. Uaktywnij tabelę.
2. Zaznacz dowolną komórkę w tabeli.
3. Z menu wybierz kolejno następujące pozycje:

   Wstaw przed

   lub

   Wstaw po

   W podmenu wybierz opcję:

   Wiersz

   lub

   Kolumna

- W każdej komórce nowego wiersza lub kolumny wstawiony zostaje znak plus (+), aby zapobiec temu, by nowy wiersz lub kolumna zostały od razu ukryte z powodu braku danych.
- W tabeli posiadającej zagnieżdżone lub warstwowe wymiary, kolumna lub wiersz są wstawiane w każdym odpowiadającym poziomie wymiaru.

Kontrola wyświetlania etykiet zmiennej i wartości

Jeśli zmienne zawierają opisowe etykiety zmiennej lub wartości, możesz sterować wyświetlaniem nazw i etykiet zmiennych oraz wartościami danych i etykietami wartości w tabelach przestawnych.

1. Uaktywnij tabelę przestawną.
2. Z menu wybierz kolejno następujące pozycje:

   Widok > **Etykiety zmiennych**

   lub

   Widok > **Etykiety wartości**
3. Zaznacz jedną z następujących opcji menu podrzędnego:

- **Nazwa** lub **Wartość**. Wyświetlane są tylko nazwy (lub wartości) zmiennych. Etykiety opisowe nie są wyświetlane.
- **Etykieta**. Wyświetlane są tylko etykiety opisowe. Nazwy (lub wartości) zmiennych nie są wyświetlane.

- **Oba**. Wyświetlane są zarówno nazwy (lub wartości), jak i etykiety opisowe.

Zmiana języka wyjściowego

Aby zmienić język w tabeli przestawnej:

1. Uaktywnij tabelę
2. Z menu wybierz kolejno następujące pozycje:

   Widok > **Język**
3. Wybierz jeden z dostępnych języków.

Zmiana języka ma wpływ tylko na tekst generowany przez aplikację: nazwy tabel, etykiety wierszy i kolumn, tekst przypisów itp. Nie ma wpływu na nazwę zmiennych, zmienne opisowe ani etykiety wartości.

Przemieszczanie się po dużych tabelach

Aby użyć okna nawigacji do przemieszczania się po dużych tabelach:

1. Uaktywnij tabelę.
2. Z menu wybierz:

   Widok > **Nawigacja**

Cofanie zmian

Możesz cofnąć większość najnowszych zmian lub wszystkie zmiany dla aktywnej tabeli przestawnej. Obie czynności mają zastosowanie wyłącznie do zmian wprowadzonych od ostatniej aktywacji tabeli.

Aby cofnąć najnowszą zmianę:

1. Z menu wybierz kolejno następujące pozycje:

   Edycja > **Cofnij**

   Aby cofnąć wszystkie zmiany:
2. Z menu wybierz kolejno następujące pozycje:

   Edycja > **Przywróć**

Praca z warstwami

Dla każdej kategorii lub kombinacji kategorii można wyświetlić osobną dwuwymiarową tabelę. Tabelę można traktować tak, jakby była złożona z leżących na sobie warstw, z widoczną tylko warstwą wierzchnią.

Tworzenie i wyświetlanie warstw

Tworzenie warstw:

1. Jeśli panele przestawiania nie są jeszcze aktywne, z menu **Tabela przestawna** wybierz kolejno następujące pozycje:

   Przestaw > **Panel przestawiania**
2. Przeciągnij element z wymiaru wierszy lub kolumn do wymiary warstw.

Przenoszenie elementów do wymiaru warstw powoduje utworzenie wielowymiarowej tabeli; wyświetlany jest jednak tylko jej pojedynczy, dwuwymiarowy „wycinek”. Tabelą widoczną jest tabela warstwy wierzchniej. Na przykład, jeśli w wymiarze warstwy znajduje się zmienna kategorialna tak/nie, to tabela wielowymiarowa ma dwie warstwy: jedną dla kategorii „tak”, a drugą dla kategorii „nie”.

Zmiana wyświetlanej warstwy

1. Z listy rozwijanej warstw (w samej tabeli przestawnej, nie na panelu przestawiania) wybierz kategorię.

Przechodzenie do kategorii w warstwie

Okno dialogowe Przejdź do kategorii w warstwie umożliwia zmianę układu warstw w tabeli przestawnej. To okno dialogowe jest szczególnie przydatne wtedy, gdy istnieje wiele warstw lub gdy wybrana warstwa ma wiele kategorii.

Wyświetlanie i ukrywanie pozycji

Można ukrywać wiele typów komórek:

- Etykiety wymiarów
- Kategorie, łącznie z komórką etykiety i komórkami danych w wierszu lub kolumnie
- Etykiety kategorii (bez ukrywania komórek danych)
- Przypisy, tytuły i nagłówki

Ukrywanie wierszy i kolumn w tabeli

Wyświetlanie ukrytych wierszy lub kolumn w tabeli

1. Z menu wybierz:

   Widok > **Pokaż kategorie**

Spowoduje to wyświetlenie wszystkich ukrytych wierszy i kolumn tabeli. (Jeżeli w wypadku danej tabeli w oknie Właściwości tabeli zaznaczona jest opcja **Ukryj puste wiersze i kolumny**, całkowicie pusty wiersz lub kolumna pozostaną ukryte).

Ukrywanie i wyświetlanie etykiet wymiaru

1. Zaznacz nagłówek (etykietę) wymiaru lub dowolną etykietę kategorii w obrębie danego wymiaru.
2. Z menu Widok lub menu kontekstowego wybierz pozycję **Ukryj opis wymiaru** lub **Pokaż opis wymiaru**.

Ukrywanie i wyświetlanie tytułów tabel

Ukrywanie tytułu:

1. Uaktywnij tabelę przestawną.
2. Wybierz tytuł.
3. Z menu Widok wybierz pozycję **Ukryj**.

   Wyświetlanie ukrytych tytułów:
4. Z menu Widok wybierz pozycję **Pokaż wszystko**.

Szablony TableLook

Szablon TableLook to zestaw właściwości definiujących wygląd tabeli. Można wybrać wcześniej zdefiniowany szablon TableLook lub utworzyć własny.

- Przed zastosowaniem szablonu TableLook lub po jego zastosowaniu można zmienić formatowanie poszczególnych komórek lub ich grup, posługując się oknem dialogowym Właściwości komórki. Zmodyfikowane formaty komórek nie zostaną zmienione nawet po zastosowaniu nowego szablonu TableLook.
- Ewentualnie można nadać wszystkim komórkom formaty zdefiniowane w bieżącym szablonie TableLook. Ta opcja resetuje wszystkie zmodyfikowane komórki. Jeśli na liście plików TableLook

zaznaczona jest opcja **Tak, jak wyświetlane**, do wszystkich zmodyfikowanych komórek zostaną zastosowane bieżące właściwości tabeli.

- W szablonach TableLook zapisane zostają tylko te właściwości tabel, które zostały zdefiniowane w oknie dialogowym Właściwości tabel. Szablony TableLook nie zawierają indywidualnych modyfikacji komórek.

Aby zastosować szablon TableLook

1. Uaktywnij tabelę przestawną.
2. Z menu wybierz kolejno następujące pozycje:

   Format > **Szablony TableLook...**
3. Z listy plików wybierz szablon TableLook. Aby wybrać plik znajdujący się w innym katalogu, kliknij przycisk **Przeglądaj**.
4. Kliknij przycisk **OK**, aby zastosować szablon TableLook do wybranej tabeli przestawnej.

Edytowanie lub tworzenie szablonów TableLook

1. W oknie dialogowym szablonów TableLook, z listy plików wybierz szablon TableLook.
2. Kliknij przycisk **Edytuj**.
3. Zmodyfikuj właściwości tabeli w zakresie wybranych atrybutów i kliknij **OK**.
4. Kliknij przycisk **Zapisz**, aby zapisać zmodyfikowany szablon TableLook lub **Zapisz jako**, aby zapisać go jako nowy szablon.

- Modyfikacje w szablonie TableLook wpływają tylko na zaznaczoną tabelę przestawną. Zmodyfikowany szablon TableLook nie zostanie zastosowany do innych tabel na nim opartych, chyba że tabele te zostaną zaznaczone i dany szablon ponownie zostanie do nich zastosowany.
- W szablonach TableLook zapisane zostają tylko te właściwości tabel, które zostały zdefiniowane w oknie dialogowym Właściwości tabel. Szablony TableLook nie zawierają indywidualnych modyfikacji komórek.

Właściwości tabeli

Okno dialogowe Właściwości tabeli umożliwia określanie ogólnych właściwości tabeli, oraz określanie stylów komórek w różnych częściach tabeli. Można:

- Określać ogólne właściwości, na przykład ukrywanie pustych wierszy lub kolumn oraz ustawienia właściwości drukowania.
- Określać format i położenie symboli przypisów.
- Ustanawiać szczególne formaty dla komórek w obszarze danych, dla etykiet wierszy i kolumn i dla innych obszarów tabeli.
- Określać szerokość i kolor linii tworzących obramowania poszczególnych obszarów tabeli.

Zmiana właściwości tabeli przestawnej

1. Z menu wybierz:

   Format > **Właściwości tabeli...**
2. Wybierz kartę (**Ogólne**, **Przypisy**, **Formaty**, **Krawędzie** lub **Drukowanie**).
3. Wybierz żądane opcje.
4. Kliknij przycisk **OK** lub **Zastosuj**.

Nowe właściwości zostaną zastosowane do zaznaczonej tabeli przestawnej.

Właściwości tabeli: ogólne

Niektóre właściwości odnoszą się do całej tabeli. Można:

- Wyświetlać lub ukrywać puste wiersze lub kolumny (za puste wiersze lub kolumny należy uważać takie, w komórkach których nie wpisano żadnych danych).
- Sterować położeniem etykiet wierszy, które mogą być zagnieżdżone lub znajdować się w lewym górnym narożniku.
- Określać maksymalną i minimalną szerokość kolumny (wyrażaną w punktach).

Zmiana ogólnych właściwości tabeli:

1. Kliknij kartę **Ogólne**.
2. Wybierz żądane opcje.
3. Kliknij przycisk **OK** lub **Zastosuj**.

Wyświetlanie wierszy tabeli

Uwaga: Ta funkcja dotyczy tylko starszej wersji tabel.

Domyślnie tabele z dużą liczbą wierszy są wyświetlane w sekcjach po 100 wierszy. Aby wybrać liczbę wierszy wyświetlanych w tabeli:

1. Wybierz **Wyświetl tabelę wierszami**.
2. Kliknij **Wyświetlanie wierszy tabeli**.

   lub
3. W menu Widok aktywnej tabeli przestawnej wybierz **Wyświetl tabelę wierszami** i **Wyświetlanie wierszy tabeli**.

Wiersze do wyświetlenia: Wybieranie maksymalnej liczby wierszy do wyświetlenia za jednym razem. Elementy nawigacji umożliwiają poruszanie różnych sekcji tabeli. Minimalną wartością jest 10. Domyślną wartością jest 100.

Tolerancja na wdowy/sieroty. Wybór maksymalnej ilości wierszy w najbardziej wewnętrznym wymiarze wierszowym tabeli przy podziale wyświetlanych widoków tabeli. Na przykład, jeśli istnieje sześć kategorii w każdej grupie najbardziej wewnętrznego wymiaru wierszowego, określenie wartości sześć zapobiegnie dzieleniu jakiejkolwiek grupy w poprzek wyświetlanego widoku. To ustawienie może spowodować, że liczba wierszy w wyświetlanym widoku przekroczy określoną maksymalną liczbę wierszy do wyświetlenia.

Właściwości tabeli: uwagi

Karta Uwagi okna dialogowego Właściwości tabeli kontroluje formatowanie przypisów i tekstu komentarzy tabeli.

Przypisy. Właściwości dotyczące symboli przypisów obejmują styl i położenie względem tekstu.

- Styl symbolu przypisu może być liczbowy (1, 2, 3 itd.) lub literowy (a, b, c itd.).
- Symbole przypisów mogą być stawiane przy tekście jako indeksy górne lub dolne.

Tekst komentarza. Do każdej tabeli można dodać tekst komentarza.

- Tekst komentarza jest wyświetlany w podpowiedzi po umieszczeniu kursora nad tabelą w Edytorze raportów.
- Programy do odczytywania zawartości ekranu czytają tekst komentarza, gdy tabela jest aktywna.
- Podpowiedź w Edytorze raportów wyświetla tylko pierwsze 200 znaków komentarza, ale lektory ekranowe czytają cały tekst.
- Podczas eksportowania wyników do formatu HTML tekst komentarza jest używany jako tekst alternatywny.

Właściwości tabeli: formaty komórek

Na potrzeby formatowania tabela jest dzielona na obszary: tytuł, warstwy, etykiety narożników, etykiety wierszy, etykiety kolumn, dane, nagłówek i przypisy. Możliwe jest modyfikowanie formatów komórek zawartych w poszczególnych obszarach tabeli. Formaty komórek obejmują charakterystykę tekstu

(czcionka, rozmiar, kolor, styl), wyrównanie w poziomie i pionie, kolory tła oraz wewnętrzne marginesy komórek.

Formaty komórek mają zastosowanie do obszarów (kategorii informacji). Nie są one charakterystyczne dla pojedynczych komórek. To rozróżnienie jest istotne podczas przestawiania tabeli.

Na przykład

- Jeśli jako format komórek etykiet kolumny zostanie określona czcionka pogrubiona, etykiety kolumn będą wyświetlane taką czcionką bez względu na to, jakie informacje są obecnie wyświetlane w wymiarze kolumn. Po przeniesieniu pozycji z wymiaru kolumn do innego wymiaru nie zachowuje ona właściwości pogrubienia etykiet kolumn.
- Jeśli nagłówki kolumn zostaną pogrubione przez wyróżnienie komórek w uaktywnionej tabeli przestawnej i kliknięcie przycisku Pogrubienie na pasku narzędzi, to zawartość tych komórek zachowa pogrubienie, niezależnie od wymiaru, do którego zostaną przeniesione, zaś pogrubienie etykiety danej kolumny nie będzie miało zastosowania do innych elementów przeniesionych do jej wymiaru.

Zmiana formatowania komórek:

1. Wybierz kartę **Formaty**.
2. Wybierz obszar z listy rozwijanej lub kliknij go na przykładowym wizerunku tabeli.
3. Określ cechy danego obszaru. Ustawienia zostaną odzwierciedlone na przykładzie obok.
4. Kliknij przycisk **OK** lub **Zastosuj**.

Zmiany kolorów wierszy

Aby zastosować inny kolor tła i/lub tekstu w celu wyróżnienia wierszy w obszarze danych tabeli:

1. Z rozwijanej listy Obszar wybierz opcję **Dane**.
2. Wybierz (zaznacz) **Zmieniaj kolory wierszy** w grupie Kolor tła.
3. Wybierz kolory, które mają być używane do wyróżniania tła i tekstu wiersza.

Zmiana kolorów wierszy obejmuje tylko obszary danych w tabeli. Nie ma wpływu na obszary etykiet wierszy lub kolumn.

Właściwości tabeli: obramowania

Możliwe jest określenie stylu i koloru linii obramowania w poszczególnych miejscach tabeli. Jeśli jako styl zostanie wybrana opcja **Brak**, w danym położeniu nie będzie widoczna żadna linia.

Zmiana obramowania tabeli:

1. Kliknij kartę **Obramowanie**.
2. Wybierz położenie obramowania, klikając jego nazwę na liście lub klikając linię w przykładzie obok.
3. Wybierz styl linii lub pozycję **Brak**.
4. Wybierz kolor.
5. Kliknij przycisk **OK** lub **Zastosuj**.

Właściwości tabeli: drukowanie

Możliwe jest określanie następujących właściwości drukowania tabel przestawnych:

- Drukowanie wszystkich warstw lub tylko wierzchniej warstwy tabeli oraz drukowanie każdej warstwy na osobnej stronie.
- Zmniejszanie rozmiaru tabeli w poziomie lub w pionie, tak aby podczas wydruku mieściła się na stronie.
- Kontrolowanie wdów/sierot – czyli minimalnej dopuszczalnej liczby wierszy i kolumn, które będą się znajdować w każdej sekcji tabeli na wydruku, jeżeli tabela jest zbyt szeroka lub zbyt długa w stosunku do zdefiniowanego rozmiaru strony.

Uwaga: Jeśli długość tabeli odpowiada zdefiniowanej długości strony, lecz nie mieści się na niej, gdyż powyżej znajdują się inne wyniki, tabela zostanie automatycznie przeniesiona na nową stronę wydruku, bez względu na ustawienia wdów i sierot.

- Dołączanie tekstu kontynuacji tabeli, jeżeli tabela nie mieści się na jednej stronie. Tekst o kontynuacji można wyświetlać u dołu każdej strony i na początku strony następnej. Jeśli żadna z opcji nie zostanie zaznaczona, tekst o kontynuacji nie będzie wyświetlany.

Kontrola właściwości drukowania tabeli przestawnej:

1. Kliknij kartę **Drukowanie**.
2. Wybierz żądane opcje drukowania.
3. Kliknij przycisk **OK** lub **Zastosuj**.

Właściwości komórki

Właściwości komórki mają zastosowanie do zaznaczonej komórki. Możliwa jest zmiana czcionki, formatu wartości, wyrównania, marginesów i kolorów. Właściwości komórki mają wyższy priorytet niż właściwości tabeli, dlatego zmiana właściwości całej tabeli nie powoduje zmiany w osobno zmodyfikowanych właściwościach komórki.

Zmiana właściwości komórki:

1. Zaznacz komórki w tabeli.
2. Z menu Format lub menu kontekstowego wybierz pozycję **Właściwości komórki**.

Czcionka i tło

Karta Czcionka i tło pozwala ustawić styl i kolor czcionki oraz kolor tła wybranych komórek w tabeli.

Formatowanie wartości

Karta Format wartości pozwala określić format wartości w wybranych komórkach. Możliwe jest określenie formatu liczb, dat, czasu lub walut, a także ustalenie liczby wyświetlanych miejsc dziesiętnych.

Wyrównanie i marginesy

Karta Wyrównanie i Marginesy pozwala zdefiniować sposób wyrównania wartości w pionie i poziomie oraz górne, dolne, lewe i prawe marginesy wybranych komórek. **Mieszany** sposób wyrównania w poziomie powoduje wyrównanie zawartości każdej komórki odpowiednio do jej typu. Na przykład daty są wyrównywane do prawej strony, a wartości tekstowe do lewej.

Przypisy i nagłówki

Do tabeli można dodać przypisy i nagłówki. Można także ukryć przypisy lub nagłówki, zmienić symbole przypisów i ponownie ponumerować przypisy.

Dodawanie przypisów i nagłówków

Dodawanie nagłówka do tabeli:

1. Z menu Wstaw wybierz polecenie **Nagłówek**.

Do każdego elementu tabeli można dołączyć przypis. Dodawanie przypisu:

1. Kliknij tytuł, komórkę lub nagłówek w uaktywnionej tabeli przestawnej.
2. Z menu Wstaw wybierz polecenie **Przypis**.
3. W podanym obszarze wstaw tekst przypisu.

Wyświetlanie lub ukrywanie nagłówka

Ukrywanie nagłówka:

1. Wybierz nagłówek.
2. Z menu Widok wybierz pozycję **Ukryj**.

Wyświetlanie ukrytych nagłówków:

1. Z menu Widok wybierz pozycję **Pokaż wszystko**.

Ukrywanie lub wyświetlanie przypisów w tabeli

Ukrywanie przypisu:

1. Kliknij prawym przyciskiem myszy komórki, które zawierają odwołanie do przypisu i zaznacz w menu kontekstowym **Ukryj przypisy**.

   lub

2. W obszarze przypisów tabeli wybierz przypis i wybierz z menu kontekstowego **Ukryj**.

   Uwaga: W przypadku tabel w starszej wersji, wybierz obszar przypisów tabeli, wybierz z menu kontekstowego polecenie **Edytuj przypis**, a następnie dla wszystkich przypisów, które chcesz ukryć, usuń zaznaczenie właściwości Widoczne.

   Jeśli komórka zawiera wiele przypisów, użyj tej ostatniej metody, aby wybiórczo ukryć przypisy.

Aby ukryć wszystkie przypisy w tabeli:

1. Wybierz wszystkie przypisy w obszarze przypisów tabeli (za pomocą: kliknij i przeciągnij lub klawisza Shift+kliknięcie, aby zaznaczyć przypisy) i zaznacz w menu Widok **Ukryj**.

Wyświetlanie ukrytych przypisów:

1. Wybierz z menu Widok **Pokaż wszystkie przypisy**.

Znacznik przypisu

Okno dialogowe Znacznik przypisu służy do zmiany znaków użytych w odsyłaczu do przypisu. Domyślnie standardowymi symbolami przypisów są kolejne litery lub liczby, w zależności od ustawień właściwości tabeli. Możesz też przypisać specjalny symbol. Na symbole specjalne nie ma wpływu zmiana numeracji przypisów lub przełączenie między liczbami a literami symboli standardowych. Wyświetlanie liczb lub liter dla standardowych symboli i umiejscowienie symboli przypisu w indeksie dolnym lub górnym jest ustawiane w karcie Przypisy znajdującej się w oknie dialogowym Właściwości tabeli.

Zmiana symboli przypisów:

1. Zaznacz przypis.
2. Z menu **Format** wybierz opcję **Znacznik przypisu**.

Symbole specjalne są ograniczone do dwóch znaków. W obszarze przypisu tabeli przypisy z symbolami specjalnymi poprzedzają przypisy z sekwencjami liter i liczb. Zmiana na symbol specjalny może spowodować przeorganizowanie listy przypisów.

Przenumerowywanie przypisów

W wyniku przestawienia tabeli przez zamianę wierszy, kolumn i warstw może zostać naruszony porządek przypisów. W celu przenumerowania przypisów:

1. Z menu Format wybierz polecenie **Przenumeruj przypisy**.

Edytowanie przypisów w tabelach starszej wersji

W przypadku tabel starszej wersji użyj okna dialogowego „Edycja przypisów”, aby wprowadzić i modyfikować zawartość przypisu, ustawienia czcionki, zmienić symbole przypisu i wybiórczo ukryć lub usunąć przypisy.

Gdy wstawisz do tabeli starszej wersji nowy przypis, automatycznie otwiera się okno dialogowe Edycja przypisów. Należy użyć okna dialogowego „Edycja przypisów”, aby edytować istniejące przypisy (bez tworzenia nowego przypisu).

Symbol Domyślnie standardowymi symbolami przypisów są kolejne litery lub liczby, w zależności od ustawień właściwości tabeli. Aby przydzielić symbol specjalny, po prostu wpisz nową wartość symbolu w kolumnie Symbol. Na symbole specjalne nie ma wpływu zmiana numeracji przypisów lub przełączenie między liczbami a literami symboli standardowych. Wyświetlanie liczb lub liter dla standardowych symboli i umiejscowienie symboli przypisu w indeksie dolnym lub górnym jest ustawiane w karcie Przypisy znajdującej się w oknie dialogowym Właściwości tabeli. Więcej informacji można znaleźć w temacie “Właściwości tabeli: uwagi” na stronie 129.

Aby zmienić symbol specjalny z powrotem na symbol standardowy, w oknie dialogowym Edycja przypisów kliknij symbol prawym przyciskiem myszy, wybierz z menu kontekstowego **Znacznik przypisu** i w oknie dialogowym Symbol przypisów zaznacz Symbol standardowy.

Przypis. Zawartość przypisu. Wyświetlacz pokazuje bieżące ustawienia czcionki i tła. Ustawienia czcionki można zmienić dla pojedynczych przypisów, używając podrzędnego okna dialogowego Format. Więcej informacji można znaleźć w temacie “Ustawienia czcionki i koloru przypisu” na stronie 133. Jeden kolor tła stosowany jest do wszystkich przypisów; można go zmienić w karcie Czcionka i tło okna dialogowego Właściwości komórki. Więcej informacji można znaleźć w temacie “ Czcionka i tło” na stronie 131.

Widoczne. Domyślnie wszystkie przypisy są widoczne. Usuń zaznaczenie pola wyboru Widoczne, aby ukryć przypis.

Ustawienia czcionki i koloru przypisu

W przypadku starszej wersji tabel możesz użyć okna dialogowego Format, aby zmienić rodzinę czcionek, styl, rozmiar i kolor przynajmniej jednego wybranego przypisu.

1. W oknie dialogowym Edycja przypisów wybierz (kliknij) jeden lub więcej przypisów w siatce Przypisy.
2. Kliknij przycisk **Format**.

Wybrana rodzina czcionek, styl, rozmiar i kolory zostaną zastosowane do wszystkich wybranych przypisów.

Kolor tła, wyrównanie oraz marginesy można ustawić w oknie dialogowym Właściwości komórki i zastosować do wszystkich przypisów. Ustawień tych nie można zmienić dla pojedynczych przypisów. Więcej informacji można znaleźć w temacie “ Czcionka i tło” na stronie 131.

Szerokość komórek danych

Okno dialogowe Ustaw szerokość komórek służy do nadawania komórkom danych jednakowej szerokości.

Ustawienie szerokości wszystkich komórek danych:

1. Z menu wybierz:

   Format > **Ustaw szerokość komórek danych...**
2. Wprowadź wartość szerokości komórki.

Zmiana szerokości kolumny

1. Przeciągnij i upuść krawędź kolumny.

Wyświetlanie ukrytych krawędzi tabeli przestawnej

W przypadku tabel z wieloma niewidocznymi krawędziami możliwe jest wyświetlanie krawędzi ukrytych. Dzięki temu wykonanie wielu operacji, takich jak zmiana szerokości kolumn, może stać się łatwiejsze.

1. Z menu Widok wybierz pozycję **Linie siatki**.

Zaznaczanie wierszy, kolumn i komórek w tabeli przestawnej

Możesz zaznaczyć cały wiersz, kolumnę lub określony zestaw komórek danych i etykiet.

Aby wybierać wiele komórek:

Wybierz > **Komórki danych i opis**

Drukowanie tabel przestawnych

Na wygląd drukowanych tabel przestawnych może wpływać kilka czynników, które można kontrolować za pomocą atrybutów tabeli przestawnej.

- W przypadku wielowymiarowych tabel przestawnych (tabel z warstwami) można wydrukować wszystkie warstwy lub tylko warstwę wierzchnią (widoczną na ekranie). Więcej informacji można znaleźć w temacie “Właściwości tabeli: drukowanie” na stronie 130.
- W przypadku długich lub szerokich tabel przestawnych można automatycznie dopasować rozmiar tabeli do rozmiaru strony lub określić położenie podziałów tabeli i podziałów strony. Więcej informacji można znaleźć w temacie “Właściwości tabeli: drukowanie” na stronie 130.

Aby zobaczyć, jak będzie wyglądać wydrukowana tabela przestawna, skorzystaj z polecenia Podgląd wydruku w menu Plik.

Kontrola podziałów tabeli w przypadku szerokich i długich tabel

Tabele przestawne, które są zbyt szerokie lub zbyt długie, aby można je było wydrukować na stronie o zadanych wymiarach, są automatycznie dzielone i drukowane w kilku częściach Można:

- Określić położenia wierszy i kolumn, w których ma nastąpić podział obszernej tabeli.
- Określić wiersze i kolumny, które powinny być zachowane razem, gdy następuje podział tabeli.
- Dostosować wielkość obszernych tabel tak, aby pasowały do zdefiniowanego rozmiaru strony.

Aby określić podziały wierszy i kolumn w drukowanych tabelach przestawnych:

1. Uaktywnij tabelę przestawną.
2. Kliknij dowolną komórkę kolumny na lewo od miejsca, w którym ma wystąpić podział lub kliknij dowolną komórkę wiersza powyżej miejsca, w którym ma wystąpić podział.
3. Z menu wybierz kolejno następujące pozycje:

   Format > **Podziały** > **Podział pionowy**

   lub

   Format > **Podziały** > **Podział poziomy**

1. Uaktywnij tabelę przestawną.
2. Kliknij dowolną komórkę kolumny na lewo od miejsca, w którym ma wystąpić podział lub kliknij dowolną komórkę wiersza powyżej miejsca, w którym ma wystąpić podział.
3. Z menu wybierz kolejno następujące pozycje:

   Format > **Podziały** > **Podział pionowy**

   lub

   Format > **Podziały** > **Podział poziomy**

Aby określić wiersze lub kolumny, które mają być zachowane razem:

1. Zaznacz nagłówki wierszy lub kolumn, które powinny być zachowane razem. Aby zaznaczyć kilka etykiet wierszy lub kolumn, kliknij i przeciągnij lub kliknij je przy wciśniętym klawiszu Shift.
2. Z menu wybierz kolejno następujące pozycje:

   Format > **Podziały** > **Scalaj w tym miejscu**

Aby przejrzeć punkty przerywające i scalone grupy:

1. Z menu wybierz kolejno następujące pozycje:

Format > **Podziały** > **Pokaż podziały**

Podziały są pokazane jako linie pionowe lub poziome. Scalone grupy pojawiają się jako oznaczone na szaro prostokątne obszary ograniczone ciemniejszą linią.

Uwaga: Wyświetlanie punktów przerywających i scalonych grup nie jest obsługiwane przez starsze wersje tabel.

Usuwanie punktów przerywających i scalonych grup

Aby usunąć jeden punkt przerywający:

1. Kliknij dowolną komórkę kolumny znajdującą się z lewej strony pionowego punktu przerywającego lub kliknij dowolną komórkę wiersza znajdującą się nad poziomym punktem przerywającym.
2. Z menu wybierz kolejno następujące pozycje:

   Format > **Podziały** > **Wyczyść punkt podziału lub grupę**

   Usuwanie scalonych grup:
3. Zaznacz etykietę kolumny lub wiersza określające grupę.
4. Z menu wybierz kolejno następujące pozycje:

   Format > **Podziały** > **Wyczyść punkt podziału lub grupę**

Wszystkie punkty przerywające i grupy scalone zostają automatycznie usunięte przy przestawianiu lub ponownym porządkowaniu dowolnego wiersza i kolumny. Zachowanie to nie dotyczy starszej wersji tabel.

Tworzenie wykresu na podstawie tabeli przestawnej

1. Kliknij dwukrotnie tabelę przestawną, aby ją uaktywnić.
2. Wybierz wiersze, kolumny lub komórki, które mają być wyświetlone na wykresie.
3. Kliknij prawym przyciskiem myszy w dowolnym miejscu zaznaczonego obszaru.
4. Z menu kontekstowego wybierz pozycję **Utwórz wykres** i wybierz typ wykresu.

Starsze wersje tabel

Możesz wybrać renderowanie tabel jako tabele starszej wersji (nazywane w wersji 19 tabelami pełno funkcyjnymi), które będą później w pełni kompatybilne z wersjami IBM SPSS Statistics wcześniejszymi niż 20. Tabele starszej wersji mogą być powolne w renderowaniu i są zalecane tylko w przypadku, gdy wymagana jest kompatybilność z wersjami starszymi niż 20. Informacje na temat tworzenia starszych wersji tabel można znaleźć w sekcji “Opcje tabel przestawnych” na stronie 137.

Opcje

Opcje

Opcje sterujące różnymi ustawieniami.

Zmiana ustawień okna Opcje

1. Z menu wybierz kolejno następujące pozycje:

Edycja > **Opcje...**

2. Kliknij kartę ustawień w celu wyświetlenia ustawień, które chcesz zmienić.
3. Wprowadź zmiany.
4. Kliknij przycisk **OK** lub **Zastosuj**.

Opcje ogólne

Maksymalna liczba wątków

Liczba wątków używanych przez procedury wielowątkowe podczas obliczania wyników. Ustawienie **Automatycznie** jest oparte na liczbie dostępnych rdzeni przetwarzania. Jeśli podczas uruchamiania procedur wielowątkowych chcesz udostępnić więcej zasobów przetwarzania dla innych aplikacji, określ niższą wartość. Ta opcja jest wyłączona w trybie analiz rozproszonych.

Raport

Wyświetlaj zero wiodące dla wartości dziesiętnych. Wyświetla zera wiodące dla wartości liczbowych, które zawierają tylko część dziesiętną. Na przykład gdy wyświetlane są zera wiodące, wartość .123 jest wyświetlana jako 0.123. To ustawienie nie wpływa na wartości liczbowe, które mają format walutowy lub procentowy. Zera wiodące nie są uwzględniane podczas zapisywania danych do pliku zewnętrznego, ale nie dotyczy to plików w formacie Fixed ASCII (*.dat).

System pomiarowy. Opcja ta umożliwia wybór jednostki miary (punkty, cale lub centymetry), która określa atrybuty, takie jak marginesy komórek tabel przestawnych, szerokości komórek, czy odstępy między tabelami na wydrukach.

Opcje raportów

Zmiany ustawień opcji wyświetlania raportów dotyczą wyłącznie raportów utworzonych po wprowadzeniu tych zmian. Zmiany tych ustawień nie mają wpływu na raport już wyświetlony w Edytorze raportów.

Stan raportu wynikowego
Ta grupa opcji umożliwia wybranie elementów, które będą automatycznie wyświetlane lub ukrywane w wynikach działania każdej procedury, a także określenie początkowego wyrównania tych elementów. Istnieje możliwość sterowania wyświetlaniem następujących elementów: dziennika, ostrzeżeń, notatek, tytułów, tabel przestawnych, wykresów, diagramów drzew i wyników tekstowych. Można również włączać i wyłączać wyświetlanie komend w dzienniku. Z pliku dziennika można kopiować składnię komendy i zapisywać ją w pliku komend.

Uwaga: Wszystkie elementy raportu wyświetlane są z wyrównaniem do lewej. Ustawienia wyrównania mają wpływ tylko na wydruki. Elementy wyśrodkowane lub wyrównane do prawej są wyróżnione niewielkim symbolem.

Tytuł
Służy do określania stylu, rozmiaru i koloru czcionki wykorzystywanej w tytułach nowych raportów. Lista **Rozmiar** czcionek zawiera zestaw predefiniowanych rozmiarów, ale można ręcznie wprowadzać inne obsługiwane rozmiary.

Tytuł strony
Umożliwia określenie stylu, rozmiaru i koloru czcionki wykorzystywanej w tytułach nowych stron i tytułach stron generowanych za pomocą komend TITLE lub SUBTITLE lub utworzonych za pomocą opcji **Nowy tytuł strony** z menu **Wstaw**. Lista **Rozmiar** czcionek zawiera zestaw predefiniowanych rozmiarów, ale można ręcznie wprowadzać inne obsługiwane rozmiary.

Wynik tekstowy
Rodzaj czcionki używanej w wynikach tekstowych. W wynikach tekstowych używana jest czcionka nieproporcjonalna (o stałej szerokości). Jeśli zostanie wybrana czcionka proporcjonalna, to wyniki zebrane w tabelach nie będą wyrównane poprawnie. Lista **Rozmiar** czcionek zawiera zestaw predefiniowanych rozmiarów, ale można ręcznie wprowadzać inne obsługiwane rozmiary.

Domyślne ustawienia strony
Kontroluje domyślne opcje orientacji i marginesy drukowania.

Opcje tabel przestawnych

Opcje: tabele przestawne ustawiają różne opcje wyświetlania tabel przestawnych.

Szablon TableLook

Z listy plików należy wybrać szablon TableLook i kliknąć przycisk **OK** lub **Zastosuj**. Można korzystać z szablonów TableLook dostarczonych z programem IBM SPSS Statistics lub za pomocą Edytora tabel przestawnych utworzyć własne (z menu Format należy wybrać pozycję Szablony **TableLook**).

- **Przeglądaj.** Umożliwia wybranie szablonu TableLook z innego katalogu.
- **Ustaw katalog TableLook.** Pozwala na zmianę domyślnego katalogu szablonów TableLook. Użyj opcji **Przeglądaj**, aby przejść do katalogu, którego chcesz używać, a następnie wybierz w tym katalogu opcję **Ustaw katalog TableLook**.

Uwaga: Szablonów TableLook utworzonych we wcześniejszych wersjach systemu IBM SPSS Statistics nie można używać w wersji 16.0 i nowszych.

Szerokość kolumn

Te opcje umożliwiają sterowanie automatycznym dostosowywaniem szerokości kolumn w tabelach przestawnych.

- **Skoryguj tylko dla etykiet.** Zaznaczenie tej opcji powoduje dopasowanie szerokości kolumny do szerokości jej etykiety. Dzięki temu tabele są bardziej zwarte, lecz wartości danych szersze niż etykieta mogą zostać obcięte.
- **Dopasuj we wszystkich tabelach do etykiet i danych.** Szerokość kolumn jest dostosowywana do większej z dwóch wartości: etykiety kolumny lub największej wartości danych. Zwiększa to szerokość tabel, ale gwarantuje, że wyświetlone zostaną wszystkie wartości.

Domyślny tryb edycji tabel

Ta opcja umożliwia określenie sposobu aktywacji tabel przestawnych w oknie Edytora raportów lub w osobnym oknie. Domyślnie dwukrotne kliknięcie tabeli przestawnej aktywuje ją w oknie programu Viewer (nie dotyczy bardzo dużych tabel). Można wybrać, aby tabele przestawne były uaktywniane w osobnym oknie lub wybrać ustawienie rozmiaru tak, aby mniejsze tabele otwierane były w oknie Edytora raportu, a większe w oddzielnym oknie.

Kopiowanie szerokich tabel do schowka w formacie RTF

Gdy tabele przestawne są wklejane w formacie Word/RTF, za szerokie tabele w stosunku do dokumentu zostaną zawinięte, pomniejszone w celu dopasowania do szerokości dokumentu albo pozostawione bez zmian.

Opcje wyników

Opcje wyników sterują ustawieniami domyślnymi dla wielu opcji danych wyjściowych.

Dostępność lektora ekranowego. Kontroluje sposób czytania wierszy i kolumn tabel przestawnych przez lektory ekranowe. Możliwe jest czytanie pełnych etykiet wierszy i kolumn dla każdej komórki danych lub czytanie tylko etykiet, które zmieniają się, gdy użytkownik przechodzi pomiędzy komórkami danych w tabeli.

Rozdział 8. Obsługa braków danych

Braki danych — przegląd

W fazie przygotowywania danych do eksploracji często chcemy wyeliminować braki, zastępując je jakimiś wartościami. **Braki danych** to wartości w zbiorze danych, które są nieznane, nie zostały zebrane lub zostały nieprawidłowo wprowadzone. Zwykle braki danych nie są poprawnymi wartościami zmiennych, w których występują. Na przykład zmienna *Płeć* powinna zawierać wartości *M* albo *K*. Napotykając w niej wartości *T* lub *Z*, możemy bezpiecznie przyjąć, że są to wartości niepoprawne, które powinny być traktowane jako brak danych. Podobnie, ujemna wartość zmiennej *Wiek* jest nieprawidłowa i powinna być interpretowana jako brak danych. Często takie oczywiście błędne wartości wprowadza się celowo (lub pozostawia się zmienne bez wartości) podczas ankiety, aby zasygnalizować w ten sposób brak odpowiedzi. Niekiedy chcemy dokładniej przeanalizować braki danych, aby ustalić, czy brak odpowiedzi (np. odmowa podania wieku przez respondenta) ma wpływ na predykcję określonego wyniku.

Niektóre techniki modelowania lepiej niż inne radzą sobie z brakami danych. Na przykład techniki C5.0 i Apriori działają dobrze w obecności wartości jawnie zadeklarowanych jako braki w węźle Typy. Inne techniki modelowania mają trudności z obsługą braków danych; w ich przypadku braki wydłużają czas szkolenia i prowadzą do generowania mniej dokładnych modeli.

Istnieje kilka typów braków danych rozpoznawanych przez program IBM SPSS Modeler:

- **Null lub systemowe braki danych.** Są to wartości niełańcuchowe pozostawione jako puste w bazie danych lub w pliku źródłowym i niezdefiniowane jako „brakujące” w węźle źródłowym ani w węźle wprowadzania danych. Systemowe braki danych są wyświetlane jako wartości **$null$**. Należy zwrócić uwagę, że puste łańcuchy tekstowe nie są traktowane jako wartości null w programie IBM SPSS Modeler, mimo że mogą być one traktowane jako wartości null przez niektóre bazy danych.
- **Puste łańcuchy i białe znaki.** Puste łańcuchy tekstowe i białe znaki (łańcuchy bez widocznych znaków) są traktowane odmiennie niż wartości null. Puste łańcuchy są w większości przypadków traktowane jako równoważne białym znakom. Na przykład po wybraniu opcji traktowania białych znaków jako pustych w węźle źródłowym albo w węźle Typ to ustawienie ma zastosowanie także do pustych łańcuchów.
- **Puste lub zdefiniowane przez użytkownika braki danych.** Są to wartości takie jak `nieznane`, 99 czy −1, zdefiniowane jawnie w węźle źródłowym lub w węźle Typ jako braki danych. Opcjonalnie można także wybrać traktowanie wartości null i białych znaków jako wartości pustej, co pozwala oznaczyć je z myślą o specjalnym ich traktowaniu i wykluczeniu ich z większości obliczeń. Można na przykład użyć funkcji `@BLANK` do traktowania tych wartości, wraz z brakami danych innego typu, jako wartości pustej.

Wczytywanie danych mieszanych. Należy zwrócić uwagę, że podczas wczytywania zmiennych z liczbowym typem składowania (liczby całkowite, rzeczywiste, czas, znacznik czasu lub data) wszelkie wartości nieliczbowe są zamieniane na *null* lub *braki systemowe*. Wynika to z faktu, że w odróżnieniu od niektórych aplikacji produkt nie zezwala na przechowywanie w zmiennej danych różnego typu. Aby uniknąć takiej automatycznej zamiany, wszelkie zmienne z danymi mieszanymi należy wczytywać jako łańcuchy, w razie potrzeby zmieniając typ składowania w węźle źródłowym lub aplikacji zewnętrznej.

Odczytywanie pustych łańcuchów Oracle. W przypadku odczytu z lub zapisu do bazy danych Oracle należy pamiętać, że inaczej niż w przypadku IBM SPSS Modeler i inaczej niż w przypadku większości innych baz danych, Oracle traktuje i przechowuje puste wartości strumienia jako równoważne wartościom null. Oznacza to, że te same dane wyodrębnione z bazy danych Oracle mogą zachowywać się w odmienny sposób w przypadku wyodrębnienia z pliku lub innej bazy danych, zaś dane mogą zwracać odmienne wyniki.

Obsługa braków danych

Sposób traktowania braków danych należy ustalić, biorąc pod uwagę znajomość uwarunkowań biznesowych lub wiedzę specjalistyczną z dziedziny będącej przedmiotem analizy. Aby skrócić czas szkolenia i zwiększyć dokładność, celowe może być usunięcie wartości pustych ze zbioru danych. Z

drugiej strony obecność wartości pustych może doprowadzić do ujawnienia nowych szans biznesowych lub dodatkowych spostrzeżeń. Wybierając technikę, należy wziąć pod uwagę następujące cechy danych:

- wielkość zbioru danych;
- liczbę zmiennych zawierających wartości puste;
- ilość brakujących informacji.

Można wskazać zasadniczo dwie strategie postępowania:

- Można wykluczyć zmienne lub rekordy z brakami danych.
- Można podstawić, zastąpić lub wymusić wartości brakujące, stosując w tym celu różne metody.

Realizację obu tych strategii można w dużej mierze zautomatyzować, korzystając z węzła Audyt danych. Można na przykład wygenerować węzeł filtrowania, który będzie wykluczał zmienne ze zbyt dużą liczbą braków danych, by były użyteczne w modelowaniu, i wygenerować Superwęzeł, który będzie podstawiał brakujące wartości wszystkich pozostałych zmiennych. Właśnie w takich sytuacjach przejawia się prawdziwy potencjał audytu danych — pozwala on bowiem nie tylko oceniać obecny stan danych, lecz także podejmować działania na podstawie wyników tej oceny.

Obsługa rekordów z brakami danych

Jeśli większość braków danych skupiona jest w niewielkiej liczbie rekordów, to można po prostu wykluczyć te rekordy. Na przykład banki zwykle przechowują szczegółowe i kompletne akta swoich kredytobiorców. Jeśli jednak przy udzielaniu kredytów własnym pracownikom bank jest mniej restrykcyjny, w danych o kredytach zaciągniętych przez pracowników niektóre pola mogą być puste. W takim przypadku można poradzić sobie z brakami danych na dwa sposoby:

- Można usunąć rekordy pracowników za pomocą węzła selekcji.
- Jeśli zbiór danych jest obszerny, można odrzucić wszystkie rekordy z wartościami pustymi.

Obsługa zmiennych z brakami danych

Jeśli większość braków danych skupiona jest w niewielkiej liczbie zmiennych, to można zająć się tymi brakami na poziomie pól, a nie na poziomie rekordów. Ta strategia umożliwi także eksperymentowanie ze względną ważnością poszczególnych zmiennych przed wybraniem strategii postępowania z brakami danych. Jeśli zmienna jest nieważna w modelowaniu, prawdopodobnie nie warto jej zachować, niezależnie od tego, ile braków danych w niej występuje.

Załóżmy na przykład, że firma zajmująca się badaniami rynku zebrała dane za pomocą ogólnego formularza złożonego z 50 pytań. Dwa z tych pytań dotyczą wieku i przekonań politycznych, czyli informacji, których wielu respondentów nie chciało podać. W takim przypadku w zmiennych *Wiek* i *Przekonania_polityczne* może występować wiele braków danych.

Poziom pomiaru zmiennej

Wybierając metodę postępowania, należy także brać pod uwagę poziom pomiaru zmiennych z brakami danych.

Zmienne liczbowe. W przypadku zmiennych typu liczbowego, na przykład *Ciągłych*, należy zawsze przed rozpoczęciem budowania modelu wyeliminować wszelkie wartości nieliczbowe, ponieważ wiele modeli nie będzie działać na zmiennych liczbowych z wartościami pustymi.

Zmienne jakościowe. W przypadku zmiennych jakościowych, np. *Nominalnych* i typu *Flaga*, modyfikowanie braków danych nie jest konieczne, ale zwiększy dokładność modelu. Na przykład model ze zmienną *Płeć* będzie działał na wartościach nic nie znaczących, takich jak *T* lub *Z*, ale usunięcie wszystkich wartości innych niż *M* i *K* zwiększy dokładność modelu.

Monitorowanie lub usuwanie pól

Aby monitorować pola ze zbyt dużą liczbą braków danych, można zastosować kilka różnych metod:

- Można użyć węzła Audyt danych w celu odfiltrowania zmiennych według kryterium jakości.

- Można użyć węzła Dobór predyktorów, aby monitorować zmienne z większym od określonego odsetkiem braków danych i uszeregować zmienne na podstawie ważności względem określonej zmiennej przewidywanej.
- Zamiast usuwać pola, można użyć węzła Typ w celu przypisania zmiennej roli **Brak**. Takie zmienne nie zostaną usunięte z danych, ale będą wyłączone z procesów modelowania.

Obsługa rekordów z systemowymi brakami danych

Czym są systemowe braki danych?

Systemowe braki danych są odzwierciedleniem wartości danych, które nie są znane lub nie mają zastosowania. W bazach danych takie wartości nazywane są często wartościami *NULL*.

Systemowe braki danych nie są tożsame z wartościami pustymi. Wartości puste są zwykle zdefiniowane w węźle Typ jako konkretne wartości lub przedziały wartości, które można traktować jako braki danych zdefiniowane przez użytkownika. Wartości puste w kontekście modelowania traktowane są inaczej.

Konstruowanie systemowych braków danych

Systemowe braki danych mogą występować w danych wczytywanych ze źródła danych (na przykład tabele bazy danych mogą zawierać wartości *NULL*).

Systemowe braki danych można konstruować, używając w wyrażeniach wartości **undef**. Na przykład następujące wyrażenie CLEM zwraca Wiek, jeśli jest mniejszy lub równy 30, albo brak danych, jeśli jest większy niż 30:

```
if Wiek > 30 then undef else Wiek endif
```

Braki danych mogą być także wynikiem łączenia zewnętrznego, dzielenia przez zero, próby obliczenia pierwiastka kwadratowego z liczby ujemnej oraz innych sytuacji.

Sposób prezentacji systemowych braków danych

Systemowe braki danych są uwidaczniane w tabelach i innych formach wyników jako wartości `$null$`.

Sprawdzanie, czy występują systemowe braki danych

Funkcja specjalna **@NULL** zwraca `true`, jeśli wartość podana jako argument jest systemowym brakiem danych, na przykład::

```
if @NULL(MojaZmienna) then 'Jest brakiem' else 'Nie jest brakiem' endif
```

Systemowe braki danych przekazywane do funkcji

Przekazanie systemowego braku danych do funkcji zwykle powoduje, że wynikiem również jest brak danych. Na przykład, jeśli wartość zmiennej `f1` w konkretnym wierszu jest systemowym brakiem danych, to wyrażenie `log(f1)` dla tego wiersza także da w wyniku systemowy brak danych. Wyjątkiem jest funkcja **@NULL**.

Systemowe braki danych w wyrażeniach zawierających operatory arytmetyczne

Zastosowanie operatora arytmetycznego do systemowego braku danych daje w wyniku systemowy brak danych. Na przykład, jeśli wartość zmiennej `f1` w konkretnym wierszu jest systemowym brakiem danych, to wyrażenie `f1 + 10` dla tego wiersza także da w wyniku systemowy brak danych.

Systemowe braki danych w wyrażeniach zawierających operatory logiczne

Gdy braki danych używane są w wyrażeniach zawierających operatory logiczne, obowiązuje logika trójwartościowa (*prawda*, *fałsz* i *brak*), którą można opisać w tabelach prawdy. Poniżej przedstawiono tabele prawdy dla typowych operatorów logicznych *not*, *and* oraz *or*.

Tabela 5. Tabela prawdy dla operatora NOT

Operand	**Operand NOT**
prawda	fałsz
fałsz	prawda
brak danych	brak danych

Tabela 6. Tabela prawdy dla operatora AND

Operand1	**Operand2**	**Operand1 AND Operand2**
prawda	prawda	prawda
prawda	fałsz	fałsz
prawda	brak danych	brak danych
fałsz	prawda	fałsz
fałsz	fałsz	fałsz
fałsz	brak danych	fałsz
brak danych	prawda	brak danych
brak danych	fałsz	fałsz
brak danych	brak danych	brak danych

Tabela 7. Tabela prawdy dla operatora OR

Operand1	**Operand2**	**Operand1 OR Operand2**
prawda	prawda	prawda
prawda	fałsz	prawda
prawda	brak danych	prawda
fałsz	prawda	prawda
fałsz	fałsz	fałsz
fałsz	brak danych	brak danych
brak danych	prawda	prawda
brak danych	fałsz	brak danych
brak danych	brak danych	brak danych

Systemowe braki danych w wyrażeniach zawierających operatory porównywania

Porównanie systemowego braku danych z wartością niebędącą systemowym brakiem danych daje w wyniku systemowy brak danych, a nie wartość prawda albo fałsz. Systemowe braki danych można porównywać ze sobą; dwa systemowe braki danych uznawane są za równe.

Systemowe braki danych w wyrażeniach if/then/else/endif

Gdy wyrażenie warunkowe zwraca systemowy brak danych, to wynikiem całego wyrażenia warunkowego jest wartość właściwa dla klauzuli `else`.

Systemowe braki danych w węźle selekcji

Gdy dla konkretnego rekordu wyrażenie selekcji daje w wyniku brak danych, rekord nie jest generowany z węzła selekcji (zasada ta obowiązuje zarówno w trybie uwzględniania, jak i odrzucania).

Systemowe braki danych w węźle łączenia

W przypadku łączenia przy użyciu klucza wszelkie rekordy z systemowymi brakami danych w polu klucza nie są łączone.

Systemowe braki danych w agregacjach

Podczas agregacji danych w kolumnach braki danych nie są uwzględniane w obliczeniach. Na przykład w kolumnie z trzema wartościami { 1, 2 i undef } suma wartości kolumny wynosi 3, zaś średnia wynosi 1,5.

Podstawianie lub wypełnianie braków danych

W przypadkach, gdy mamy tylko kilka braków danych, użyteczna może być możliwość wstawienia wartości, które zastąpią te braki. Operację tę można przeprowadzić z raportu Audyt danych, który umożliwia określenie opcji dla konkretnych zmiennych, a następnie wygenerować Superwęzeł, który podstawi wartości przy użyciu jednej z wielu dostępnych metod. Jest to sposób zapewniający największą elastyczność, który ponadto umożliwia określenie sposobu postępowania z dużą liczbą zmiennych w jednym węźle.

Wprowadzanie braków danych umożliwiają następujące metody:

Stała. Zastępuje stałą wartość (średnią dla zmiennej, środek zakresu lub wskazaną stałą).

Losowa. Podstawia wartość losową w oparciu o rozkład normalny lub jednostajny.

Wyrażenie. Umożliwia określenie wyrażenia niestandardowego. Można na przykład zastąpić wartości zmienną globalną utworzoną za pośrednictwem węzła wartości globalnych.

Algorytm. Podstawia wartość przewidywaną przez model w oparciu o algorytm C&RT. Dla każdej zmiennej wprowadzonej tą metodą dostępny będzie osobny model C&RT oraz węzeł wypełniania zastępujący wartości puste i wartości null wartością przewidywaną przez model. Następnie stosowany jest węzeł filtrowania umożliwiający usuwanie zmiennych predykcyjnych wygenerowanych przez model.

Alternatywnym sposobem jest wymuszenie wartości konkretnych zmiennych za pomocą węzła Typ — wartości niedozwolone dla danego typu zmiennej są wysyłane do kolumny *Sprawdź* w celu **Wymuszenia** wartości zmiennych, które obecnie mają wartości puste i wymagają zmiany.

Funkcje języka CLEM do obsługi braków danych

Istnieje kilka funkcji służących do obsługi braków danych. Następujące funkcje używane są często w węzłach selekcji i wypełniania do usuwania lub zastępowania braków danych:

- `count_nulls(LISTA)`
- `@BLANK(ZMIENNA)`
- `@NULL(ZMIENNA)`
- `undef`

Funkcji `@` można używać razem z funkcją `@FIELD` do wykrywania wartości pustych lub null w jednej lub wielu zmiennych. Zmienne mogą być po prostu oznaczane flagą w wypadku obecności wartości pustych lub null albo wypełniane wartościami zastępczymi bądź używane w różnych innych operacjach.

Możliwe jest określenie liczebności wartości null wśród zmiennych na liście, na przykład:

```
count_nulls(['okreskarty' 'okreskarty2' 'okreskarty3'])
```

W przypadku używania funkcji akceptujących listę zmiennych jako argumentu wejściowego można używać funkcji specjalnych `@FIELDS_BETWEEN` i `@FIELDS_MATCHING`, tak jak przedstawiono to w poniższym przykładzie:

```
count_nulls(@FIELDS_MATCHING('karta*'))
```

Funkcji `undef` można używać do przypisywania zmiennym systemowych braków danych prezentowanych jako **$null$**. Na przykład, aby zastąpić brakiem danych dowolną wartość liczbową, można użyć instrukcji warunkowej, takiej jak:

```
if not(Wiek > 17) or not(Wiek < 66) then undef else Wiek endif
```

Spowoduje ona zastąpienie brakami danych (prezentowanymi jako **$null$**) wszelkich wartości niemieszczących się w określonym przedziale. Korzystając z funkcji `not()`, można wychwytywać wszystkie wartości liczbowe inne niż określone, w tym wartości ujemne. Więcej informacji można znaleźć w temacie “Funkcje obsługujące wartości puste i null” na stronie 198.

Uwaga na temat odrzucania rekordów

Używając węzła selekcji do odrzucania rekordów, należy pamiętać, że w instrukcjach używana jest logika trójwartościowa, a wartości null są automatycznie uwzględniane w instrukcjach selekcji. Aby wykluczyć wartości null (systemowe braki danych) z wyrażenia selekcji, należy jawnie tego zażądać, używając w wyrażeniu klauzuli `and not`. Na przykład, aby wyselekcjonować i uwzględnić wszystkie rekordy, w których typem leku jest Lek C, należałoby użyć instrukcji:

```
Lek = 'lekC' and not(@NULL(Lek))
```

Wcześniejsze wersje produktu w takich sytuacjach wykluczały wartości null.

Rozdział 9. Tworzenie wyrażeń CLEM

Informacje o produkcie CLEM

Język CLEM (ang. Control Language for Expression Manipulation) to oferujący imponujące możliwości język do analizowania i manipulowania danymi przechodzącymi przez strumienie IBM SPSS Modeler. Narzędzia do eksploracji danych wykorzystują język CLEM w szerokim zakresie w operacjach strumieniowych, do wykonywania zadań zarówno tak prostych, jak wyliczanie zysku na podstawie kosztów i przychodów, jak i tak złożonych, jak transformowanie treści blogów w zestaw zmiennych i rekordów zawierający użyteczne informacje.

Język CLEM jest używany w programie IBM SPSS Modeler do:

- Porównywania i oceny warunków dotyczących zmiennych w rekordach.
- Wyliczania wartości dla nowych zmiennych.
- Wyliczania nowych wartości dla istniejących zmiennych.
- Uzasadniania kolejności rekordów.
- Wstawiania danych z rekordów do raportów.

Wyrażenia CLEM są nieodzowne do przygotowania danych w programie IBM SPSS Modeler i mogą być używane w szerokim zakresie węzłów — od operacji na rekordach i zmiennych (Selekcja, Równoważenie, Wypełnienie) po wykresy i dane wynikowe (Analiza, Raport, Tabela). Można na przykład użyć wyrażenia CLEM w formule, takiej jak iloraz.

Wyrażenia CLEM mogą być również używane na potrzeby globalnego wyszukiwania i zastępowania. Na przykład wyrażenie `@NULL(@FIELD)` może być używane w węźle wypełniania do zastąpienia **braków danych** o wartości całkowitej równej 0. (W celu zastąpienia **brakujących wartości użytkownika**, zwanych również wartościami pustymi, należy użyć funkcji `@BLANK`.)

Możliwe jest także tworzenie bardziej złożonych wyrażeń CLEM. Można na przykład wyliczyć nowe zmienne w oparciu o warunkowy zestaw reguł, na przykład jako nową kategorię wartości utworzoną na podstawie następujących wyrażeń: `If: CardID = @OFFSET(CardID,1), Then: @OFFSET(ValueCategory,1), Else: 'exclude'`.

W tym przykładzie funkcję `@OFFSET` zastosowano do wyrażenia następującej zależności: „Jeśli wartość zmiennej *CardID* dla danego rekordu jest taka sama, jak dla poprzedniego rekordu, wówczas zwracaj wartość zmiennej o nazwie *ValueCategory* dla poprzedniego rekordu. W przeciwnym wypadku przypisz łańcuch „exclude”. Innymi słowy, jeśli *CardID* dla sąsiednich rekordów są takie same, powinny one zostać przypisane do tej samej kategorii wartości. (Rekordy z łańcuchem exclude można później selekcjonować, korzystając z węzła Selekcja.)

CLEM — przykłady

Poniżej przedstawiono prawidłowe przykłady składni i typów wyrażeń, które można formułować w języku CLEM.

Wyrażenia proste

Formuły mogą mieć prostą postać, taką jak niniejsza formuła wyliczająca nową zmienną w oparciu o wartości pól *Przed* i *Po*:

```
(Przed - Po) / Przed * 100.0
```

Należy zauważyć, że nazwy zmiennych nie są ujęte w cudzysłowy, gdy odnoszą się do wartości zmiennej.

Poniższe wyrażenie po prostu zwraca logarytm każdej z wartości zmiennej *wynagrodzenie*.

```
log(wynagrodzenie)
```

Wyrażenia złożone

Wyrażenia mogą jednak być również długie i bardziej złożone. Poniższe wyrażenie zwraca wartość *true*, jeśli wartość dwóch zmiennych (*$KX-Kohonen* i *$KY-Kohonen*) zawiera się w określonym przedziale. Należy zauważyć, że tutaj nazwy zmiennych są ujęte w pojedyncze cudzysłowy, ponieważ nazwy te zawierają znaki specjalne.

```
('$KX-Kohonen' >= -0.2635771036148072 i '$KX-Kohonen' <= 0.3146203637123107
i '$KY-Kohonen' >= -0.18975617885589602 i
'$KY-Kohonen' <= 0.17674794197082522) -> T
```

Wiele funkcji, np. funkcje łańcuchowe, wymaga wprowadzenia kilku parametrów i zastosowania prawidłowej składni. W poniższym przykładzie funkcja `subscrs` zwraca pierwszy znak zmiennej *produce_ID*, wskazując, czy pozycja jest ekologiczna, genetycznie zmodyfikowana czy standardowa. Wyniki działania wyrażenia są opisywane za pomocą funkcji `-> `wynik``.

```
subscrs(1,produce_ID) -> `c`
```

Podobnie, następnym wyrażeniem jest:

```
stripchar(`3`,`123`) -> `12`
```

Należy pamiętać, że znaki są zawsze ujmowane w pojedyncze apostrofy odwrotne.

Łączenie funkcji w wyrażeniu

Często wyrażenia CLEM składają się z kombinacji funkcji. Poniższa funkcja stanowi kombinację wyrażeń `subscr` oraz `lowertoupper`, zwraca ona pierwszy znak pola *produce_ID* i przekształca go w wielką literę.

```
lowertoupper(subscr(1,produce_ID)) -> `C`
```

To wyrażenie można również zapisać w skróconej postaci:

```
lowertoupper(produce_ID(1)) -> `C`
```

Inną często używaną kombinacją funkcji jest:

```
locchar_back(`n`, (length(web_page)), web_page)
```

To wyrażenie pozwala na wyszukanie znaku `n` w wartości zmiennej *web_page*. Wartość ta jest odczytywana od tyłu, począwszy od ostatniego znaku wartości zmiennej. Dodanie funkcji `length` powoduje, że wyrażenie dynamicznie oblicza długość bieżącej wartości, a nie korzysta z liczby stałej, np. 7, która będzie nieprawidłowa w przypadku wartości krótszych niż siedem znaków.

Funkcje specjalne

Użytkownik może korzystać z wielu funkcji specjalnych (poprzedzonych znakiem @). Często używane funkcje to na przykład:

```
@BLANK('referrer ID') -> T
```

Funkcje specjalne są często używane w połączeniu z innymi funkcjami. Dzięki temu można oznaczać puste zmienne wsadowo, a nie pojedynczo.

```
@BLANK(@FIELD)-> T
```

Dodatkowe przykłady omówiono w dokumentacji języka CLEM. Więcej informacji można znaleźć w temacie “ CLEM — przegląd informacji” na stronie 161.

Wartości i typy danych

Wyrażenia języka CLEM są podobne do formuł budowanych z wartości, nazw zmiennych, operatorów i funkcji. Najprostszym poprawnym wyrażeniem CLEM jest wartość lub nazwa zmiennej. Oto przykłady poprawnych wartości:

```
3
1.79
'banan'
```

Oto przykłady nazw zmiennych:

```
Product_ID
'$P-NextField'
```

gdzie *Product* jest nazwą zmiennej z zestawu danych dotyczących rynku, a *'$P-NextField'* jest nazwą parametru, zaś wartością wyrażenia jest wartość wskazanej zmiennej. Zwykle nazwy zmiennych rozpoczynają się od litery i mogą także zawierać cyfry i znaki podkreślenia (_). Można używać nazw naruszających te reguły, jednak wówczas muszą one być ujęte w znaki cudzysłowu. W języku CLEM wartości mogą być:

- łańcuchami — na przykład `"c1"`, `"Type 2"`, `"dowolny tekst"`
- liczbami całkowitymi — na przykład `12`, `0`, `–189`
- liczbami rzeczywistymi — na przykład `12.34`, `0.0`, `–0.0045`
- datą/godziną — na przykład `05/12/2002`, `12/05/2002`, `12/05/02`

Możliwe jest stosowanie także następujących elementów:

- kodów znaków — na przykład `` `a` `` lub `3`
- list elementów — na przykład `[1 2 3]`, `['Typ 1' 'Typ 2']`

Kody znaków i listy zwykle nie są stosowane jako wartości zmiennych. Zwykle używane są jako argumenty funkcji CLEM.

Reguły używania cudzysłowów

Mimo że oprogramowanie zapewnia dużą elastyczność w określania zmiennych, wartości, parametrów i łańcuchów używanych w wyrażeniu CLEM, użytkownik powinien zapoznać się z poniższymi, ogólnymi „dobrymi praktykami” dotyczącymi tworzenia wyrażeń:

- **Łańcuchy** — łańcuchy zapisuj zawsze w znakach podwójnego cudzysłowu (`"Type 2"` lub `"value"`). Można stosować pojedyncze cudzysłowy, ale należy się liczyć z możliwością pomyłki ze zmiennymi ujętymi w znaki cudzysłowu.
- **Znaki** — zawsze używaj apostrofów odwrotnych, w taki sposób: `` ` ``. Przykładem może być zapis znaku d w funkcji `stripchar(`d`,"drugA")`. Jedynym wyjątkiem jest odwołanie do konkretnego znaku w łańcuchu za pośrednictwem liczby całkowitej. Tutaj na przykład występuje odwołanie do znaku numer 5: `lowertoupper("druga"(5)) –> "A"`. *Uwaga*: na standardowej klawiaturze brytyjskiej i amerykańskiej (USA) klawisz apostrofu odwrotnego (kod Unicode 0060) znajduje się bezpośrednio pod klawiszem Esc.
- **Zmienne** — w wyrażeniach CLEM zmienne zwykle zapisywane są bez cudzysłowu (`subscr(2,arrayID)) –> CHAR`). W razie potrzeby można używać pojedynczych cudzysłowów, aby uwzględnić spacje lub inne znaki specjalne (`'Order Number'`). Zmienne ujęte w cudzysłowy, ale niezdefiniowane w zestawie danych, zostaną błędnie odczytane jako łańcuchy.
- **Parametry** — zawsze używaj pojedynczych cudzysłowów (`'$P-threshold'`).

Wyrażenia i warunki

Wyrażenia CLEM mogą zwracać wynik (w przypadku wywodzenia nowych wartości), na przykład:

```
Weight * 2.2
Age + 1
sqrt(Signal-Echo)
```

Lub mogą one kończyć się wynikiem *true* lub *false* (w przypadku wyboru warunku), na przykład:

```
Drug = "drugA"
Age < 16
not(PowerFlux) and Power > 2000
```

Istnieje możliwość połączenia operatorów i funkcji w programie CLEM w dowolny sposób, na przykład:

```
sqrt(abs(Signal)) * max(T1, T2) + Baseline
```

Pierwszeństwo nawiasów i operatorów określa kolejność, w jakiej wyrażenie jest interpretowane. W tym przykładzie kolejność interpretowania jest następująca:

- Interpretowane jest `abs(Signal)`, a następnie do wyniku interpretowania stosowane jest `sqrt`.
- Interpretowane jest `max(T1, T2)`.
- Dwa wyniki są następnie przez siebie mnożone: **x** ma pierwszeństwo przed +.
- Na zakończenie następuje dodanie `Baseline` do wyniku.

Malejący porządek pierwszeństwa (to jest od operacji wykonywanych jako pierwsze do operacji wykonywanych jako ostatnie) jest następujący:

- Argumenty funkcji
- Wywołania funkcji
- **xx**
- **x / mod div rem**
- `+ -`
- `> < >= <= /== == = /=`

W przypadku potrzeby pominięcia pierwszeństwa lub w przypadku wątpliwości co do prawidłowej kolejności oceny można użyć nawiasów, które pozwalają w sposób jasny określić kolejność, na przykład,

```
sqrt(abs(Signal)) * (max(T1, T2) + Baseline)
```

Parametry strumienia, sesji i superwęzła

Parametry można definiować w celu wykorzystania ich w wyrażeniach CLEM oraz w skryptach. Stanowią one w efekcie zdefiniowane przez użytkownika zmienne, zapisane i utrwalone w bieżącym strumieniu, sesji lub superwęźle, i można uzyskać do nich dostęp za pośrednictwem interfejsu użytkownika, a także skryptów. W przypadku zapisania strumienia wszelkie parametry ustawione dla tego strumienia również są zapisywane. (To odróżnia je od lokalnych zmiennych skryptu, które mogą być używane tylko w skrypcie, w którym zostały zadeklarowane). Parametry są często używane w skryptach do sterowania zachowaniem skryptów dzięki dostarczaniu informacji na temat zmiennych i wartości niewymagających ustawienia na stałe w skrypcie.

Zasięg parametru zależy od tego, gdzie jest on ustawiany:

- Parametry strumienia można ustawić w skrypcie strumienia lub we właściwościach strumienia, są one również dostępne we wszystkich węzłach w strumieniu. Są one wyświetlane na liście Parametry Konstruktora wyrażeń.

- Parametry sesji można ustawić w autonomicznym skrypcie lub w oknie dialogowym parametrów sesji. Są one dostępne we wszystkich strumieniach używanych w bieżącej sesji (we wszystkich węzłach wymienionych na karcie Strumienie w panelu Zarządzanie).

Parametry można także ustawić dla superwęzłów — w takim przypadku są one widoczne tylko dla węzłów opakowanych w ramach tego superwęzła.

Korzystanie z parametrów w wyrażeniach CLEM

Parametry w wyrażeniach CLEM mają postać `$P-pname`, gdzie `pname` jest nazwą parametru. Korzystając z nich w wyrażeniach CLEM, należy ujmować je w pojedyncze cudzysłowy, na przykład `'$P-scale'`.

Dostępne parametry są łatwo dostępne z poziomu Konstruktora wyrażeń. Aby wyświetlić bieżące parametry:

1. W dowolnym oknie dialogowym akceptującym wyrażenia CLEM kliknij przycisk Konstruktor wyrażeń.
2. Z listy Zmienne wybierz opcję **Parametry**.

Parametry można wybrać z listy, a następnie wstawić do wyrażenia CLEM. Więcej informacji można znaleźć w temacie "Wybór zmiennych, parametrów i zmiennych globalnych" na stronie 157.

Praca z łańcuchami

W przypadku łańcuchów można zastosować wiele operacji, w tym między innymi poniższe operacje.

- Przekształcanie wielkich liter na małe w łańcuchu – `uppertolower(ZNAK)`.
- Usuwanie określonych znaków, np. `` `ID_` `` lub `` `$` ``, ze zmiennej łańcucha – `stripchar(ZNAK,ŁAŃCUCH)`.
- Określanie długości (liczby znaków) łańcucha – `length(ŁAŃCUCH)`.
- Kontrola sortowania alfabetycznego wartości łańcucha – `alphabefore(ŁAŃCUCH1, ŁAŃCUCH2)`.
- Usuwanie początkowych lub końcowych spacji z wartości – `trim(ŁAŃCUCH)`, `trim_start(ŁAŃCUCH)` lub `trimend(ŁAŃCUCH)`.
- Wyodrębnianie pierwszych lub ostatnich *n* znaków łańcucha – `startstring(DŁUGOŚĆ, ŁAŃCUCH)` lub `endstring(DŁUGOŚĆ, ŁAŃCUCH)`. Przykładowo: użytkownik dysponuje zmienną *element* łączącą nazwę produktu z 4-znakowym kodem ID (`ACME CAMERA-D109`). Aby utworzyć nową zmienną zawierającą tylko 4-znakowy kod, poniższą formułę należy wprowadzić w węźle wyliczenia:

```
endstring(4, element)
```

- Dopasowanie określonego wzorca – `ŁAŃCUCH matches WZORZEC`. Przykładowo: aby wybrać osoby, w przypadku których w nazwie stanowiska gdziekolwiek występuje ciąg "market", należy poniższą funkcję zastosować w węźle wyboru:

```
job_title matches "*market*"
```

- Zastępowanie wszystkich wystąpień podłańcucha w łańcuchu – `replace(PODŁAŃCUCH, NOWYPODŁAŃCUCH, ŁAŃCUCH)`. Przykładowo: aby zastąpić wszystkie wystąpienia nieobsługiwanych znaków, np. pionowej kreski (`|`), znakiem średnika przed rozpoczęciem eksploracji danych, należy użyć funkcji `replace` w węźle wypełniania. W obszarze **Wypełnij zmienne:** należy wybrać wszystkie zmienne, w których mogą występować dane znaki. W przypadku warunku **Zamień:** należy wybrać opcję **Zawsze** i określić następujący warunek w obszarze **Zamień na:**

```
replace('|',';',@FIELD)
```

- Wyliczanie zmiennej typu flaga w oparciu o obecność określonych podłańcuchów. Przykładowo: za pomocą funkcji łańcuchowej w węźle wyliczenia można utworzyć oddzielną zmienną typu flaga dla każdej odpowiedzi, korzystając z wyrażenia:

```
hassubstring(museums,"museum_of_design")
```

Więcej informacji można znaleźć w temacie “Funkcje łańcuchowe” na stronie 179.

Obsługa wartości pustych i braków danych

Zastępowanie braków danych lub pustych wartości to często wykonywane zdanie przygotowawcze w ramach eksploracji danych. W języku CLEM dostępnych jest wiele narzędzi automatyzujących obsługę pustych wartości. Działania związane z wartościami pustymi są najczęściej wykonywane w węźle wypełniania. Jednak poniższych funkcji można użyć w dowolnym węźle zapewniającym obsługę wyrażeń CLEM:

- `@BLANK(FIELD)` - tej funkcji można użyć do określenia rekordów, które zawierają puste wartości w określonej zmiennej, np. *Wiek*.
- `@NULL(FIELD)` - tej funkcji można użyć do określenia rekordów, które zawierają systemowe braki danych w określonych zmiennych. W języku IBM SPSS Modeler systemowe braki danych są wyświetlane jako wartości **$null$**.

Więcej informacji można znaleźć w temacie “Funkcje obsługujące wartości puste i null” na stronie 198.

Praca z liczbami

W języku IBM SPSS Modeler dostępnych jest wiele standardowych operacji wykonywanych na wartościach numerycznych, np.:

- Obliczanie sinusa określonego kąta – `sin(LICZ)`
- Obliczanie logarytmu naturalnego zmiennej numerycznej – `log(LICZ)`
- Obliczanie sumy dwóch liczb `LICZ1 + LICZ2`

Więcej informacji można znaleźć w temacie “Funkcje numeryczne” na stronie 174.

Praca z godzinami i datami

Formaty daty i czasu mogą się różnić w zależności od źródła danych i ustawień regionalnych. Formaty te są określone dla danego strumienia i można je skonfigurować w oknie dialogowym właściwości strumienia. Poniżej przedstawiono często używane funkcje pozwalające na pracę ze zmiennymi data/czas.

Obliczanie minionego czasu

Czas, który upłynął od daty bazowej, można obliczyć w prosty sposób, korzystając z funkcji podobnych do funkcji zaprezentowanej poniżej. Ta funkcja zwraca czas w miesiącach od daty bazowej do daty reprezentowanej przez łańcuch DATA w postaci liczby rzeczywistej. Jest to wartość określona w oparciu o długość miesiąca wynoszącą 30,0 dni.

```
date_in_months(Data)
```

Porównywanie wartości data/czas

Wartości zmiennych data/czas można porównywać w wielu rekordach, korzystając z funkcji podobnych do poniższej. Ta funkcja zwraca wartość *true*, jeśli łańcuch daty DATA1 reprezentuje datę przypadającą przed datą reprezentowaną przez łańcuch daty DATA2. W przeciwnym razie ta funkcja zwraca wartość 0.

```
date_before(Data1, Data2)
```

Obliczanie różnic

Użytkownik może również obliczyć różnicę między dwoma punktami w czasie i dwiema datami, korzystając z funkcji takich jak:

```
date_weeks_difference(Data1, Data2)
```

Ta funkcja zwraca czas w tygodniach, który upłynął od daty reprezentowanej przez łańcuch daty DATA1 do daty reprezentowanej przez łańcuch daty DATA2 w postaci liczby rzeczywistej. Jest to wartość określona w oparciu o długość tygodnia wynoszącą 7,0 dni. Jeśli wartość DATA2 przypada przed wartością DATA1, ta funkcja zwraca liczbę ujemną.

Dzisiejsza data

Aktualną datę można dodać do zbioru danych za pomocą funkcji @TODAY. Ta data jest dodawana jako łańcuch do określonej zmiennej lub jako nowa zmienna z zastosowaniem formatu wybranego w oknie dialogowym właściwości strumienia. Więcej informacji można znaleźć w temacie “Funkcje daty i czasu” na stronie 186.

Podsumowanie wielu zmiennych

Język CLEM jest wyposażony w wiele funkcji zwracających statystyki podsumowujące wiele zmiennych. Funkcje te mogą być szczególnie użyteczne podczas analizowania danych z ankiet, w przypadku których wiele odpowiedzi na pytanie można przechowywać w wielu zmiennych. Więcej informacji można znaleźć w temacie “Praca z danymi opisującymi wielokrotne odpowiedzi” na stronie 152.

Funkcje porównawcze

Użytkownik może porównywać wartości w wielu zmiennych, korzystając z funkcji min_n i max_n, np.:

```
max_n(['oplatazakarte1' 'oplatazakarte2''oplatazakarte3''oplatazakarte4'])
```

Można również korzystać z wielu funkcji zliczających w celu uzyskania liczby wartości spełniających określone kryteria, nawet jeśli wartości te są przechowywane w wielu zmiennych. Przykładowo: aby określić liczbę kart, którymi klienci dysponowali przez ponad pięć lat:

```
count_greater_than(5, ['okreskarty' 'okreskarty2' 'okreskarty3'])
```

Aby określić liczbę wartości null w ramach tego samego zbioru zmiennych:

```
count_nulls(['okreskarty' 'okreskarty2' 'okreskarty3'])
```

W tym przykładzie jest określana liczba kart, a nie liczba ich właścicieli. Więcej informacji można znaleźć w temacie “Funkcje porównawcze” na stronie 171.

Liczbę wystąpień danej wartości w wielu zmiennych można określić za pomocą funkcji count_equal. W poniższym przykładzie określana jest liczba zmiennych na liście zawierających wartość Y.

```
count_equal("Y",[Odpowiedź1, Odpowiedź2, Odpowiedź3])
```

Po wprowadzeniu tych wartości w zmiennych na liście funkcja zwraca liczebność wartości Y, tak jak przedstawiono to poniżej.

Tabela 8. Wartości funkcji

Odpowiedź1	Odpowiedź2	Odpowiedź3	Liczebność
Y	N	Y	2
Y	N	N	1

Funkcje numeryczne

Użytkownik może uzyskać statystyki w wielu zmiennych, korzystając z funkcji sum_n, mean_n i sdev_n, np.:

```
sum_n(['saldokarty1' 'saldokarty2''saldokarty3'])
```

```
mean_n(['saldokarty1' 'saldokarty2''saldokarty3'])
```

Więcej informacji można znaleźć w temacie “Funkcje numeryczne” na stronie 174.

Tworzenie listy zmiennych

W przypadku funkcji akceptujących listę zmiennych można użyć funkcji specjalnych `@FIELDS_BETWEEN(początek, koniec)` i `@FIELDS_MATCHING(wzorzec)` w charakterze danych wejściowych. Przykładowo: zakładając, że kolejność zmiennych jest zgodna z pokazaną w przykładzie `sum_n`, poniższa operacja jest równoważna operacji z tamtego przykładu:

```
sum_n(@FIELDS_BETWEEN(saldokarty1, saldokarty3))
```

Aby określić liczbę wartość null we wszystkich zmiennych rozpoczynających się znakami "*karta*":

```
count_nulls(@FIELDS_MATCHING('karta*'))
```

Więcej informacji można znaleźć w temacie “Funkcje specjalne” na stronie 199.

Praca z danymi opisującymi wielokrotne odpowiedzi

Dane zapewniające wielokrotne odpowiedzi można analizować z użyciem wielu funkcji porównawczych, np.:

- `value_at`
- `first_index / last_index`
- `first_non_null / last_non_null`
- `first_non_null_index / last_non_null_index`
- `min_index / max_index`

Przykładowo: przyjmijmy, że w ankiecie zapytano o pierwszą, drugą i trzecią co do ważności przyczynę dokonania konkretnego zakupu (np. cena, rekomendacja osobista, recenzja, lokalny dostawca, inne). W takiej sytuacji można określić istotność ceny, wyznaczając indeks zmiennej, w której cena została uwzględniona pierwszy raz:

```
first_index("cena", [Przyczyna1 Przyczyna2 Przyczyna3])
```

Podobnie: załóżmy, że klientów poproszono o stworzenie rankingu trzech samochodów uszeregowanych zgodnie z prawdopodobieństwem zakupu, a odpowiedzi zakodowano w oddzielnych zmiennych w następujący sposób:

Tabela 9. Przykładowy ranking samochodów

id klienta	**samochód1**	**samochód2**	**samochód3**
101	1	3	2
102	3	2	1
103	2	3	1

W takiej sytuacji można określić indeks zmiennej dla najbardziej lubianego samochodu (miejsce 1. lub najniższy poziom w rankingu), korzystając z funkcji `min_index`:

```
min_index(['samochód1' 'samochód2' 'samochód3'])
```

Więcej informacji można znaleźć w temacie “Funkcje porównawcze” na stronie 171.

Zestaw wielokrotnych odpowiedzi – odniesienia

Za pomocą funkcji specjalnej `@MULTI_RESPONSE_SET` można utworzyć odniesienia do wszystkich zmiennych w zestawie wielokrotnych odpowiedzi. Przykładowo: jeśli trzy zmienne *samochód*

z poprzedniego przykładu zostaną zawarte w zestawie wielokrotnych odpowiedzi o nazwie *rankingi_samochodów*, to użycie poniższej funkcji zapewni uzyskanie takich samych wyników:

```
max_index(@MULTI_RESPONSE_SET("rankingi_samochodów"))
```

Konstruktor wyrażeń

Wyrażenia CLEM można wpisywać ręcznie lub tworzyć za pośrednictwem Konstruktora wyrażeń, który pozwala wyświetlać pełną listę funkcji CLEM i operatorów, a także zmienne zawierające dane z bieżącego strumienia, co umożliwia szybkie budowanie wyrażeń bez potrzeby zapamiętywania dokładnych nazw zmiennych czy funkcji. Ponadto elementy sterujące Konstruktora umożliwiają ujmowanie zmiennych i wartości we własne znaki cudzysłowu, ułatwiając tym samym tworzenie wyrażeń z prawidłową składnią.

Uwaga: Konstruktor wyrażeń nie jest obsługiwany w przypadku ustawień parametrów lub skryptów.

Uwaga: W celu zmiany źródła danych należy uprzednio sprawdzić, czy Konstruktor wyrażeń nadal obsługuje wybrane przez użytkownika funkcje. Ponieważ nie wszystkie bazy danych obsługują wszystkie funkcje, w przypadku uruchomienia takiej funkcji na nowej bazie danych może wystąpić błąd.

Uzyskiwanie dostępu do Konstruktora wyrażeń

Konstruktor wyrażeń jest dostępny we wszystkich węzłach, w których używane są wyrażenia CLEM, w tym w węźle selekcji, ważenia, wyliczania, filtrowania, analizy, raportowania i tabeli. Konstruktora wyrażeń można otworzyć, klikając przycisk kalkulatora po prawej stronie pola formularza.

Tworzenie wyrażeń

Konstruktor wyrażeń oferuje nie tylko kompletną listę zmiennych, funkcji i operatorów, lecz także dostęp do wartości danych w przypadku, gdy są one określone.

Aby utworzyć wyrażenie, korzystając z Konstruktora wyrażeń

1. Wpisz zmienną wyrażenia, korzystając z list funkcji i zmiennych.

   lub

2. Wybierz wymagane zmienne i funkcje z list przewijania.
3. Kliknij dwukrotnie lub jednokrotnie żółty przycisk strzałki, aby dodać zmienną lub funkcję do wyrażenia.
4. Użyj przycisków operandów w środkowej części okna dialogowego w celu wstawienia operacji do wyrażenia.

Wybieranie funkcji

Na liście funkcji wyświetlane są wszystkie dostępne funkcje i operatory CLEM. Przewiń, aby wybrać funkcję z listy, lub, w celu ułatwienia wyszukiwania, użyj listy rozwijanej podzbioru funkcji lub operatorów. Dostępne funkcje są pogrupowane w kategorie celem ułatwienia wyszukiwania.

Większość z nich opisano w sekcji referencyjnej opisu języka CLEM. Aby uzyskać więcej informacji, zobacz “Funkcje — informacje” na stronie 167.

Pozostałe kategorie wymieniono poniżej.

- **Funkcje ogólne** to zbiór najczęściej używanych funkcji.
- **Ostatnio używane** to lista funkcji CLEM używanych w bieżącej sesji.
- **Funkcje @** to lista wszystkich funkcji specjalnych, których nazwy są poprzedzone symbolem „@”.

  Uwaga: Funkcje `@DIFF1(ZMIENNA1,ZMIENNA2)` i `@DIFF2(ZMIENNA1,ZMIENNA2)` wymagają, aby obie zmienne były tego samego typu (np. obie typu Liczba całkowita, Liczba typu long lub Liczba rzeczywista).

- **Funkcje bazodanowe**. Jeśli strumień zawiera połączenie z bazą danych (za pośrednictwem węzła Źródło bazy danych), ten wybór powoduje wygenerowanie listy funkcji dostępnych z tej bazy danych, w tym funkcji zdefiniowanych przez użytkownika (UDF). Aby uzyskać więcej informacji, zobacz "Funkcje bazodanowe" na stronie 154.
- **Agregaty bazy danych**. Jeśli strumień obejmuje połączenie z bazą danych (za pośrednictwem węzła Źródło bazy danych), ten wybór powoduje wygenerowanie listy opcji agregacji dostępnych z poziomu tej bazy danych. Opcje te są dostępne w Konstruktorze wyrażeń węzła Agregacja.
- **Agregaty okna bazy danych**. Jeśli strumień obejmuje połączenie z bazą danych (za pośrednictwem węzła Źródło bazy danych), ten wybór powoduje wygenerowanie listy opcji agregacji dla okna, których można użyć w tej bazie danych. Opcje te są dostępne w Konstruktorze wyrażeń wyłącznie w węzłach należących do palety **Operacje na zmiennych**.

  Uwaga: Ponieważ program SPSS Modeler uzyskuje **Funkcje agregujące okna** z widoku systemowej bazy danych, dostępność opcji zależy od zachowania bazy danych.

  Mimo określenia „agregujące" opcje te nie są przeznaczone do użycia w węźle Agregacja; są one znacznie bardziej przydatne w przypadku takich węzłów, jak Wyliczanie lub Selekcja. Dzieje się tak, ponieważ ich wyniki są wartościami skalarnymi, nie zaś „prawdziwymi" agregacjami; oznacza to, że nie powodują one zmniejszenia ilości danych w wyniku w ten sam sposób, w jaki odbywa się to w przypadku węzła Agregacja. Na przykład można użyć tego rodzaju agregacji w celu udostępnienia średniej ruchomej, takich jak „średnia dla bieżącego wiersza oraz dla wszystkich poprzednich wierszy".
- **Agregaty wbudowane**. Jest to lista wszystkich możliwych trybów agregacji, jakie mogą być używane.
- **Operatory** to lista wszystkich operatorów, jakich można użyć podczas konstruowania wyrażeń. Operatory są także dostępne za pośrednictwem przycisków w środkowej części okna dialogowego.
- **Wszystkie funkcje** to pełna lista dostępnych funkcji CLEM.

Po wybraniu grupy funkcji należy kliknąć dwukrotnie, aby wstawić funkcję do pola wyrażenia w punkcie wskazanym przez położenie kursora.

Funkcje bazodanowe

Funkcje bazodanowe mogą być wymienione na listach dla wielu miejscach; w poniższej tabeli przedstawiono miejsca przeszukiwane przez program SPSS Modeler. Ta tabela może być używana przez administratorów baz danych w celu zapewnienia użytkownikom uprawnień dostępu do wymaganych obszarów, tak aby mogli oni korzystać z różnych funkcji.

Ponadto w tabeli wymieniono warunki decydujące o dostępności funkcji w oparciu o typ bazy danych i funkcji.

Uwaga: W przypadku korzystania z funkcji bazodanowych Amazon Redshift może być konieczne udzielenie użytkownikowi przez administratora bazy danych uprawnień do następujących sześciu obiektów bazy danych. Pierwsze cztery to tabele katalogu systemowego, a pozostałe dwa to schematy.

- pg_type
- pg_proc
- pg_namespace
- pg_aggregate
- information_schema
- pg_catalog

Tabela 10. Funkcje bazodanowe w Konstruktorze wyrażeń

Baza danych	Typ funkcji	Gdzie znaleźć funkcję	Warunki służące do filtrowania funkcji
Db2 LUW	UDF	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE ma wartość F oraz FUNCTIONTYPE ma wartość S

Tabela 10. Funkcje bazodanowe w Konstruktorze wyrażeń (kontynuacja)

Baza danych	Typ funkcji	Gdzie znaleźć funkcję	Warunki służące do filtrowania funkcji
Db2 LUW	UDA	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE ma wartość F oraz FUNCTIONTYPE ma wartość C
Db2 iSeries	UDF	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE ma wartość F oraz FUNCTION_TYPE ma wartość S
Db2 iSeries	UDA	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE ma wartość F oraz FUNCTION_TYPE ma wartość C
Db2 z/OS	UDF	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE ma wartość F oraz FUNCTIONTYPE ma wartość S
Db2 z/OS	UDA	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE ma wartość F oraz FUNCTIONTYPE ma wartość C
SQL Server	UDF	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE ma wartość FN lub FS
SQL Server	UDA	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE ma wartość AF
Oracle	UDF	ALL_ARGUMENTS ALL_PROCEDURES	Spełnione są wszystkie poniższe warunki: • OBJECT_TYPE ma wartość FUNCTION • AGGREGATE ma wartość NO • PLS_TYPE jest różne od NULL
Oracle	UDA	ALL_ARGUMENTS ALL_PROCEDURES	Spełnione są wszystkie poniższe warunki: • ARGUMENT_NAME ma wartość NULL • AGGREGATE ma wartość YES • PLS_TYPE jest różne od NULL
Teradata	UDF	DBC.FUNCTIONS DBC.ALLRIGHTS	Spełnione są wszystkie poniższe warunki: • FUNCTIONTYPE ma wartość F • COLUMNNAME ma wartość RETURN0 • SPPARAMETERTYPE ma wartość 0 • ACCESSRIGHT ma wartość EF
Teradata	UDA	DBC.FUNCTIONS DBC.ALLRIGHTS	Spełnione są wszystkie poniższe warunki: • FUNCTIONTYPE ma wartość A • COLUMNNAME ma wartość RETURN0 • SPPARAMETERTYPE ma wartość 0 • ACCESSRIGHT ma wartość EF

Tabela 10. Funkcje bazodanowe w Konstruktorze wyrażeń (kontynuacja)

Baza danych	Typ funkcji	Gdzie znaleźć funkcję	Warunki służące do filtrowania funkcji
Netezza	UDF	####..._V_FUNCTION NZA.._V_FUNCTION INZA.._V_FUNCTION	W przypadku ####.._V_FUNCTION mają zastosowanie następujące warunki: • RESULT nie zawiera łańcucha z takimi wartościami, jak: TABLE% • FUNCTION nie zawiera łańcucha z takimi wartościami, jak: '/_%' escape '/' • VARARGS ma wartość FALSE W przypadku zarówno NZA.._V_FUNCTION, jak i INZA.._V_FUNCTION, mają zastosowanie następujące warunki: • RESULT nie zawiera łańcucha z takimi wartościami, jak: TABLE% • FUNCTION nie zawiera łańcucha z takimi wartościami, jak: '/_%' escape '/' • BUILTIN ma wartość f • VARARGS ma wartość FALSE
Netezza	UDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	Spełnione są oba poniższe warunki: • AGGTYPE ma wartość ANY lub GROUPED • VARARGS ma wartość FALSE
Netezza	WUDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	W przypadku ####.._V_AGGREGATE mają zastosowanie następujące warunki: • AGGTYPE ma wartość ANY lub ANALYTIC • AGGREGATE jest różne od MAX_LABEL • VARARGS ma wartość FALSE W przypadku zarówno NZA.._V_FUNCTION, jak i INZA.._V_FUNCTION, mają zastosowanie następujące warunki: • AGGTYPE ma wartość ANY lub ANALYTIC • BUILTIN ma wartość f • VARARGS ma wartość FALSE

Klucz do terminów używanych w tabeli

- UDF Funkcja zdefiniowana przez użytkownika
- UDA Funkcja agregująca zdefiniowana przez użytkownika
- WUDA Funkcja agregująca okna zdefiniowana przez użytkownika
- #### Baza danych, z którą użytkownik ma obecnie połączenie.

Wybór zmiennych, parametrów i zmiennych globalnych

Lista zmiennych zawiera wszystkie zmienne dostępne w danym miejscu w strumieniu danych. Przewijając, można wybrać zmienną z listy. Kliknij dwukrotnie lub jednokrotnie żółty przycisk strzałki, aby dodać zmienną do wyrażenia.

Więcej informacji można znaleźć w temacie "Parametry strumienia, sesji i superwęzła" na stronie 148.

Obok zmiennych można także wybierać następujące elementy:

Zestawy wielokrotnych odpowiedzi. Więcej informacji zawiera podręcznik *IBM SPSS Modeler — węzły źródłowe, procesowe i wyników*.

Ostatnio używane to lista zmiennych, zestawów wielokrotnych odpowiedzi, parametrów i wartości globalnych używanych w bieżącej sesji.

Parametry. Więcej informacji można znaleźć w temacie "Parametry strumienia, sesji i superwęzła" na stronie 148.

Wartości globalne. Więcej informacji zawiera podręcznik *IBM SPSS Modeler — węzły źródłowe, procesowe i wyników*.

Wyświetlanie lub wybieranie wartości

Wartości zmiennych można wyświetlać z wielu miejsc w systemie, w tym z Konstruktora wyrażeń, z raportów z kontroli danych oraz podczas edytowania przyszłych wartości w węźle Przedziały czasowe. Należy zwrócić uwagę, że aby możliwe było użycie danych, muszą one zostać w pełni określone w źródle lub w węźle Typ; konieczne jest określenie miejsc składowania, typów i wartości.

W celu wyświetlenia wartości dla zmiennej z Kreatora wyrażeń z węzła Przedziały czasu, wybierz żądaną zmienną i kliknij Przycisk wyboru wartości, aby otworzyć okno dialogowe z listą wartości dla wybranej zmiennej. Można następnie wybrać wartość i kliknąć opcję **Wstaw** w celu wklejenia wartości do bieżącego wyrażenia lub listy.

Rysunek 16. Przycisk wyboru wartości

W przypadku zmiennych flagi i zmiennych nominalnych wymienione są wszystkie zdefiniowane wartości. Dla zmiennych ciągłych (zakres liczb) wyświetlane są wartości minimalne i maksymalne.

Sprawdzanie wyrażeń CLEM

Kliknij opcję **Sprawdź** w Konstruktorze wyrażeń (prawy dolny róg) w celu sprawdzenia poprawności wyrażenia. Wyrażenia niesprawdzone są wyświetlane w kolorze czerwonym. W przypadku znalezienia błędów wyświetlany jest komunikat wskazujący przyczynę.

Sprawdzane są następujące elementy:

- Poprawność cudzysłowów otaczających wartości i nazwy zmiennych
- Prawidłowość użycia parametrów i zmiennych globalnych
- Poprawność użycia operatorów
- Obecność zmiennych, do których odwołuje się wyrażenie
- Obecność i definicja zmiennych globalnych, do których odwołuje się wyrażenie

W przypadku napotkania błędów w składni należy spróbować utworzyć wyrażenie, korzystając z przycisków list i operatorów, bez wprowadzania wyrażeń ręcznie. Metoda ta umożliwia automatyczne dodawanie właściwych znaków cudzysłowu dla zmiennych i wartości.

Należy zwrócić uwagę na następujące ograniczenia przy tworzeniu wyrażeń w programie IBM Analytical Decision Management. Wyrażenia nie mogą zawierać żadnego z poniższych elementów:

- Odwołanie do parametru strumienia IBM SPSS Modeler
- Odwołanie do parametru globalnego strumienia IBM SPSS Modeler
- Odwołanie do funkcji bazy danych
- Odwołanie do następujących funkcji specjalnych lub funkcji @:
  - @TARGET
  - @PREDICTED
  - @FIELD
  - @PARTITION_FIELD
  - @TRAINING_PARTITION
  - @TESTING_PARTITION
  - @VALIDATION_PARTITION

Uwaga: Nazwy zmiennych zawierające separatory muszą być ujęte w apostrofy. Aby apostrofy były automatycznie dodawane, można tworzyć wyrażenia, korzystając z list i przycisków operatorów, zamiast wpisywać je ręcznie. Obecność następujących znaków w nazwach zmiennych może powodować błędy:
• ! "# $% & '() = ~ |-^ ¥ @" "+ *" "<>? . ,/ :; → (strzałka), □ ▵ (symbole graficzne itp.)

Znajdowanie i zastępowanie

Okno dialogowe Znajdź i zamień jest dostępne tylko w miejscach, w których można edytować tekst skryptu lub wyrażenia, takich jak edytor skryptów czy Konstruktor wyrażeń CLEM, lub podczas definiowania szablonu w węźle Raport. Podczas edytowania tekstu w dowolnym z tych obszarów naciśnij kombinację klawiszy Ctrl+F, aby uzyskać dostęp do okna dialogowego, upewniając się wcześniej, że kursor znajduje się w obszarze tekstowym i ma fokus. W przypadku pracy w węźle wypełniania można uzyskać dostęp do okna dialogowego z dowolnego z obszarów tekstowych na karcie Ustawienia lub ze zmiennej tekstowej w Konstruktorze wyrażeń.

1. Po umieszczeniu kursora w obszarze tekstowym naciśnij kombinację klawiszy Ctrl+F, aby uzyskać dostęp do okna dialogowego Znajdź i zamień.
2. Wprowadź tekst, który chcesz wyszukać, lub wybierz z listy rozwijanej ostatnio wyszukiwany element.
3. Wprowadź nowy tekst odpowiednio do potrzeb.
4. Kliknij opcję **Znajdź następne**, aby rozpocząć wyszukiwanie.
5. Kliknij opcję **Zamień**, aby zastąpić bieżący wybór, lub opcję **Zamień wszystko**, aby zmienić wszystkie lub tylko wybrane wystąpienia.
6. Okno dialogowe zamyka się po każdej operacji. Naciśnij klawisz F3 z poziomu dowolnego obszaru tekstowego, aby powtórzyć ostatnią operację wyszukiwania, lub naciśnij kombinację klawiszy Ctrl+F, aby ponownie uzyskać dostęp do tego okna dialogowego.

Opcje wyszukiwania

Uwzględnij wielkość liter. Określa, czy operacja wyszukiwania rozróżnia wielkość liter; na przykład czy *myvar* stanowi dopasowanie dla *myVar*. Tekst zastępujący jest wstawiany zawsze dokładnie tak, jak został wprowadzony, niezależnie od tego ustawienia.

Tylko całe słowa. Określa, czy operacja wyszukiwania dopasowuje tylko tekst odzwierciedlający całe słowa. W przypadku zaznaczenia tej opcji wyszukiwanie tekstu *spider* nie zwróci takich wyników, jak *spiderman* czy *spider-man*.

Wyrażenia regularne. Określa, czy używana jest składnia wyrażenia regularnego (patrz następna sekcja). Zaznaczenie tej opcji powoduje, że opcja **Tylko całe słowa** jest nieaktywna, a jej wartość jest ignorowana.

Tylko wybrany tekst. Wpływa na zasięg wyszukiwania w przypadku korzystania z opcji **Zamień wszystko**.

Składnia wyrażenia regularnego

Wyrażenia regularne umożliwiają wyszukiwanie znaków specjalnych, takich jak znaki tabulacji czy znaki podziału wiersza, klasy lub zakresy znaków, np. od *a* do *d*, dowolne cyfry lub znaki alfanumeryczne, oraz granice, takie jak początek czy koniec wiersza. Obsługiwane są następujące typy wyrażeń.

Tabela 11. Dopasowania znaków

Znaki	**Dopasowania**
x	Znak x
\\	Ukośnik lewy
\0n	Znak o wartości ósemkowej 0n (0 <= n <= 7)
\0nn	Znak o wartości ósemkowej 0nn (0 <= n <= 7)
\0mnn	Znak o wartości ósemkowej 0mnn (0 <= m <= 3, 0 <= n <= 7)
\xhh	Znak o wartości szesnastkowej 0xhh
\uhhhh	Znak o wartości szesnastkowej 0xhhhh
\t	Znak tabulacji ('\u0009')
\n	Znak nowego wiersza (podziału wiersza) ('\u000A')
\r	Znak powrotu karetki ('\u000D')
\f	Znak nowej strony ('\u000C')
\a	Znak alertu (sygnału dźwiękowego) ('\u0007')
\e	Znak zmiany znaczenia ('\u001B')
\cx	Znak sterujący odpowiadający x

Tabela 12. Dopasowywanie klas znaków

Klasy znaków	**Dopasowania**
[abc]	a, b lub c (klasa prosta)
[^abc]	Dowolny znak z wyjątkiem a, b lub c (odejmowanie)
[a-zA-Z]	Od a do z lub od A do Z włącznie (przedział)
[a-d[m-p]]	Od a do d lub od m do p (suma przedziałów). Można też przedstawić ten zapis w postaci [a-dm-p]
[a-z&&[def]]	Od a do z i d, e lub f (część wspólna)
[a-z&&[^bc]]	Od a do z, z wyjątkiem b i c (odejmowanie). Można też przedstawić ten zapis w postaci [ad-z]
[a-z&&[^m-p]]	Od a do z i nie m do p (odejmowanie). Można też przedstawić ten zapis w postaci [a-lq-z]

Tabela 13. Predefiniowane klasy znaków

Predefiniowane klasy znaków	**Dopasowania**
.	Dowolny znak (może odpowiadać lub może nie odpowiadać znakom końca wiersza)
\d	Dowolna cyfra: [0-9]
\D	Dowolny znak inny niż cyfra: [^0-9]

Tabela 13. Predefiniowane klasy znaków (kontynuacja)	
Predefiniowane klasy znaków	**Dopasowania**
\s	Biały znak: [\t\n\x0B\f\r]
\S	Znak inny niż biały znak: [^\s]
\w	Znak należący do słowa: [a-zA-Z_0-9]
\W	Znak nienależący do słowa: [^\w]

Tabela 14. Dopasowania granic	
Dopasowania granic	**Dopasowania**
^	Początek wiersza
$	Koniec wiersza
\b	Granica słowa
\B	Granica inna niż słowa
\A	Początek danych wejściowych
\Z	Koniec danych wejściowych, z wyjątkiem finalnego znaku kończącego, o ile występuje
\z	Koniec danych wejściowych

Rozdział 10. CLEM — informacje o języku

CLEM — przegląd informacji

W tej sekcji opisano język CLEM (Control Language for Expression Manipulation). Jest to potężny mechanizm służący do analizowania danych stosowanych w strumieniach IBM SPSS Modeler oraz do manipulowania tymi danymi. Języka CLEM można używać w węzłach i w ten sposób wykonywać operacje polegające na ocenie wyrażeń, obliczaniu wartości lub wstawianiu danych w raportach.

Wyrażenia CLEM składają się z wartości, nazw zmiennych, operatorów i funkcji. Dzięki zastosowaniu prawidłowej składni można utworzyć wiele niezwykle wydajnych operacji umożliwiających obsługę danych.

Typy danych CLEM

Typy danych CLEM mogą składać się z:

- Liczb całkowitych
- Liczb rzeczywistych
- Znaków
- Łańcuchów
- List
- Zmiennych
- Daty/godziny

Zasady używania cudzysłowów

Oprogramowanie IBM SPSS Modeler zapewnia dużą elastyczność w określaniu zmiennych, wartości, parametrów i łańcuchów używanych w wyrażeniu CLEM, użytkownik powinien zapoznać się z poniższymi, ogólnymi „dobrymi praktykami” dotyczącymi tworzenia wyrażeń:

- Łańcuchy: zawsze używaj podwójnych znaków cudzysłowu w łańcuchach, np. `"Typ 2"`. Można stosować pojedyncze cudzysłowy, ale należy się liczyć z możliwością pomyłki ze zmiennymi ujętymi w znaki cudzysłowu.
- Zmienne: pojedynczych cudzysłowów należy używać wyłącznie, gdy jest to wymagane w celu ujęcia w cudzysłów spacji lub innych znaków specjalnych, np. takich jak `'Numer zamówienia'`. Zmienne ujęte w cudzysłowy, ale niezdefiniowane w zestawie danych, zostaną błędnie odczytane jako łańcuchy.
- Parametry: zawsze używaj pojedynczych znaków cudzysłowu w parametrach, np. `'$P-threshold'`.
- Znaki: zawsze używaj pojedynczych apostrofów odwrotnych (`), np. `stripchar(`d`, "drugA")`.

Dokładne opisy niniejszych zasad zawarto w kolejnych tematach.

Liczby całkowite

Liczby całkowite są reprezentowane jako sekwencja cyfr dziesiętnych. Opcjonalnie można umieścić symbol minus (–) przed liczbą całkowitą, aby określić liczbę jako ujemną — na przykład, 1234, 999, –77.

Język CLEM obsługuje liczby całkowite o arbitralnie przyjętej precyzji. Maksymalna wielkość liczby całkowitej zależy od używanej platformy. Jeśli wartości są zbyt duże, aby mogły zostać wyświetlone w polu liczby całkowitej, zmiana typu zmiennej na typ `Real` zwykle powoduje przywrócenie wartości.

Liczby rzeczywiste

Termin *Liczba rzeczywista* oznacza liczbę zmiennopozycyjną. Liczby rzeczywiste są przedstawiane w postaci jednej lub kilku cyfr, kropki dziesiętnej oraz jednej lub kilku cyfr po kropce. Liczby rzeczywiste w języku CLEM są liczbami podwójnej precyzji.

Opcjonalnie można umieścić symbol minus (–) przed liczbą rzeczywistą, aby określić liczbę jako ujemną — na przykład: `1.234`, `0.999`, `–77.001`. Aby przedstawić liczbę rzeczywistą w zapisie wykładniczym, użyj postaci <*liczba*> e <*wykładnik*> — na przykład `1234.0e5`, `1.7e–2`. Podczas odczytu przez aplikację IBM SPSS Modeler łańcuchów liczbowych z plików i automatycznym przekształceniu ich na liczby akceptowane są liczby bez wiodącej cyfry przed znakiem dziesiętnym lub bez cyfry po znaku dziesiętnym, na przykład `999.` czy `,11`. W wyrażeniach CLEM takie postacie są jednak niedozwolone.

Uwaga: Podczas tworzenia odwołań do liczb rzeczywistych w wyrażeniach CLEM konieczne jest użycie kropki jako separatora dziesiętnego, niezależnie od ustawień dla bieżącego strumienia czy ustawień regionalnych. Należy wpisać na przykład

```
Na > 0.6
```

nie zaś

```
Na > 0,6
```

Dotyczy to również sytuacji, w których przecinek wybrano jako symbol dziesiętny w oknie dialogowym właściwości strumienia, i jest spójne z ogólnymi wytycznymi mówiącymi, że składnia kodu powinna być niezależna od wszelkich ustawień narodowych czy obowiązujących konwencji.

Znaki

Znaki (zazwyczaj oznaczane jako CHAR) są najczęściej używane w wyrażeniach CLEM do wykonywania testów na łańcuchach. Przykładowo: za pomocą funkcji `isuppercode` można określić, czy pierwszy znak łańcucha jest wielką literą. Poniższe wyrażenie CLEM wykonuje test na pierwszym znaku łańcucha:

```
isuppercode(subscrs(1, "MójŁańcuch"))
```

Aby w wyrażeniu CLEM zawrzeć kod (w odróżnieniu od lokalizacji) danego znaku, należy zastosować pojedyncze apostrofy odwrotne `` `<*znak*>` `` — na przykład: `` `A` ``, `` `Z` ``.

Uwaga: nie istnieje typ składowania zmiennych CHAR, więc jeśli zmienna jest wyliczana lub wypełniana z zastosowaniem wyrażenia, którego wynikiem działania jest CHAR, to ten wynik zostanie przekształcony do postaci łańcucha.

Łańcuchy

Ogólnie: łańcuchy należy umieszczać w podwójnych cudzysłowach. Przykładowe łańcuchy: `"c35product2"` i `"referrerID"`. Aby uwzględnić znaki specjalne w łańcuchu, należy zastosować znak ukośnika odwrotnego, np.: `"\$65443"`. (Aby użyć znaku ukośnika odwrotnego, należy go poprzedzić drugim ukośnikiem odwrotnym: `\\`). Łańcuch można ująć w pojedyncze cudzysłowy, ale wyniku nie będzie można odróżnić od zmiennej ujętej w cudzysłów (`'referrerID'`). Więcej informacji można znaleźć w temacie “Funkcje łańcuchowe” na stronie 179.

Listy

Lista jest uporządkowaną sekwencją elementów, przy czym elementy te mogą należeć do różnych typów. Listy ujmuje się w nawiasy kwadratowe ([]). Oto przykłady list: `[1 2 4 16]` i `["abc" "def"]`. Listy nie są używane jako wartości zmiennych IBM SPSS Modeler. Służą do przekazywania argumentów do funkcji, takich jak `member` i `oneof`.

Uwaga: Listy mogą być złożone tylko z obiektów statycznych (np. łańcuchów, liczb lub nazw zmiennych) i nie mogą zawierać wywołań funkcji.

Zmienne

Nazwy w wyrażeniach CLEM niebędące nazwami funkcji są traktowane jako nazwy zmiennych. Można je zapisać po prostu jako Power, val27, state_flag itd., lecz jeśli nazwa zaczyna się cyfrą lub obejmuje znaki inne niż znaki alfabetu, na przykład spacje (z wyjątkiem znaku podkreślenia), należy ująć tę nazwę w pojedyńcze znaki cudzysłowu –– na przykład, 'Power Increase', '2nd answer', '#101', '$P-NextField'.

Uwaga: Zmienne ujęte w cudzysłowy, ale niezdefiniowane w zestawie danych, zostaną błędnie odczytane jako łańcuchy.

Daty

Obliczenia dat są dokonywane na podstawie daty bazowej, określonej w oknie dialogowym właściwości strumienia. Domyślna data bazowa to 1 stycznia 1900.

Język CLEM obsługuje następujące formaty daty.

Tabela 15. Formaty daty języka CLEM

Format	**Przykłady**
DDMMRR	150163
MMDDRR	011563
RRMMDD	630115
RRRRMMDD	19630115
RRRRDDD	Czterocyfrowy rok wraz z trzycyfrową liczbą oznaczającą dzień roku; na przykład 2000032 oznacza 32-gi dzień roku 2000 lub 1 lutego 2000.
DAY	Dzień tygodnia w bieżących ustawieniach regionalnych — na przykład Poniedziałek, Wtorek, itd..., w języku polskim.
MONTH	Miesiąc w bieżących ustawieniach regionalnych — na przykład Styczeń, Luty, ...
DD/MM/RR	15/01/63
DD/MM/RRRR	15/01/1963
MM/DD/RR	01/15/63
MM/DD/RRRR	01/15/1963
DD-MM-RR	15-01-63
DD-MM-RRRR	15-01-1963
MM-DD-RR	01-15-63
MM-DD-RRRR	01-15-1963
DD.MM.RR	15.01.63
DD.MM.RRRR	15.01.1963
MM.DD.RR	01.15.63
MM.DD.RRRR	01.15.1963
DD-MMM-RR	15-STY-63, 15-sty-63, 15-Sty-63
DD/MMM/RR	15/STY/63, 15/sty/63, 15/Sty/63

Tabela 15. Formaty daty języka CLEM (kontynuacja)	
Format	**Przykłady**
DD.MMM.RR	15.STY.63, 15.sty.63, 15.Sty.63
DD-MMM-RRRR	15-STY-1963, 15-sty-1963, 15-Sty-1963
DD/MMM/RRRR	15/STY/1963, 15/sty/1963, 15/Sty/1963
DD.MMM.RRRR	15.STY.1963, 15.sty.1963, 15.Sty.1963
MON YYYY	Sty 2004
q Q YYYY	Cyfra (1–4) oznacza kwartał, litera *Q* kwartał, zaś liczba czterocyfrowa rok — na przykład 25 grudnia 2004 będzie przedstawiany jako 4 Q 2004.
tt WK RRRR	Dwucyfrowa liczba przedstawiająca tydzień roku, litery *WK* oraz czterocyfrowe oznaczenie roku. Tydzień roku jest obliczany przy założeniu, że pierwszy dzień tygodnia to poniedziałek, oraz że jako pierwszy tydzień roku uznaje się tydzień, w którym przypada choć jeden dzień danego roku.

Czas

W języku CLEM obsługiwane są poniższe formaty czasu.

Tabela 16. Formaty czasu w języku CLEM	
Format	**Przykłady**
HHMMSS	120112, 010101, 221212
HHMM	1223, 0745, 2207
MMSS	5558, 0100
HH:MM:SS	12:01:12, 01:01:01, 22:12:12
HH:MM	12:23, 07:45, 22:07
MM:SS	55:58, 01:00
(H)H:(M)M:(S)S	12:1:12, 1:1:1, 22:12:12
(H)H:(M)M	12:23, 7:45, 22:7
(M)M:(S)S	55:58, 1:0
HH.MM.SS	12.01.12, 01.01.01, 22.12.12
HH.MM	12.23, 07.45, 22.07
MM.SS	55.58, 01.00
(H)H.(M)M.(S)S	12.1.12, 1.1.1, 22.12.12
(H)H.(M)M	12.23, 7.45, 22.7
(M)M.(S)S	55.58, 1.0

Operatory CLEM

Dostępne są następujące operatory.

Tabela 17. Operatory języka CLEM

Operacja	**Komentarze**	**Pierwszeństwo (patrz następna sekcja)**
or	Używana między dwoma wyrażeniami CLEM. Zwraca wartość prawda, jeśli jedno lub oba wyrażenia są prawdziwe.	10
and	Używana między dwoma wyrażeniami CLEM. Zwraca wartość prawda, jeśli oba wyrażenia są prawdziwe.	9
=	Używana między dowolnymi dwoma elementami, które można ze sobą porównać. Zwraca wartość true, jeśli ELEMENT1 jest równy elementowi ELEMENT2.	7
==	Identyczna z =.	7
/=	Używana między dowolnymi dwoma elementami, które można ze sobą porównać. Zwraca wartość true, jeśli ELEMENT1 *nie* jest równy ELEMENTOWI2.	7
/==	Identyczna z /=.	7
>	Używana między dowolnymi dwoma elementami, które można ze sobą porównać. Zwraca wartość true, jeśli wartość ELEMENTU1 jest dokładnie większa od wartości ELEMENTU2.	6
>=	Używana między dowolnymi dwoma elementami, które można ze sobą porównać. Zwraca wartość true w przypadku rekordów, w których ELEMENT1 jest większy lub równy ELEMENTOWI2.	6
<	Używana między dowolnymi dwoma elementami, które można ze sobą porównać. Zwraca wartość true, jeśli wartość ELEMENTU1 jest dokładnie mniejsza od wartości ELEMENTU2	6
<=	Używana między dowolnymi dwoma elementami, które można ze sobą porównać. Zwraca wartość true, jeśli ELEMENT1 jest mniejszy lub równy ELEMENTOWI2.	6
&&=_0	Używana między dwiema liczbami całkowitymi. Równoważny logicznemu wyrażeniu LCAŁK1 && LCAŁK2 = 0.	6
&&/=_0	Używana między dwiema liczbami całkowitymi. Równoważny logicznemu wyrażeniu LCAŁK1 && LCAŁK2 /= 0.	6
+	Dodaje dwie liczby: LICZ1 + LICZ2.	5

Tabela 17. Operatory języka CLEM (kontynuacja)

Operacja	Komentarze	Pierwszeństwo (patrz następna sekcja)
><	Łączy dwa łańcuchy; na przykład ŁAŃCUCH1 >< ŁAŃCUCH2.	5
-	Odejmuje jedną liczbę od drugiej: LICZ1 - LICZ2. Może być również używana przed liczbą: - LICZBA.	5
*	Używana do mnożenia dwu liczb: LICZ1 * LICZ2.	4
&&	Używana między dwiema liczbami całkowitymi. Wynikiem jest bitowe 'and' liczb całkowitych LCAŁK1 i LCAŁK2.	4
&&~~	Używana między dwiema liczbami całkowitymi. Wynikiem jest bitowe "and" dla liczby LCAŁK1 oraz bitowe dopełnienie LCAŁK2.	4
||	Używana między dwiema liczbami całkowitymi. Wynikiem jest bitowe ‘inclusive or’ LCAŁK1 i LCAŁK2.	4
~~	Używana przed liczbą całkowitą. Generuje bitowe dopełnienie LCAŁK.	4
||/&	Używana między dwiema liczbami całkowitymi. Wynikiem jest bitowe ‘exclusive or’ LCAŁK1 i LCAŁK2.	4
LCAŁK1 << N	Używana między dwiema liczbami całkowitymi. Tworzy bitowy wzorzec LCAŁK przesuniętej w lewo o N miejsc.	4
LCAŁK1 >> N	Używana między dwiema liczbami całkowitymi. Tworzy bitowy wzorzec LCAŁK przesuniętej w prawo o N miejsc.	4
/	Używana do dzielenia jednej liczby przez drugą: LICZ1 / LICZ2.	4
**	Używana między dwiema liczbami: PODSTAWA ** POTĘGA. Zwraca PODSTAWĘ podniesioną do potęgi POTĘGA.	3
rem	Używana między dwiema liczbami całkowitymi: LCAŁK1 rem LCAŁK2. Zwraca resztę, LCAŁK1 - (LCAŁK1 div LCAŁK2) * LCAŁK2.	2
div	Używana między dwiema liczbami całkowitymi: LCAŁK1 div LCAŁK2. Wykonuje dzielenie całkowitoliczbowe.	2

Pierwszeństwo operatorów

Pierwszeństwo determinuje analizowanie składni złożonych wyrażeń, w szczególności wyrażeń bez nawiasów z więcej niż jednym operatorem. Na przykład

```
3 + 4 * 5
```

jest przetwarzane jako `3 + (4 * 5)`, nie zaś jako `(3 + 4) * 5`, ponieważ pierwszeństwo względne określa, że operacja * ma być przetwarzana przed operacją +. Każdy operator w języku CLEM ma przypisaną wartość pierwszeństwa; im niższa jest ta wartość, tym ważniejszy staje się on na liście analizowania składni, co oznacza, że będzie on przetwarzany przed innymi operatorami o wyższych wartościach określających pierwszeństwo.

Funkcje — informacje

Następujące funkcje CLEM są dostępne podczas opracowywania danych w oprogramowaniu IBM SPSS Modeler. Funkcje te można wprowadzać jako kody w wielu oknach dialogowych, takich jak np. węzły Wyliczanie, lub korzystać z Kreatora wyrażeń i tworzyć poprawne wyrażenia CLEM bez potrzeby zapamiętywania list funkcji ani nazw zmiennych.

Tabela 18. Funkcje CLEM przeznaczone do użycia z danymi IBM SPSS Modeler

Typ funkcji	**Opis**
Informacje	Służą do uzyskiwania informacji o wartościach zmiennych. Przykładowo: funkcja `is_string` zwraca wartość true dla wszystkich rekordów typu łańcuch.
Konwersja	Służą do tworzenia nowych zmiennych lub przekształcenia typu składowania. Przykładowo: funkcja `to_timestamp` przekształca wybraną zmienną na znacznik czasu.
Porównywanie	Służą do porównywania wartości zmiennych ze sobą lub z określonym łańcuchem. Przykładowo: `<=` służy do sprawdzania, czy wartości dwóch zmiennych są mniejsze od siebie czy równe.
Logiczne	Służą do wykonywania operacji logicznych, takich jak `if`, `then`, `else`.
Numeric	Służą do wykonywania obliczeń numerycznych, takich jak logarytm naturalny z wartości zmiennych.
Trygonometryczne	Służą do wykonywania obliczeń trygonometrycznych, takich jak arcus cosinus określonego kąta.
Prawdopodobieństwo	Zwracają prawdopodobieństwa oparte na różnych rozkładach, takie jak prawdopodobieństwo, że wartość rozkładu Student *t* jest mniejsza niż określona wartość.
Przestrzenne	Służą do wykonywania obliczeń przestrzennych z użyciem danych geoprzestrzennych.
Bitowe	Służą do przetwarzania liczb całkowitych w postaci ciągów bitów.
Losowe	Służą do losowego wybierania elementów lub generowania liczb.
Łańcuch	Służą do wykonywania różnych operacji na łańcuchach, np. funkcja `stripchar`, pozwala na usuwanie określonych znaków.
SoundEx	Służą do wyszukiwania łańcuchów, których dokładna treść nie jest znana. Wyszukiwanie jest oparte na założeniach fonetycznych dotyczących wymowy określonych liter.
Data i czas	Służą do wykonywania różnych operacji dotyczących zmiennych daty, godziny i znaczników czasu.

Tabela 18. Funkcje CLEM przeznaczone do użycia z danymi IBM SPSS Modeler (kontynuacja)	
Typ funkcji	**Opis**
Sekwencja podstawień	Służą do badania kolejności rekordów w zbiorze danych lub wykonywania operacji opartych na kolejności.
Globalne	Służą do uzyskiwania dostępu do wartości globalnych utworzonych za pomocą węzła wartości globalnych. Przykładowo: `@MEAN` służy do odwoływania się do średniej ze wszystkich wartości w przypadku zmiennej w całym zbiorze danych.
Puste i wartości null	Służą do oznaczania, wypełniania braków danych użytkownika lub systemowych braków danych, a także do uzyskiwania dostępu do tego typu danych. Przykładowo: `@BLANK(ZMIENNA)` służy do ustawiania flagi true w przypadku rekordów zawierających braki danych.
Funkcje specjalne	Służą do wyznaczania określonych danych z przeznaczeniem do zbadania. Przykładowo: zmienna `@FIELD` służy do wyliczania wielu zmiennych.

Konwencje zastosowane w opisach funkcji

W tym podręczniku, odwołując się do elementów funkcji, zastosowano następujące konwencje.

Tabela 19. Konwencje zastosowane w opisach funkcji	
Konwencja	**Opis**
BOOL	Wartość logiczna lub flaga (prawda lub fałsz).
LICZ, *LICZ1*, *LICZ2*	Dowolna liczba.
LICZBA RZECZYWISTA, *LICZBA RZECZYWISTA1*, *LICZBA RZECZYWISTA2*	Dowolna liczba rzeczywista, taka jak `1.234` lub `–77.01`.
LCAŁK, *LCAŁK1*, *LCAŁK2*	Dowolna liczba całkowita, taka jak `1` lub `–77`.
ZNAK	Kod znaku, na przykład `` `A` ``.
ŁAŃCUCH	Łańcuch, taki jak `"referrerID"`.
LISTA	Lista elementów, taka jak `["abc" "def"]`.
ELEMENT	Zmienna, taka jak `Customer` lub `extract_concept`.
DATA	Zmienna daty, taka jak `start_date`, gdzie wartości są w formacie `DD-MMM-RRRR`.
CZAS	Zmienna czasu, taka jak `power_flux`, gdzie wartości są w formacie `GGMMSS`.

Funkcje w tym podręczniku przedstawiono w taki sposób, że funkcja znajduje się w jednej kolumnie, typ wyniku (liczba całkowita, łańcuch itd.) w drugiej, zaś opis (tam, gdzie jest dostępny) w trzeciej. Na przykład poniżej znajduje się opis funkcji `rem`.

Tabela 20. opis funkcji rem		
Funkcja	**Wynik**	**Opis**
`LCAŁK1 rem LCAŁK2`	*Liczba*	Zwraca resztę z dzielenia *LCAŁK1* przez *LCAŁK2*. Przykładowo: `LCAŁK1 – (LCAŁK1 div LCAŁK2) * LCAŁK2`.

Szczegółowe informacje dotyczące konwencji użycia, takie jak sposób wyświetlania na liście lub określenie znaków funkcji wymieniono w innym miejscu. Więcej informacji można znaleźć w temacie “ Typy danych CLEM” na stronie 161.

Funkcje informacyjne

Funkcje informacyjne służą do uzyskiwania informacji na temat wartości danej zmiennej. Zazwyczaj są one używane do wyliczania zmiennych typu flaga. Przykładowo: za pomocą funkcji @BLANK można utworzyć zmienną typu flaga oznaczającą rekordy, których wartości są puste dla wybranej zmiennej. Można również sprawdzić typ składowania zmiennej, stosując dowolną funkcję typu składowania, np. `is_string`.

Tabela 21. Funkcje informacyjne CLEM

Funkcja	Wynik	Opis
`@BLANK(ZMIENNA)`	*Boolean*	Zwraca wartość true w przypadku wszystkich rekordów, których wartości są puste zgodnie z regułami obsługi wartości pustych określonymi w poprzedzającym węźle Typ węzła źródłowego (karta Typy).
`@NULL(ELEMENT)`	*Boolean*	Zwraca wartość true dla wszystkich rekordów, których wartości są niezdefiniowane. Wartości niezdefiniowane są systemowymi wartościami null wyświetlanymi w IBM SPSS Modeler jako `$null$`.
`is_date(ELEMENT)`	*Boolean*	Zwraca wartość true dla wszystkich rekordów typu data.
`is_datetime(ELEMENT)`	*Boolean*	Zwraca wartość true dla wszystkich rekordów typu data, czas lub znacznik czasu.
`is_integer(ELEMENT)`	*Boolean*	Zwraca wartość true dla wszystkich rekordów typu liczba całkowita.
`is_number(ELEMENT)`	*Boolean*	Zwraca wartość true dla wszystkich rekordów typu liczba.
`is_real(ELEMENT)`	*Boolean*	Zwraca wartość true dla wszystkich rekordów typu liczba rzeczywista.
`is_string(ELEMENT)`	*Boolean*	Zwraca wartość true dla wszystkich rekordów typu łańcuch.
`is_time(ELEMENT)`	*Boolean*	Zwraca wartość true dla wszystkich rekordów typu czas.
`is_timestamp(ELEMENT)`	*Boolean*	Zwraca wartość true dla wszystkich rekordów typu znacznik czasu.

Funkcje przekształcania

Funkcje przekształcania umożliwiają tworzenie nowych zmiennych i przekształcanie typu składowania istniejących plików. Przykładowo: można tworzyć nowe łańcuchy, łącząc ze sobą kilka łańcuchów lub rozdzielając je. Aby połączyć dwa łańcuchy, należy użyć operatora ><. Przykładowo: jeśli wartością zmiennej `Lokalizacja` jest `"BRAMLEY"`, to operacja `"xx" >< Lokalizacja` zwraca ciąg `"xxBRAMLEY"`. Wynikiem działania >< jest zawsze łańcuch, nawet jeśli argumenty nie są łańcuchami. Więc jeśli zmienna V1 wynosi 3, a zmienna V2 wynosi 5, to `V1 >< V2` zwraca `"35"` (łańcuch, a nie liczbę).

Funkcje przekształcania (i inne funkcje wymagające określonego typu danych wejściowych, np. wartość daty lub godziny) są uzależnione od bieżących formatów w oknie dialogowym Opcje łańcucha. Przykładowo: aby wykonać przekształcenia zmiennej łańcuchowej o wartościach *Sty 2003*, *Lut 2003* itd., jako domyślny format daty strumienia należy wybrać odpowiedni format daty **MIE RRRR**.

Tabela 22. Funkcje przekształcania CLEM

Funkcja	Wynik	Opis
ELEMENT1 >< ELEMENT2	*Łańcuch*	Łączy łańcuchy z dwóch zmiennych i zwraca łańcuch wynikowy w postaci *ELEMENT1ELEMENT2*.
to_integer(ELEMENT)	*Liczba całkowita*	Przekształca typ składowania określonej zmiennej na liczbę całkowitą.
to_real(ELEMENT)	*Liczba rzeczywista*	Przekształca typ składowania określonej zmiennej na liczbę rzeczywistą.
to_number(ELEMENT)	*Liczba*	Przekształca typ składowania określonej zmiennej na liczbę.
to_string(ELEMENT)	*Łańcuch*	Przekształca typ składowania określonej zmiennej na łańcuch. Po użyciu tej funkcji w celu wykonania przekształcenia liczby rzeczywistej do postaci łańcucha zwraca ona wartość z dokładnością do 6 cyfr po przecinku.
to_time(ELEMENT)	*Czas*	Przekształca typ składowania określonej zmiennej na godzinę.
to_date(ELEMENT)	*Date*	Przekształca typ składowania określonej zmiennej na datę.
to_timestamp(ELEMENT)	*Znacznik czasu*	Przekształca typ składowania określonej zmiennej na znacznik czasu.
to_datetime(ELEMENT)	*Datetime*	Przekształca typ składowania określonej zmiennej na datę, godzinę i znacznik czasu.
datetime_date(ELEMENT)	*Date*	Zwraca wartość daty dla *liczby*, *łańcucha* lub *znacznika czasu*. Jest to jedyna funkcja, która pozwala na wsteczne przekształcenie liczby (w sekundach) na datę. Jeśli ELEMENT jest łańcuchem, data jest tworzona poprzez przeanalizowanie łańcucha w bieżącym formacie daty. Format daty opisany w oknie dialogowym właściwości strumienia musi być poprawny, aby ta funkcja zadziałała prawidłowo. Jeśli ELEMENT jest liczbą, jest ona interpretowana jako liczba sekund, które upłynęły od daty podstawowej. Ułamkowe części wartości dni są obcinane. Jeśli ELEMENT jest znacznikiem czasu, zwracana jest część daty znacznika czasu. Jeśli ELEMENT jest datą, jest ona zwracana w niezmienionej postaci.
stb_centroid_latitude(ELEMENT)	*Liczba całkowita*	Zwraca liczbę całkowitą dla szerokości geograficznej odpowiadającą środkowi ciężkości argumentu geohash.
stb_centroid_longitude(ELEMENT)	*Liczba całkowita*	Zwraca liczbę całkowitą dla długości geograficznej odpowiadającą środkowi ciężkości argumentu geohash.

Tabela 22. Funkcje przekształcania CLEM (kontynuacja)		
Funkcja	**Wynik**	**Opis**
`to_geohash(ELEMENT)`	*Łańcuch*	Zwraca łańcuch geohash odpowiadający szerokości i długości geograficznej w gęstości określonej w liczbie bitów. Geohash to kod służący do określania zestawu współrzędnych geograficznych w oparciu o szczegółowe wartości szerokości i długości geograficznej. Trzy parametry `to_geohash` są następujące: • *szerokość geograficzna*: zakres (-180, 180), a jednostki to stopnie w układzie współrzędnych WGS84 • *długość geograficzna*: zakres (-90, 90), a jednostki to stopnie w układzie współrzędnych WGS84 • *bity*: liczba bitów, w których zostanie zapisany łańcuch hash. Zakres [1,75]. To wpływa na długość zwracanego ciągu znaków (1 znak jest używany na każde 5 bitów) oraz dokładność łańcucha hash. Na przykład 5 bitów (1 znak) reprezentuje około 2500 kilometrów, a 45 bitów (9 znaków) reprezentuje około 2,3 metra.

Funkcje porównawcze

Funkcje porównawcze służą do porównywania wartości zmiennych ze sobą lub z określonym łańcuchem. Przykładowo: można sprawdzić łańcuch pod kątem równości, używając operatora =. Przykładowa weryfikacja równości łańcucha: `Class = "class 1"`.

W przypadku wykonywania porównania liczbowego określenie *większe* oznacza wartość bliższą nieskończoności po stronie dodatniej, a określenie *mniejsze* — bliżej nieskończoności po stronie ujemnej. Czyli wszystkie liczby ujemne są mniejsze od dowolnej liczby dodatniej.

Tabela 23. Funkcje porównawcze CLEM		
Funkcja	**Wynik**	**Opis**
`count_equal(ELEMENT 1, LISTA)`	*Liczba całkowita*	Zwraca liczbę wartości z listy równych wartości *ELEMENT1* lub wartość null, jeśli wartość *ELEMENT1* ma wartość null.
`count_greater_than (ELEMENT1, LISTA)`	*Liczba całkowita*	Zwraca liczbę wartości z listy większych od wartości *ELEMENT1* lub wartość null, jeśli *ELEMENT1* ma wartość null.
`count_less_than(ELE MENT1, LISTA)`	*Liczba całkowita*	Zwraca liczbę wartości z listy mniejszych od wartości *ELEMENT1* lub wartość null, jeśli *ELEMENT1* ma wartość null.
`count_not_equal(ELE MENT1, LISTA)`	*Liczba całkowita*	Zwraca liczbę wartości z listy nierównych wartości *ELEMENT1* lub wartość null, jeśli *ELEMENT1* ma wartość null.
`count_nulls(LISTA)`	*Liczba całkowita*	Zwraca liczbę wartości null z listy.
`count_non_nulls(LIS TA)`	*Liczba całkowita*	Zwraca liczbę wartości innego typu niż null z listy.
`date_before(DATA1, DATA2)`	*Boolean*	Służy do sprawdzania kolejności dat. Zwraca wartość true, jeśli *DATA1* następuje przed *DATĄ2*.

Tabela 23. Funkcje porównawcze CLEM (kontynuacja)

Funkcja	Wynik	Opis
`first_index(ELEMENT, LISTA)`	*Liczba całkowita*	Zwraca indeks pierwszej zmiennej zawierającej ELEMENT na liście zmiennych LISTA lub 0, jeśli nie znaleziono danej wartości. Ta funkcja jest obsługiwana tylko w przypadku łańcuchów, liczb całkowitych i rzeczywistych.
`first_non_null(LISTA)`	*Any*	Zwraca pierwszą wartość inną niż null z listy. Obsługuje wszystkie typy składowania.
`first_non_null_index(LISTA)`	*Liczba całkowita*	Zwraca indeks pierwszej zmiennej na określonej LIŚCIE zawierającej wartość inną niż null lub 0, jeśli wszystkie wartości są równe null. Wszystkie typy składowania są obsługiwane.
`ELEMENT1 = ELEMENT2`	*Boolean*	Zwraca wartość true w przypadku rekordów, w których *ELEMENT1* jest równy *ELEMENT2*.
`ELEMENT1 /= ELEMENT2`	*Boolean*	Zwraca wartość true, jeśli dwa łańcuchy nie są takie same, lub 0, jeśli są takie same.
`ELEMENT1 < ELEMENT2`	*Boolean*	Zwraca wartość true w przypadku rekordów, w których *ELEMENT1* jest mniejszy niż *ELEMENT2*.
`ELEMENT1 <= ELEMENT2`	*Boolean*	Zwraca wartość true w przypadku rekordów, w których *ELEMENT1* jest równy *ELEMENT2* lub mniejszy.
`ELEMENT1 > ELEMENT2`	*Boolean*	Zwraca wartość true w przypadku rekordów, w których *ELEMENT1* jest większy niż *ELEMENT2*.
`ELEMENT1 >= ELEMENT2`	*Boolean*	Zwraca wartość true w przypadku rekordów, w których *ELEMENT1* jest równy *ELEMENT2* lub większy.
`last_index(ELEMENT, LISTA)`	*Liczba całkowita*	Zwraca indeks ostatniej zmiennej zawierającej element ELEMENT na liście zmiennych LISTA lub 0, jeśli nie znaleziono danej wartości. Ta funkcja jest obsługiwana tylko w przypadku łańcuchów, liczb całkowitych i rzeczywistych.
`last_non_null(LISTA)`	*Any*	Zwraca ostatnią wartość inną niż null z listy. Obsługuje wszystkie typy składowania.
`last_non_null_index(LISTA)`	*Liczba całkowita*	Zwraca indeks ostatniej zmiennej na określonej LIŚCIE zawierającej wartość inną niż null lub 0, jeśli wszystkie wartości są równe null. Wszystkie typy składowania są obsługiwane.
`max(ELEMENT1, ELEMENT2)`	*Any*	Zwraca większy z dwóch elementów: *ELEMENT1* lub *ELEMENT2*.
`max_index(LISTA)`	*Liczba całkowita*	Zwraca indeks zmiennej zawierającej maksymalną wartość z listy zmiennych numerycznych lub 0, jeśli wszystkie wartości są równe null. Przykładowo: jeśli trzecia zmienna na liście zawiera wartość maksymalną, zwracana jest wartość indeksu 3. Jeśli wiele zmiennych zawiera wartość maksymalną, zwracana jest wartość pierwsza (od lewej).
`max_n(LISTA)`	*Liczba*	Zwraca wartość maksymalną z listy zmiennych numerycznych lub null, jeśli wszystkie wartości zmiennych są typu null.

Tabela 23. Funkcje porównawcze CLEM (kontynuacja)		
Funkcja	**Wynik**	**Opis**
`member(ELEMENT, LISTA)`	*Boolean*	Zwraca wartość true, jeśli *ELEMENT* jest członkiem określonej *LISTY*. W przeciwnym razie zwraca wartość false. Można również określić listę nazw zmiennych.
`min(ELEMENT1, ELEMENT2)`	*Any*	Zwraca mniejszy z dwóch elementów: *ELEMENT1* lub *ELEMENT2*.
`min_index(LISTA)`	*Liczba całkowita*	Zwraca indeks zmiennej zawierającej minimalną wartość z listy zmiennych numerycznych lub 0, jeśli wszystkie wartości są typu NULL. Przykładowo: jeśli trzecia zmienna na liście zawiera wartość minimalną, zwracana jest wartość indeksu 3. Jeśli wiele zmiennych zawiera wartość minimalną, zwracana jest wartość pierwsza (od lewej).
`min_n(LISTA)`	*Liczba*	Zwraca wartość minimalną z listy zmiennych numerycznych lub null, jeśli wszystkie wartości zmiennych są typu null.
`time_before(CZAS1, CZAS2)`	*Boolean*	Służy do sprawdzania kolejności wartości czasu. Zwraca wartość true, jeśli *CZAS1* następuje przed wartością *CZAS2*.
`value_at(LCAŁK, LISTA)`		Zwraca wartość zmiennej znajdującej się na liście na pozycji LCAŁK lub wartość NULL, jeśli pozycja wykracza poza zakres prawidłowych wartości (czyli jest mniejsza niż 1 lub większa niż liczba zmiennych na liście). Obsługuje wszystkie typy składowania.

Funkcje logiczne

Wyrażenia CLEM służące do wykonywania operacji logicznych.

Tabela 24. Funkcje logiczne CLEM		
Funkcja	**Wynik**	**Opis**
`WAR1 and WAR2`	*Boolean*	Ta operacja stanowi koniunkcję logiczną i zwraca wartość true, jeśli obydwa warunki *WAR1* i *WAR2* mają wartość true. Jeśli wartością warunku *WAR1* jest false, to warunek *WAR2* nie jest oceniany. W ten sposób można określić koniunkcje, w przypadku których warunek *WAR1* najpierw sprawdza, czy operacja w warunku *WAR2* jest dozwolona. Przykładowo: `length(Label) >=6` i `Label(6) = 'x'`.
`WAR1 or WAR2`	*Boolean*	Ta operacja stanowi alternatywę logiczną i zwraca wartość true, jeśli jeden z warunków *WAR1* lub *WAR2* ma wartość true lub obydwa mają wartość true. Jeśli wartością warunku *WAR1* jest true, to warunek *WAR2* nie jest oceniany.
`not(WAR)`	*Boolean*	Ta operacja stanowi negację logiczną i zwraca wartość true, jeśli warunek *WAR* ma wartość false. W przeciwnym razie ta operacja zwraca wartość 0.
`if WAR then WYR1 else WYR2 endif`	*Any*	Ta operacja jest obliczeniem warunkowym. Jeśli warunek *WAR* ma wartość true, ta operacja zwraca wynik *WYR1*. W przeciwnym razie zwracany jest wynik *WYR2*.

Tabela 24. Funkcje logiczne CLEM (kontynuacja)		
Funkcja	**Wynik**	**Opis**
`if COND1 then EXPR1 elseif COND2 then EXPR2 else EXPR_N endif`	*Any*	Ta operacja jest rozgałęzionym obliczeniem warunkowym. Jeśli wartością warunku *WAR1* jest true, ta operacja zwraca wynik *WYR1*. W przeciwnym razie jeśli wartością warunku *WAR2* jest true, ta operacja zwraca wynik *WYR2*. W przeciwnym razie zwracany jest wynik *WYR_N*.

Funkcje numeryczne

CLEM zawiera wiele często używanych funkcji numerycznych.

Tabela 25. Funkcje numeryczne CLEM		
Funkcja	**Wynik**	**Opis**
`–LICZBA`	*Liczba*	Służy do negowania wartości *LICZBA*. Zwraca odpowiednią liczbę z odwrotnym znakiem.
`LICZ1 + LICZ2`	*Liczba*	Zwraca sumę *LICZ1* i *LICZ2*.
`LICZ1 –LICZ2`	*Liczba*	Zwraca wartość po odjęciu *LICZBY2* od *LICZBY1*.
`LICZ1 * LICZ2`	*Liczba*	Zwraca wartość mnożenia *LICZBY1* razy *LICZ2*.
`LICZ1 / LICZ2`	*Liczba*	Zwraca wartość dzielenia *LICZBY1* przez *LICZBĘ2*.
`LCAŁK1 div LCAŁK2`	*Liczba*	Służy do dzielenia liczb całkowitych. Zwraca wartość dzielenia *LCAŁK1* przez *LCAŁK2*.
`LCAŁK1 rem LCAŁK2`	*Liczba*	Zwraca resztę z dzielenia *LCAŁK1* przez *LCAŁK2*. Przykładowo: `LCAŁK1 – (LCAŁK1 div LCAŁK2) * LCAŁK2`.
`LCAŁK1 mod LCAŁK2`	*Liczba*	Ta funkcja nie jest już obsługiwana. Należy korzystać z funkcji `rem`.
`BASE ** POWER`	*Liczba*	Zwraca podstawę *BASE* podniesioną do potęgi *POWER*. Obydwie wartości mogą być dowolnymi liczbami (wartość *BASE* nie może być zerem, jeśli wartość *POWER* wynosi zero typu innego niż całkowitoliczbowy). Jeśli wartość *POWER* jest liczbą całkowitą, obliczenia są wykonywane poprzez kolejne mnożenie potęg *BASE*. Więc jeśli wartość *BASE* jest liczbą całkowitą, wynik będzie liczbą całkowitą. Jeśli wartością *POWER* jest liczba całkowitą równą 0, wynikiem jest zawsze 1 tego samego typu, co *BASE*. W przeciwnym razie, jeśli wartość *POWER* nie jest liczbą całkowitą, wynik jest obliczany jako `exp(POWER * log(BASE))`.
`abs(LICZ)`	*Liczba*	Zwraca wartość bezwzględną wartości *LICZ*, która zawsze jest liczbą tego samego typu.
`exp(LICZ)`	*Liczba rzeczywista*	Zwraca wartość *e* podniesioną do potęgi *LICZ*, gdzie *e* jest podstawą logarytmu naturalnego.
`fracof(LICZ)`	*Liczba rzeczywista*	Zwraca ułamkową część wartości *LICZ* zdefiniowaną jako `LICZ–intof(LICZ)`.
`intof(LICZ)`	*Liczba całkowita*	Skraca argument do liczby całkowitej. Zwraca liczbę całkowitą z tym samym znakiem co wartość *LICZ* i największą wartością, aby `abs(LCAŁK) <= abs(LICZ)`.

Tabela 25. Funkcje numeryczne CLEM (kontynuacja)		
Funkcja	**Wynik**	**Opis**
`log(LICZ)`	*Liczba rzeczywista*	Zwraca logarytm naturalny (o podstawie *e*) z wartości *LICZ*, która w żadnym wypadku nie może być zerem.
`log10(LICZ)`	*Liczba rzeczywista*	Zwraca logarytm naturalny o podstawie 10 z wartości *LICZ*, która w żadnym wypadku nie może być zerem. Ta funkcja jest definiowana jako `log(LICZ) / log(10)`
`negate(LICZ)`	*Liczba*	Służy do negowania wartości *LICZBA*. Zwraca odpowiednią liczbę z odwrotnym znakiem.
`round(LICZ)`	*Liczba całkowita*	Zaokrągla *LICZBĘ* do liczby całkowitej przez wzięcie `intof(LICZBA+0,5)`, jeżeli *LICZBA* jest dodatnia lub `intof(LICZBA-0,5)`, jeżeli *LICZBA* jest ujemna.
`sign(LICZ)`	*Liczba*	Służy do określania znaku wartości *LICZ*. Ta operacja zwraca –1, 0 lub 1, jeśli wartość *LICZ* jest liczbą całkowitą. Jeśli wartość *LICZ* jest liczbą rzeczywistą, zwraca –1,0, 0,0 lub 1,0, w zależności od tego czy wartość *LICZ* jest ujemna, zerowa czy dodatnia.
`sqrt(LICZ)`	*Liczba rzeczywista*	Zwraca pierwiastek kwadratowy z wartości *LICZ*. Wartość *LICZ* musi być dodatnia.
`sum_n(LISTA)`	*Liczba*	Zwraca sumę wartości z listy zmiennych numerycznych lub null, jeśli wszystkie wartości zmiennych są typu null.
`mean_n(LISTA)`	*Liczba*	Zwraca wartość średnią z listy zmiennych numerycznych lub null, jeśli wszystkie wartości zmiennych są typu null.
`sdev_n(LISTA)`	*Liczba*	Zwraca odchylenie standardowe z listy zmiennych numerycznych lub null, jeśli wszystkie wartości zmiennych są typu null.

Funkcje trygonometryczne

Wszystkie funkcje w tej sekcji mogą przyjąć kąt jako argument lub zwrócić kąt jako wynik. W obydwu przypadkach jednostkami kąta (radiany lub stopnie) steruje ustawienie odpowiedniej opcji strumienia.

Tabela 26. Funkcje trygonometryczne CLEM		
Funkcja	**Wynik**	**Opis**
`arccos(LICZ)`	*Liczba rzeczywista*	Wylicza arcus cosinus określonego kąta.
`arccosh(LICZ)`	*Liczba rzeczywista*	Wylicza hiperboliczny arcus cosinus określonego kąta.
`arcsin(LICZ)`	*Liczba rzeczywista*	Wylicza arcus sinus określonego kąta.
`arcsinh(LICZ)`	*Liczba rzeczywista*	Wylicza hiperboliczny arcus sinus określonego kąta.
`arctan(LICZ)`	*Liczba rzeczywista*	Wylicza arcus tangens określonego kąta.

Tabela 26. Funkcje trygonometryczne CLEM (kontynuacja)		
Funkcja	**Wynik**	**Opis**
`arctan2(LICZ_Y, LICZ_X)`	*Liczba rzeczywista*	Wylicza arcus tangens `LICZ_Y/LICZ_X`. Wykorzystuje znaki tych dwóch liczb do wyliczenia informacji o kwadrancie. Wynikiem jest liczba rzeczywista w zakresie `- pi < KĄT <= pi (radiany) – 180 < KĄT <= 180 (stopnie)`
`arctanh(LICZ)`	*Liczba rzeczywista*	Wylicza hiperboliczny arcus tangens określonego kąta.
`cos(LICZ)`	*Liczba rzeczywista*	Wylicza cosinus określonego kąta.
`cosh(LICZ)`	*Liczba rzeczywista*	Wylicza hiperboliczny cosinus określonego kąta.
`pi`	*Liczba rzeczywista*	Ta stała jest najlepszym rzeczywistym przybliżeniem liczby pi.
`sin(LICZ)`	*Liczba rzeczywista*	Wylicza sinus określonego kąta.
`sinh(LICZ)`	*Liczba rzeczywista*	Wylicza hiperboliczny sinus określonego kąta.
`tan(LICZ)`	*Liczba rzeczywista*	Wylicza tangens określonego kąta.
`tanh(LICZ)`	*Liczba rzeczywista*	Wylicza hiperboliczny tangens określonego kąta.

Funkcje prawdopodobieństwa

Funkcje prawdopodobieństwa zwracają prawdopodobieństwa oparte na różnych rozkładach, takie jak prawdopodobieństwo, że wartość rozkładu Student *t* jest mniejsza niż określona wartość.

Tabela 27. Funkcje prawdopodobieństwa CLEM		
Funkcja	**Wynik**	**Opis**
`cdf_chisq(LICZ, SS)`	*Liczba rzeczywista*	Zwraca informacje o prawdopodobieństwie tego, że wartość rozkładu chi-kwadrat w określonych stopniach swobody będzie mniejsza niż określona liczba.
`cdf_f(LICZ, SS1, SS2)`	*Liczba rzeczywista*	Zwraca informacje o prawdopodobieństwie tego, że wartość rozkładu *F* w stopniach swobody *DF1* i *DF2* będzie mniejsza niż określona liczba.
`cdf_normal(LICZ, ŚREDNIA, ODCHSTD)`	*Liczba rzeczywista*	Zwraca informacje o prawdopodobieństwie tego, że wartość rozkładu normalnego ze zdefiniowaną wartością średnią i odchyleniem standardowym będzie mniejsza niż określona liczba.
`cdf_t(LICZ, SS)`	*Liczba rzeczywista*	Zwraca informacje o prawdopodobieństwie tego, że wartość rozkładu Student *t* w określonych stopniach swobody będzie mniejsza niż określona liczba.

Funkcje przestrzenne

Funkcje przestrzenne mogą być używane z danymi geoprzestrzennymi. Przykładowo: umożliwiają one obliczanie odległości między dwoma punktami, obliczanie obszaru wielokąta itd. Mogą również wystąpić

sytuacje wymagające połączenia wielu zbiorów danych geoprzestrzennych opartych na predykacie przestrzennym (within, close to itd.). Tę operację można wykonać za pomocą warunku łączenia.

Funkcje przestrzenne operują w układzie współrzędnych określonym w menu **Narzędzia** > **Właściwości strumienia** > **Opcje** > **Geoprzestrzenne**.

Uwaga: Funkcje przestrzenne nie są stosowane w przypadku danych trójwymiarowych. Jeśli dane trójwymiarowe zostaną zaimportowane do strumienia, tylko dwa pierwsze wymiary są używane przez funkcje. Wartości na osi z są ignorowane.

Tabela 28. Funkcje przestrzenne CLEM

Funkcja	Wynik	Opis
`close_to(KSZTAŁT,KSZTAŁT,LICZ)`	*Boolean*	Testuje, czy 2 kształty mieszczą się w określonej ODLEGŁOŚCI od siebie. W przypadku użycia rzutowanego układu współrzędnych wartość ODLEGŁOŚĆ jest wyrażona w metrach. Jeśli żaden układ współrzędnych nie jest używany, jest to jednostka dowolna.
`crosses(KSZTAŁT,KSZTAŁT)`	*Boolean*	Testuje, czy 2 kształty przecinają się ze sobą. Ta funkcja ma zastosowanie w przypadku 2 kształtów liniowych lub 1 liniowego i 1 typu wielokąt.
`overlap(KSZTAŁT,KSZTAŁT)`	*Boolean*	Testuje, czy 2 wieloboki mają część wspólną oraz czy ta część wspólna występuje wewnątrz obu kształtów.
`within(KSZTAŁT,KSZTAŁT)`	*Boolean*	Testuje, czy całość KSZTAŁTU1 jest zawarta w WIELOKĄCIE.
`area(KSZTAŁT)`	*Liczba rzeczywista*	Zwraca pole określonego WIELOKĄTA. W przypadku rzutowanego układu współrzędnych funkcja zwraca metry kwadratowe. Jeśli żaden układ współrzędnych nie jest używany, jest to jednostka dowolna. Kształt musi być WIELOKĄTEM lub ZBIOREM WIELOKĄTÓW.
`num_points(KSZTAŁT,LISTA)`	*Liczba całkowita*	Zwraca liczbę punktów zmiennej punktowej (MULTIPUNKT) zawartych w granicach WIELOKĄTA. KSZTAŁT1 musi być WIELOKĄTEM lub ZBIOREM WIELOKĄTÓW.
`distance(KSZTAŁT,KSZTAŁT)`	*Liczba rzeczywista*	Zwraca odległość między kształtami KSZTAŁT1 a KSZTAŁT2. W przypadku rzutowanego układu współrzędnych funkcja zwraca metry. Jeśli żaden układ współrzędnych nie jest używany, jest to jednostka dowolna. Typ wartości KSZTAŁT1 i KSZTAŁT2 może być dowolnym typem geopomiaru.

Operacje bitowe na liczbach całkowitych

Te funkcje pozwalają na manipulowanie liczbami całkowitymi jako ciągami bitów w kodzie uzupełnień do dwóch, gdzie pozycja bitu N ma wagę 2**N. Bity są numerowane od 0. Te operacje działają tak jakby bit znaku liczby całkowitej był rozszerzany w nieskończoność w lewą stronę. Więc wszędzie powyżej jego najbardziej znaczącego bitu dodatnia liczba całkowita ma 0 bitów, a ujemna liczba całkowita ma 1 bit.

Tabela 29. Operacja bitowe CLEM na liczbach całkowitych

Funkcja	Wynik	Opis
`~~ LCAŁK1`	*Liczba całkowita*	Tworzy bitowe dopełnienie liczby całkowitej *LCAŁK1*. Czyli wynik zawiera 1 na każdej pozycji bitowej, która zawiera 0 w *LCAŁK1*. Zawsze prawdziwe jest twierdzenie `~~ LCAŁK = -(LCAŁK + 1)`.
`LCAŁK1 || LCAŁK2`	*Liczba całkowita*	Wynikiem tej operacji jest bitowe "inclusive or" z *LCAŁK1* i *LCAŁK2*. Czyli wynik zawiera 1 na każdej pozycji bitowej, która zawiera 1 w *LCAŁK1* lub *LCAŁK2* lub w obydwu.
`LCAŁK1 ||/& LCAŁK2`	*Liczba całkowita*	Wynikiem tej operacji jest bitowe "exclusive or" z *LCAŁK1* i *LCAŁK2*. Czyli wynik zawiera 1 na każdej pozycji bitowej, która zawiera 1 w *LCAŁK1* lub *LCAŁK2*, ale nie w obydwu.
`LCAŁK1 && LCAŁK2`	*Liczba całkowita*	Wynikiem jest bitowe "and" liczb całkowitych *LCAŁK1* i *LCAŁK2*. Czyli wynik zawiera 1 na każdej pozycji bitowej, na której 1 występuje i w *LCAŁK1*, i w *LCAŁK2*.
`LCAŁK1 &&~~ LCAŁK2`	*Liczba całkowita*	Wynikiem jest bitowe "and" dla liczby *LCAŁK1* oraz bitowe dopełnienie *LCAŁK2*. Czyli wynik zawiera 1 na każdej pozycji bitowej, która zawiera 1 w *LCAŁK1* i 0 w *LCAŁK2*. Odpowiada to wartości `LCAŁK1 && (~~LCAŁK2)` i jest przydatne do zerowania bitów *LCAŁK1* określonych w *LCAŁK2*.
`LCAŁK << N`	*Liczba całkowita*	Tworzy ciąg bitów *LCAŁK1* przesunięty w lewo o *N* miejsc. Wartość ujemna *N* powoduje przesunięcie w prawo.
`LCAŁK >> N`	*Liczba całkowita*	Tworzy ciąg bitów *LCAŁK1* przesunięty w prawo o *N* miejsc. Wartość ujemna *N* powoduje przesunięcie w lewo.
`LCAŁK1 &&=_0 LCAŁK2`	*Boolean*	Funkcja równoważna logicznemu wyrażeniu `LCAŁK1 && LCAŁK2 /== 0`, ale bardziej wydajna w działaniu.
`LCAŁK1 &&/=_0 LCAŁK2`	*Boolean*	Funkcja równoważna logicznemu wyrażeniu `LCAŁK1 && LCAŁK2 == 0`, ale bardziej wydajna w działaniu.
`integer_bitcount(LCAŁK)`	*Liczba całkowita*	Zlicza liczbę bitów 1 lub 0 w dopełnieniu do dwóch reprezentacji *LCAŁK*. Jeśli wartość *LCAŁK* jest nieujemna, *N* jest liczbą bitów 1. Jeśli wartość *LCAŁK* jest ujemna, N jest liczbą bitów 0. Ze względu na rozszerzenie znaku istnieje nieskończona liczba bitów 0 w nieujemnej liczbie całkowitej lub bitów 1 w ujemnej liczbie całkowitej. Zawsze `integer_bitcount(LCAŁK) = integer_bitcount(-(LCAŁK+1))`.
`integer_leastbit(LCAŁK)`	*Liczba całkowita*	Zwraca pozycję bitu *N* najmniej znaczącego zbioru bitów w liczbie całkowitej *LCAŁK*. *N* jest najwyższą potęgą 2, przez którą *LCAŁK* dzieli się bez reszty.

Tabela 29. Operacja bitowe CLEM na liczbach całkowitych (kontynuacja)		
Funkcja	**Wynik**	**Opis**
`integer_length(LCAŁK)`	*Liczba całkowita*	Zwraca długość bitową *LCAŁK* jako liczbę całkowitą w kodzie uzupełnień do dwóch. Czyli *N* jest najmniejszą liczbą całkowitą taką, że `LCAŁK < (1 << N)`, `jeśli LCAŁK >= 0 LCAŁK >= (–1 << N)`, `jeśli LCAŁK < 0`. Jeśli wartość *LCAŁK* jest nieujemna, reprezentacja *LCAŁK* w postaci liczby całkowitej bez znaku wymaga zmiennej o długości co najmniej *N* bitów. Co najmniej *N+1* bitów jest wymaganych w celu utworzenia reprezentacji *LCAŁK* jako liczby całkowitej ze znakiem (bez względu na znak).
`testbit(LCAŁK, N)`	*Boolean*	Testuje bit w pozycji *N* w liczbie całkowitej *LCAŁK* i zwraca stan bitu *N* jako wartość logiczną, która ma wartość true dla 1 i false dla 0.

Funkcje losowe

Poniższe funkcje służą do losowego wybierania elementów lub generowania liczb.

Tabela 30. Funkcje losowe CLEM		
Funkcja	**Wynik**	**Opis**
`oneof(LISTA)`	*Any*	Zwraca losowo wybrany element *LISTY*. Elementy listy powinny być wprowadzone w następujący sposób: `[ELEMENT1,ELEMENT2,...,ELEMENT_N]`. Można również określić listę nazw zmiennych.
`random(LICZ)`	*Liczba*	Zwraca losową liczbę z rozkładu jednostajnego tego samego typu (*LCAŁK* lub *RZECZ*), od 1 do *LICZ*. Jeśli użyta zostanie liczba całkowita, zwrócona zostanie tylko liczba całkowita. Jeśli użyta zostanie liczba rzeczywista (dziesiętna), zwracane są liczby rzeczywiste (dokładność dziesiętna określana przez opcje strumienia). Największa liczba rzeczywista zwracana przez tę funkcję może się równać wartości *LICZ*.
`random0(LICZ)`	*Liczba*	Ma takie same właściwości jak `random(LICZ)`, ale zaczyna losowanie od 0. Największa liczba losowa zwracana przez funkcję nigdy nie będzie równa wartości *LICZ*.

Funkcje łańcuchowe

W języku CLEM na łańcuchach można wykonać następujące operacje:

- Porównywanie łańcuchów
- Tworzenie łańcuchów
- Uzyskiwanie dostępu do znaków

W języku CLEM, łańcuch to sekwencja znaków ujętych w podwójne cudzysłowy (`"łańcuch w cudzysłowie"`). Znaki (ZNAK) mogą być dowolnymi, pojedynczymi znakami alfanumerycznymi. Są one deklarowane w wyrażeniach CLEM za pomocą pojedynczych apostrofów odwrotnych w postaci `` `<znak>` ``, np. `` `z` ``, `` `A` `` lub `` `2` ``. Znaki spoza zakresu lub indeksy ujemne łańcucha spowodują występowanie nieoczekiwanych zachowań.

Uwaga: Porównania między łańcuchami, które nie korzystają ze wstępnego generowania kodu SQL, mogą zwracać odmienne wyniki w przypadku, gdy istnieją spacje końcowe.

Tabela 31. Funkcje łańcuchowe CLEM

Funkcja	Wynik	Opis
`allbutfirst(N, ŁAŃCUCH)`	*Łańcuch*	Zwraca łańcuch, który jest *ŁAŃCUCHEM* z usuniętą liczbą *N* pierwszych znaków.
`allbutlast(N, ŁAŃCUCH)`	*Łańcuch*	Zwraca łańcuch, który jest *ŁAŃCUCHEM* z usuniętymi ostatnimi znakami.
`alphabefore(ŁAŃCUCH1, ŁAŃCUCH2)`	*Boolean*	Służy do sprawdzania kolejności alfabetycznej łańcuchów. Zwraca wartość true, jeśli *ŁAŃCUCH1* poprzedza *ŁAŃCUCH2*.
`endstring(DŁUGOŚĆ, ŁAŃCUCH)`	*Łańcuch*	Wyodrębnia ostatnich *N* znaków z określonego łańcucha. Jeśli długość łańcucha jest równa określonej długości (lub krótsza), zmiany nie są wprowadzane.
`hasendstring(ŁAŃCUCH, PODŁAŃCUCH)`	*Liczba całkowita*	Ta funkcja działa tak samo jak `isendstring(PODŁAŃCUCH, ŁAŃCUCH)`.
`hasmidstring(ŁAŃCUCH, PODŁAŃCUCH)`	*Liczba całkowita*	Ta funkcja działa tak samo jak `ismidstring(PODŁAŃCUCH, ŁAŃCUCH)` (osadzony podłańcuch).
`hasstartstring(ŁAŃCUCH, PODŁAŃCUCH)`	*Liczba całkowita*	Ta funkcja działa tak samo jak `isstartstring(PODŁAŃCUCH, ŁAŃCUCH)`.
`hassubstring(ŁAŃCUCH, N, PODŁAŃCUCH)`	*Liczba całkowita*	Ta funkcja działa tak samo jak `issubstring(PODŁAŃCUCH, N, ŁAŃCUCH)`, gdzie w przypadku *N* jest przywracana wartość domyślna 1.
`count_substring(ŁAŃCUCH, PODŁAŃCUCH)`	*Liczba całkowita*	Zwraca liczbę wystąpień określonego podłańcucha w łańcuchu. Na przykład `count_substring("foooo.txt", "oo")` zwraca 3.
`hassubstring(ŁAŃCUCH, PODŁAŃCUCH)`	*Liczba całkowita*	Ta funkcja działa tak samo jak `issubstring(PODŁAŃCUCH, 1, ŁAŃCUCH)`, gdzie w przypadku *N* jest przywracana wartość domyślna 1.
`isalphacode(ZNAK)`	*Boolean*	Zwraca wartość true, jeżeli *ZNAK* w łańcuchu (często określonym jako nazwa zmiennej) jest literą. W przeciwnym razie ta funkcja zwraca wartość 0. Przykładowo: `isalphacode(produce_num(1))`.
`isendstring(PODŁAŃCUCH, ŁAŃCUCH)`	*Liczba całkowita*	Jeśli łańcuch *ŁAŃCUCH* kończy się podłańcuchem *PODŁAŃCUCH*, to ta funkcja zwraca całkowitoliczbowy indeks dolny *PODŁAŃCUCHA* w *ŁAŃCUCHU*. W przeciwnym razie ta funkcja zwraca wartość 0.

Tabela 31. Funkcje łańcuchowe CLEM (kontynuacja)

Funkcja	Wynik	Opis
`islowercode(ZNAK)`	*Boolean*	Zwraca wartość true, jeżeli *ZNAK* w łańcuchu (często określonego jako nazwa zmiennej) jest małą literą. W przeciwnym razie ta funkcja zwraca wartość 0. Przykładowo: obydwa wyrażenia `islowercode(``)` i `islowercode(country_name(2))` są prawidłowe.
`ismidstring(PODŁAŃCUCH, ŁAŃCUCH)`	*Liczba całkowita*	Jeśli *PODŁAŃCUCH* jest podłańcuchem *ŁAŃCUCHA*, ale nie zaczyna się w pierwszym znaku *ŁAŃCUCHA* ani nie kończy na ostatnim, ta funkcja zwraca indeks, od którego rozpoczyna się podłańcuch. W przeciwnym razie ta funkcja zwraca wartość 0.
`isnumbercode(ZNAK)`	*Boolean*	Zwraca wartość true, jeżeli *ZNAK* w łańcuchu (często określonym jako nazwa zmiennej) jest cyfrą. W przeciwnym razie ta funkcja zwraca wartość 0. Przykładowo: `isnumbercode(product_id(2))`.
`isstartstring(PODŁAŃCUCH, ŁAŃCUCH)`	*Liczba całkowita*	Jeśli łańcuch *ŁAŃCUCH* zaczyna się podłańcuchem *PODŁAŃCUCH*, to ta funkcja zwraca indeks 1. W przeciwnym razie ta funkcja zwraca wartość 0.
`issubstring(PODŁAŃCUCH, N, ŁAŃCUCH)`	*Liczba całkowita*	Przeszukuje łańcuch *ŁAŃCUCH*, rozpoczynając od *N*-tego znaku, w poszukiwaniu podłańcucha równego *PODŁAŃCUCH*. Jeśli podłańcuch zostanie znaleziony, ta funkcja zwraca całkowitoliczbowy indeks, od którego rozpoczyna się zgodny podłańcuch. W przeciwnym razie ta funkcja zwraca wartość 0. Jeśli wartość *N* nie jest podana, przywracana jest wartość domyślna funkcji wynosząca 1.
`issubstring(PODŁAŃCUCH, ŁAŃCUCH)`	*Liczba całkowita*	Przeszukuje łańcuch *ŁAŃCUCH*, rozpoczynając od *N*-tego znaku, w poszukiwaniu podłańcucha równego *PODŁAŃCUCH*. Jeśli podłańcuch zostanie znaleziony, ta funkcja zwraca całkowitoliczbowy indeks, od którego rozpoczyna się zgodny podłańcuch. W przeciwnym razie ta funkcja zwraca wartość 0. Jeśli wartość *N* nie jest podana, przywracana jest wartość domyślna funkcji wynosząca 1.
`issubstring_count(PODŁAŃCUCH, N, ŁAŃCUCH):`	*Liczba całkowita*	Zwraca indeks *N*-tego wystąpienia *PODŁAŃCUCHA* w określonym *ŁAŃCUCHU*. Jeśli istnieje mniej niż *N* wystąpień *PODŁAŃCUCHA*, zwracana jest wartość 0.

Tabela 31. Funkcje łańcuchowe CLEM (kontynuacja)

Funkcja	Wynik	Opis
`issubstring_lim(PODŁAŃCUCH, N, OGRSTART, OGRKOŃ, ŁANCUCH)`	*Liczba całkowita*	Ta funkcja działa tak samo jak funkcja `issubstring`, lecz dopasowanie jest ograniczone do fragmentu od indeksu *OGRSTART* do indeksu *OGRKOŃ*. Ograniczenia *OGRSTART* lub *OGRKOŃ* można wyłączyć, wprowadzając wartość false dla dowolnego argumentu; przykładowo: `issubstring_lim(PODŁAŃCUCH, N, false, false, ŁAŃCUCH)` odpowiada `issubstring`.
`isuppercode(ZNAK)`	*Boolean*	Zwraca wartość true, jeżeli *ZNAK* jest wielką literą. W przeciwnym razie ta funkcja zwraca wartość 0. Przykładowo: obydwa wyrażenia `isuppercode(``)` i `isuppercode(country_name(2))` są prawidłowe.
`last(ZNAK)`	*Łańcuch*	Zwraca ostatni *ZNAK ŁAŃCUCHA* (który musi mieć przynajmniej jeden znak).
`length(ŁAŃCUCH)`	*Liczba całkowita*	Zwraca długość *ŁAŃCUCHA*, czyli liczbę zawartych w nim znaków.
`locchar(ZNAK, N, ŁAŃCUCH)`	*Liczba całkowita*	Służy do określania położenia znaków w zmiennych symbolicznych. Ta funkcja przeszukuje łańcuch *ŁAŃCUCH* pod kątem znaku *ZNAK*, rozpoczynając przeszukiwanie od *N*-tego znaku w *ŁAŃCUCHU*. Ta funkcja zwraca wartość wskazującą położenie (od *N*) znalezionego znaku. Jeśli znak nie zostanie znaleziony, funkcja zwraca wartość 0. Jeśli wartość przesunięcia funkcji *(N)* jest nieprawidłowa (przykładowo: wartość przesunięcia jest większa niż długość łańcucha), funkcja zwraca wartość $`null`$. Przykładowo: `locchar(`n`, 2, web_page)` przeszukuje zmienną *web_page* pod kątem znaku `n`, rozpoczynając od drugiego znaku w wartości zmiennej. *Uwaga*: dany znak należy ujmować w pojedyncze apostrofy odwrotne.

Tabela 31. Funkcje łańcuchowe CLEM (kontynuacja)

Funkcja	Wynik	Opis
`locchar_back(ZNAK, N, ŁAŃCUCH)`	*Liczba całkowita*	Działanie funkcji jest podobne do `locchar`, jednak wyszukiwanie jest wykonywane do tyłu, począwszy od *N*-tego znaku. Przykładowo: `locchar_back(`n`, 9, web_page)` przeszukuje zmienną *web_page*, począwszy od 9. znaku i przechodzi w tył w kierunku początku łańcucha. Jeśli wartość przesunięcia funkcji jest nieprawidłowa (przykładowo: wartość przesunięcia jest większa niż długość łańcucha), funkcja zwraca wartość `$null$`. Zaleca się używanie funkcji `locchar_back` wraz z funkcją `length(<field>)`, pozwoli to na dynamiczne używanie długości bieżącej wartości zmiennej. Przykładowo: `locchar_back(`n`, (length(web_page)), web_page)`.

`lowertoupper(ZNAK)`
 	
`lowertoupper (ŁAŃCUCH)`
	ZNAK lub *ŁAŃCUCH*	Można wprowadzić łańcuch lub znak. Funkcja zwróci nowy element tego samego typu, a wszystkie znaki zapisane małymi literami zostaną przekształcone na wielkie litery. Przykładowo: `lowertoupper(`a`)`, `lowertoupper("Mój łańcuch")` i `lowertoupper(nazwa_zmiennej(2))` są prawidłowymi wyrażeniami.

`matches`	*Boolean*	Zwraca wartość true, jeśli łańcuch odpowiada określonemu wzorcowi. Wzorcem musi być literał łańcuchowy, nie może to być nazwa zmiennej zawierająca wzorzec. Znak zapytania (?) można uwzględnić we wzorcu. Ten znak odpowiada dokładnie jednemu znakowi. Znak gwiazdki (*) odpowiada zerowej lub dowolnej liczbie znaków. Aby wyszukać rzeczywisty znak zapytania lub gwiazdki (a nie używać tych znaków jako symboli wieloznacznych), należy je poprzedzić znakiem ukośnika odwrotnego (\).
`replace(PODŁAŃCUCH, NOWYPODŁAŃCUCH, ŁAŃCUCH)`	*Łańcuch*	W określonym *ŁAŃCUCHU* zastępuje wszystkie wystąpienia *PODŁAŃCUCHA* za pomocą *NOWEGOPODŁAŃCUCHA*.
`replicate(LICZBA, ŁAŃCUCH)`	*Łańcuch*	Zwraca łańcuch składający się z oryginalnego łańcucha skopiowanego określoną liczbę razy.

Tabela 31. Funkcje łańcuchowe CLEM (kontynuacja)

Funkcja	Wynik	Opis
`stripchar(ZNAK,ŁAŃCUCH)`	*Łańcuch*	Umożliwia usuwanie określonych znaków z łańcucha lub zmiennej. Ta funkcja służy na przykład do usuwania z danych dodatkowych symboli, takich jak oznaczenia waluty, w celu uzyskania prostej liczby lub nazwy. Przykładowo: użycie składni `stripchar(`$`, 'Cost')` zwraca nową zmienną pozbawioną znaku dolara we wszystkich wartościach. *Uwaga*: dany znak należy ujmować w pojedyncze apostrofy odwrotne.
`skipchar(ZNAK, N, ŁAŃCUCH)`	*Liczba całkowita*	Przeszukuje łańcuch *ŁAŃCUCH* pod kątem znaku innego niż *ZNAK*, rozpoczynając przeszukiwanie od *N*-tego znaku. Ta funkcja zwraca podłańcuch liczby całkowitej wskazujący punkt, w którym znaleziono znak lub wartość 0, jeśli wszystkie znaki od znaku *N*-tego w przód są wartościami typu *ZNAK*. Jeśli wartość przesunięcia funkcji jest nieprawidłowa (przykładowo: wartość przesunięcia jest większa niż długość łańcucha), funkcja zwraca wartość $null$. Funkcja `locchar` jest często używana wraz z funkcjami `skipchar` w celu określenia wartości *N* (punktu rozpoczęcia przeszukiwania łańcucha). Przykładowo: `skipchar(`s`, (locchar(`s`, 1, "MójŁańcuch")), "MójŁańcuch")`.
`skipchar_back(ZNAK, N, ŁAŃCUCH)`	*Liczba całkowita*	Działanie funkcji jest podobne do `skipchar`, jednak wyszukiwanie jest wykonywane **do tyłu**, począwszy od *N*-tego znaku.
`startstring(DŁUGOŚĆ, ŁAŃCUCH)`	*Łańcuch*	Wyodrębnia pierwszych *N* znaków z określonego łańcucha. Jeśli długość łańcucha jest równa określonej długości (lub krótsza), zmiany nie są wprowadzane.
`strmember(ZNAK, ŁAŃCUCH)`	*Liczba całkowita*	Funkcja działa tak samo jak `locchar(ZNAK, 1, ŁAŃCUCH)`. Zwraca podłańcuch liczby całkowitej od miejsca wystąpienia pierwszego *ZNAKU*, lub wartość 0. Jeśli wartość przesunięcia funkcji jest nieprawidłowa (przykładowo: wartość przesunięcia jest większa niż długość łańcucha), funkcja zwraca wartość $null$.

Tabela 31. Funkcje łańcuchowe CLEM (kontynuacja)		
Funkcja	**Wynik**	**Opis**
`subscrs(N, ŁAŃCUCH)`	*ZNAK*	Zwraca *N*-ty znak *ZNAK* łańcucha wejściowego *ŁAŃCUCH*. Tę funkcję można również zapisać w skróconej postaci jako `ŁAŃCUCH(N)`. Przykładowo: `lowertoupper("nazwa"(1))` jest prawidłowym wyrażeniem.
`substring(N, DŁU, ŁAŃCUCH)`	*Łańcuch*	Zwraca łańcuch *PODŁAŃCUCH*, zawierający *DŁU* znaków *ŁAŃCUCHA*, począwszy od znaku *N*.
`substring_between(N1, N2, ŁAŃCUCH)`	*Łańcuch*	Zwraca podłańcuch *ŁAŃCUCHA*, który rozpoczyna się od indeksu *N1* i kończy w *N2*.
`trim(ŁAŃCUCH)`	*Łańcuch*	Usuwa spacje początkowe i końcowe z określonego łańcucha.
`trim_start(ŁAŃCUCH)`	*Łańcuch*	Usuwa spacje początkowe z określonego łańcucha.
`trimend(ŁAŃCUCH)`	*Łańcuch*	Usuwa spacje końcowe z określonego łańcucha.
`unicode_char(LICZ)`	*ZNAK*	Wartość wejściowa musi być wartością dziesiętną, a nie szesnastkową. Zwraca znak o kodzie *LICZ* w kodowaniu Unicode.
`unicode_value(ZNAK)`	*LICZ*	Zwraca kod Unicode dla *ZNAKU*

`uppertolower(ZNAK)`
 	
`uppertolower (ŁAŃCUCH)`
	ZNAK lub *ŁAŃCUCH*	Można wprowadzić łańcuch lub znak. Funkcja zwróci nowy element tego samego typu, a wszystkie znaki zapisane wielkimi literami zostaną przekształcone na małe litery.
		
		Uwaga: łańcuchy należy ująć w podwójne cudzysłowy; znaki należy ujmować w pojedyncze apostrofy odwrotne. Proste nazwy zmiennych należy wprowadzić bez cudzysłowów.

Funkcje SoundEx

Za pomocą metody SoundEx można wyszukiwać łańcuchy, gdy znana jest wymowa, ale nie dokładna pisownia. Tę metodę opracowano w 1918 roku. Pozwala ona na wyszukiwanie słów brzmiących podobnie. Działanie metody jest oparte na założeniach fonetycznych dotyczących wymowy określonych liter. Metoda ta służy do wyszukiwania nazw w bazach danych, np. gdy pisownia i wymowa podobnych nazw jest różna. Podstawowy algorytm SoundEx opisano w wielu źródłach i pomimo znanych ograniczeń (np. litery początkowe, takie jak ph i f nie są wyszukiwane, pomimo ich identycznego brzmienia), jest on obsługiwany w większości baz danych.

Tabela 32. Funkcje soundex CLEM		
Funkcja	**Wynik**	**Opis**
`soundex(ŁAŃCUCH)`	*Liczba całkowita*	Zwraca 4-znakowy kod SoundEx dla określonego *ŁAŃCUCHA*.

Tabela 32. Funkcje soundex CLEM (kontynuacja)		
Funkcja	**Wynik**	**Opis**
soundex_difference(ŁAŃCUCH1, ŁAŃCUCH2)	*Liczba całkowita*	Zwraca liczbę całkowitą w zakresie od 0 do 4 oznaczającą liczbę znaków, które są takie same (w kodowaniu SoundEx) w dwóch łańcuchach, gdzie 0 oznacza brak podobieństwa, a 4 oznacza wysokie podobieństwo lub identyczność łańcuchów.

Funkcje daty i czasu

Oprogramowanie CLEM oferuje wiele funkcji zapewniających obsługę zmiennych łańcuchowych reprezentujących daty i godziny. Używane formaty daty i godziny są dostosowane do określonych strumieni; można je ustawić w odpowiednich oknach dialogowych właściwości strumieni. Funkcje daty i czasu analizują łańcuchy dat i godzin zgodnie z aktualnie wybranym formatem.

Jeśli użytkownik określi rok w dacie tylko z użyciem dwóch cyfr (bez informacji o wieku), w oprogramowaniu IBM SPSS Modeler używany jest bieżący wiek ustawiony w oknie dialogowym właściwości strumienia.

Uwaga: Jeśli funkcja daty zostanie przesłana do serwera SQL lub IBM SPSS Analytic Server, w gałęzi następującej po źródle danych Analytic Server wszystkie łańcuchy formatów danych (to_date) w zakresie danej daty muszą odpowiadać formatowi daty określonemu w strumieniu SPSS Modeler.

Tabela 33. Funkcje daty i czasu CLEM		
Funkcja	**Wynik**	**Opis**
@TODAY	*Łańcuch*	Jeśli użytkownik wybierze opcję **Cofnij do dni/godzin** w oknie dialogowym właściwości strumienia, ta funkcja zwraca bieżącą datę jako łańcuch z zastosowaniem bieżącego formatu daty. Jeśli użytkownik korzysta z dwucyfrowego formatu zapisu daty i nie wybierze opcji **Cofnij do dni/godzin**, ta funkcja na bieżącym serwerze zwraca wartość $null$.
to_time(ELEMENT)	*Czas*	Przekształca typ składowania określonej zmiennej na godzinę.
to_date(ELEMENT)	*Date*	Przekształca typ składowania określonej zmiennej na datę.
to_timestamp(ELEMENT)	*Znacznik czasu*	Przekształca typ składowania określonej zmiennej na znacznik czasu.
to_datetime(ELEMENT)	*Datetime*	Przekształca typ składowania określonej zmiennej na datę, godzinę i znacznik czasu.

Tabela 33. Funkcje daty i czasu CLEM (kontynuacja)

Funkcja	Wynik	Opis
`datetime_date(ELEMENT)`	*Date*	Zwraca wartość daty dla *liczby*, *łańcucha* lub *znacznika czasu*. Jest to jedyna funkcja, która pozwala na wsteczne przekształcenie liczby (w sekundach) na datę. Jeśli ELEMENT jest łańcuchem, data jest tworzona poprzez przeanalizowanie łańcucha w bieżącym formacie daty. Format daty opisany w oknie dialogowym właściwości strumienia musi być poprawny, aby ta funkcja zadziałała prawidłowo. Jeśli ELEMENT jest liczbą, jest ona interpretowana jako liczba sekund, które upłynęły od daty podstawowej. Ułamkowe części wartości dni są obcinane. Jeśli ELEMENT jest znacznikiem czasu, zwracana jest część daty znacznika czasu. Jeśli ELEMENT jest datą, jest ona zwracana w niezmienionej postaci.
`date_before(DATA1, DATA2)`	*Boolean*	Zwraca wartość true, jeśli wartość *DATA1* reprezentuje datę lub znacznik czasu przypadające przed datą reprezentowaną przez wartość *DATE2*. W przeciwnym razie ta funkcja zwraca wartość 0.
`date_days_difference(DATA1, DATA2)`	*Liczba całkowita*	Zwraca liczbę całkowitą oznaczającą czas w dniach od daty lub znacznika czasu reprezentowanych przez wartość *DATA1* do wartości reprezentowanej przez wartość *DATA2* w postaci liczby całkowitej. Jeśli wartość *DATA2* przypada przed wartością *DATA1*, ta funkcja zwraca liczbę ujemną.
`date_in_days(DATA)`	*Liczba całkowita*	Zwraca czas w dniach od daty bazowej do danej daty lub znacznika czasu reprezentowanych przez wartość *DATA* w postaci liczby całkowitej. Jeśli wartość *DATA* przypada przed datą bazową, ta funkcja zwraca liczbę ujemną. Poprawne działanie obliczenia jest uzależnione od wprowadzenia prawidłowej daty. Przykładowo: nie należy określać daty 29 lutego 2001. Rok 2001 nie był rokiem przestępnym, więc taka data nie istniała.
`date_in_months(DATA)`	*Liczba rzeczywista*	Zwraca czas w miesiącach od daty bazowej do daty lub znacznika czasu reprezentowanych przez wartość *DATA* w postaci liczby rzeczywistej. Jest to wartość zaokrąglona w oparciu o przyjętą średnią długość miesiąca wynoszącą 30,4375 dnia. Jeśli wartość *DATA* przypada przed datą bazową, ta funkcja zwraca liczbę ujemną. Poprawne działanie obliczenia jest uzależnione od wprowadzenia prawidłowej daty. Przykładowo: nie należy określać daty 29 lutego 2001. Rok 2001 nie był rokiem przestępnym, więc taka data nie istniała.

Tabela 33. Funkcje daty i czasu CLEM (kontynuacja)

Funkcja	Wynik	Opis
`date_in_weeks(DATA)`	*Liczba rzeczywista*	Zwraca czas w tygodniach od daty bazowej do daty lub znacznika czasu reprezentowanych przez wartość *DATA* w postaci liczby rzeczywistej. Jest to wartość określona w oparciu o długość tygodnia wynoszącą 7,0 dni. Jeśli wartość *DATA* przypada przed datą bazową, ta funkcja zwraca liczbę ujemną. Poprawne działanie obliczenia jest uzależnione od wprowadzenia prawidłowej daty. Przykładowo: nie należy określać daty 29 lutego 2001. Rok 2001 nie był rokiem przestępnym, więc taka data nie istniała.
`date_in_years(DATA)`	*Liczba rzeczywista*	Zwraca czas w latach od daty bazowej do daty lub znacznika czasu reprezentowanych przez wartość *DATA* w postaci liczby rzeczywistej. Jest to wartość zaokrąglona w oparciu o przyjętą średnią długość roku wynoszącą 365,25 dnia. Jeśli wartość *DATA* przypada przed datą bazową, ta funkcja zwraca liczbę ujemną. Poprawne działanie obliczenia jest uzależnione od wprowadzenia prawidłowej daty. Przykładowo: nie należy określać daty 29 lutego 2001. Rok 2001 nie był rokiem przestępnym, więc taka data nie istniała.
`date_months_difference (DATA1, DATA2)`	*Liczba rzeczywista*	Zwraca liczbę całkowitą oznaczającą czas w miesiącach od daty lub znacznika czasu reprezentowanych przez wartość *DATA1* do wartości reprezentowanej przez *DATA2* w postaci liczby rzeczywistej. Jest to wartość zaokrąglona w oparciu o przyjętą średnią długość miesiąca wynoszącą 30,4375 dnia. Jeśli wartość *DATA2* przypada przed wartością *DATA1*, ta funkcja zwraca liczbę ujemną.
`datetime_date(ROK, MIESIĄC, DZIEŃ)`	*Date*	Tworzy wartość daty z danego *ROKU*, *MIESIĄCA* i *DNIA*. Argumenty muszą być liczbami całkowitymi.
`datetime_day(DATA)`	*Liczba całkowita*	Zwraca dzień miesiąca na podstawie danej *DATY* lub znacznika czasu. Wynik jest liczbą całkowitą w przedziale od 1 do 31.
`datetime_day_name(DZIEŃ)`	*Łańcuch*	Zwraca pełną nazwę danego *DNIA*. Argument musi być liczbą całkowitą w przedziale od 1 (niedziela) do 7 (sobota).
`datetime_hour(CZAS)`	*Liczba całkowita*	Zwraca godzinę na podstawie *CZASU* lub znacznika czasu. Wynik jest liczbą całkowitą w przedziale od 0 do 23.
`datetime_in_seconds(CZAS)`	*Liczba rzeczywista*	Zwraca sekcję sekund przechowywaną w wartości *CZAS*.
`datetime_in_seconds(DATA)`, `datetime_in_seconds(DATA_CZAS)`	*Liczba rzeczywista*	Zwraca skumulowaną liczbę przekształconą na sekundy, obliczoną jako różnica między bieżącą *DATĄ* lub *DATĄGODZINĄ* i datą bazową (1900-01-01).
`datetime_minute(CZAS)`	*Liczba całkowita*	Zwraca minutę na podstawie *CZASU* lub znacznika czasu. Wynik jest liczbą całkowitą w przedziale od 0 do 59.

Tabela 33. Funkcje daty i czasu CLEM (kontynuacja)

Funkcja	Wynik	Opis
`datetime_month(DATA)`	*Liczba całkowita*	Zwraca miesiąc na podstawie danej *DATY* lub znacznika czasu. Wynik jest liczbą całkowitą w przedziale od 1 do 12.
`datetime_month_name (MIESIĄC)`	*Łańcuch*	Zwraca pełną nazwę danego *MIESIĄCA*. Argument musi być liczbą całkowitą w przedziale od 1 do 12.
`datetime_now`	*Znacznik czasu*	Zwraca aktualny czas jako znacznik czasu.
`datetime_second(CZAS)`	*Liczba całkowita*	Zwraca sekundę na podstawie *CZASU* lub znacznika czasu. Wynik jest liczbą całkowitą w przedziale od 0 do 59.
`datetime_day_short_name (DZIEŃ)`	*Łańcuch*	Zwraca skróconą nazwę danego *DNIA*. Argument musi być liczbą całkowitą w przedziale od 1 (niedziela) do 7 (sobota).
`datetime_month_short_nam e (MIESIĄC)`	*Łańcuch*	Zwraca skróconą nazwę danego *MIESIĄCA*. Argument musi być liczbą całkowitą w przedziale od 1 do 12.
`datetime_time(GODZINA, MINUTA, SEKUNDA)`	*Czas*	Zwraca wartość czasu dla określonej *GODZINY*, *MINUTY* i *SEKUNDY*. Argumenty muszą być liczbami całkowitymi.
`datetime_time(ELEMENT)`	*Czas*	Zwraca wartość czasu danego *ELEMENTU*.
`datetime_timestamp(ROK, MIESIĄC, DZIEŃ, GODZINA, MINUTA, SEKUNDA)`	*Znacznik czasu*	Zwraca wartość znacznika czasu dla danego *ROKU*, *MIESIĄCA*, *DNIA*, *GODZINY*, *MINUTY* i *SEKUNDY*.
`datetime_timestamp(DATA, CZAS)`	*Znacznik czasu*	Zwraca wartość znacznika czasu dla danej *DATY* i *CZASU*.
`datetime_timestamp (NUMER)`	*Znacznik czasu*	Zwraca wartość znacznika czasu dla danej liczby sekund.
`datetime_weekday(DATA)`	*Liczba całkowita*	Zwraca dzień tygodnia na podstawie danej *DATY* lub znacznika czasu.
`datetime_year(DATA)`	*Liczba całkowita*	Zwraca rok na podstawie danej *DATY* lub znacznika czasu. Wynik jest liczbą całkowitą np. 2002.
`date_weeks_difference (DATA1, DATA2)`	*Liczba rzeczywista*	Zwraca liczbę całkowitą oznaczającą czas w tygodniach od daty lub znacznika czasu reprezentowanych przez wartość *DATA1* do wartości reprezentowanej przez *DATA2* w postaci liczby rzeczywistej. Jest to wartość określona w oparciu o długość tygodnia wynoszącą 7,0 dni. Jeśli wartość *DATA2* przypada przed wartością *DATA1*, ta funkcja zwraca liczbę ujemną.
`date_years_difference (DATA1, DATA2)`	*Liczba rzeczywista*	Zwraca liczbę całkowitą oznaczającą czas w latach od daty lub znacznika czasu reprezentowanych przez wartość *DATA1* do wartości reprezentowanej przez *DATA2* w postaci liczby rzeczywistej. Jest to wartość zaokrąglona w oparciu o przyjętą średnią długość roku wynoszącą 365,25 dnia. Jeśli wartość *DATA2* przypada przed wartością *DATA1*, ta funkcja zwraca liczbę ujemną.

Tabela 33. Funkcje daty i czasu CLEM (kontynuacja)

Funkcja	Wynik	Opis
`date_from_ywd(ROK, TYDZIEŃ, DZIEN)`	*Liczba całkowita*	Przekształca rok, tydzień w roku oraz dzień tygodnia na datę, używając normy ISO 8601.
`date_iso_day(DATA)`	*Liczba całkowita*	Zwraca dzień tygodnia z daty, używając normy ISO 8601.
`date_iso_week(DATA)`	*Liczba całkowita*	Zwraca tydzień roku z daty, używając normy ISO 8601.
`date_iso_year(DATA)`	*Liczba całkowita*	Zwraca rok z daty, używając normy ISO 8601.
`time_before(CZAS1, CZAS2)`	*Boolean*	Zwraca wartość true, jeśli wartość *CZAS1* reprezentuje czas lub znacznik czasu przypadające przed czasem reprezentowanym przez wartość *CZAS2*. W przeciwnym razie ta funkcja zwraca wartość 0.
`time_hours_difference (CZAS1, CZAS2)`	*Liczba rzeczywista*	Zwraca liczbę rzeczywistą oznaczającą różnicę wyrażoną w godzinach pomiędzy czasem lub znacznikiem czasu *CZAS1* a *CZAS2*. Jeśli użytkownik wybierze opcję **Cofnij do dni/godzin** w oknie dialogowym właściwości strumienia, jako odniesienie do poprzedniego dnia jest wybierana wyższa wartość *CZAS1*. Jeśli ta opcja nie zostanie wybrana, wyższa wartość *CZAS1* powoduje, że zwracana wartość jest ujemna.
`time_in_hours(CZAS)`	*Liczba rzeczywista*	Zwraca czas w godzinach reprezentowany jako wartość *CZAS* w postaci liczby rzeczywistej. Przykładowo: w przypadku formatu czasu `GGMM` wyrażenie `time_in_hours('0130')` zwraca 1.5. Wartość *CZAS* może odzwierciedlać czas lub znacznik czasu.
`time_in_mins(CZAS)`	*Liczba rzeczywista*	Zwraca czas w minutach reprezentowany jako wartość *CZAS* w postaci liczby rzeczywistej. Wartość *CZAS* może odzwierciedlać czas lub znacznik czasu.
`time_in_secs(CZAS)`	*Liczba całkowita*	Zwraca czas w sekundach reprezentowany jako wartość *CZAS* w postaci liczby całkowitej. Wartość *CZAS* może odzwierciedlać czas lub znacznik czasu.
`time_mins_difference(CZAS1, CZAS2)`	*Liczba rzeczywista*	Zwraca liczbę rzeczywistą oznaczającą wyrażoną w minutach różnicę pomiędzy czasem lub znacznikiem czasu *CZAS1* a *CZAS2*. Jeśli użytkownik wybierze opcję **Cofnij do dni/godzin** w oknie dialogowym właściwości strumienia, jako odniesienie do poprzedniego dnia jest wybierana wyższa wartość *CZAS1* (lub poprzedniej godziny, jeśli minuty i sekundy określono w bieżącym formacie). Jeśli ta opcja nie zostanie wybrana, wyższa wartość *CZAS1* powoduje, że zwracana wartość jest ujemna.

Tabela 33. Funkcje daty i czasu CLEM (kontynuacja)		
Funkcja	**Wynik**	**Opis**
`time_secs_difference(CZAS1, CZAS2)`	*Liczba całkowita*	Zwraca liczbę całkowitą oznaczającą wyrażoną w sekundach różnicę pomiędzy czasami lub znacznikami czasu reprezentowaną przez wartości *CZAS1* i *CZAS2*. Jeśli użytkownik wybierze opcję **Cofnij do dni/godzin** w oknie dialogowym właściwości strumienia, jako odniesienie do poprzedniego dnia jest wybierana wyższa wartość *CZAS1* (lub poprzedniej godziny, jeśli minuty i sekundy określono w bieżącym formacie). Jeśli ta opcja nie zostanie wybrana, wyższa wartość *CZAS1* powoduje, że zwracana wartość jest ujemna.

Konwertowanie wartości daty i czasu

Funkcje przekształcania (i inne funkcje wymagające określonego typu danych wejściowych, np. wartość daty lub czasu) są uzależnione od bieżących formatów określonych w oknie dialogowym Opcje strumienia. Przykładowo: jeśli istnieje zmienna o nazwie *DATA* i jest ona przechowywana jako łańcuch z wartościami *Sty 2003*, *Lut 2003* itd., w następujący sposób można wykonać jej przekształcenie do postaci składowania daty:

```
to_date(DATA)
```

Aby operacja przekształcenia została wykonana prawidłowo, jako domyślny format daty strumienia należy wybrać odpowiedni format daty **MIE RRRR**.

Przykładowa operacja przekształcenia wartości łańcucha na daty za pomocą węzła wypełniania znajduje się w strumieniu *broadband_create_models.str* zainstalowanym w folderze *Demos* w podfolderze *streams*.

Daty przechowywane jako liczby. *DATA* w poprzednim przypadku jest nazwą zmiennej, a `to_date` to funkcja CLEM. Jeśli daty są przechowywane jako liczby, można przeliczyć je za pomocą funkcji `datetime_date`, która interpretuje liczbę jako liczbę sekund od daty podstawowej.

```
datetime_date(DATA)
```

Po przeliczeniu daty na liczbę sekund (i odwrotnie) można wykonać obliczenia, takie jak określenie bieżącej daty minus lub plus stała liczba dni, np.:

```
datetime_date((date_in_days(DATA)-7)*60*60*24)
```

Funkcje sekwencji

W przypadku niektórych operacji istotna jest kolejność zdarzeń. Aplikacja umożliwia pracę z następującymi sekwencjami rekordów:

- Sekwencje i szeregi czasowe
- Funkcje sekwencji
- Indeksowanie rekordów
- Uśrednianie, sumowanie i porównywanie wartości
- Monitorowanie zmian – różnicowanie
- `@SINCE`
- Wartości przesunięcia
- Dodatkowe narzędzia kolejności

W wielu przypadkach każdy rekord w strumieniu można uznać za pojedynczą obserwację, niezależną od wszystkich innych. Zazwyczaj w takich sytuacjach kolejność rekordów nie jest ważna.

Jednak w niektórych klasach problemów kolejność rekordów jest niezwykle istotna. Dotyczy to głównie szeregów czasowych, w przypadku których kolejność rekordów odzwierciedla uszeregowaną sekwencję zdarzeń lub wystąpień. Każdy rekord to obraz stanu w danej chwili. Wiele najistotniejszych informacji może jednak być zawartych nie w wartościach chwilowych, ale w sposobie, w jaki te wartości się zmieniają i zachowują w czasie.

Oczywiście kluczowym parametrem nie zawsze musi być czas. Przykładowo: rekordy mogą odzwierciedlać analizy wykonane w określonych odległościach na linii.

Funkcje specjalne i dotyczące kolejności można rozpoznać dzięki następującym wyznacznikom:

- Wszystkie rozpoczynają się od znaku `@`.
- Ich nazwy są zapisane wielkimi literami.

Funkcje sekwencji mogą się odnosić do rekordu aktualnie przetwarzanego przez węzeł, rekordów, które przeszły przez węzeł, a nawet (w jednym przypadku) rekordów, które dopiero przejdą przez węzeł. Funkcje takie można swobodnie łączyć z innymi składnikami wyrażeń CLEM, choć niektóre z nich są obwarowane ograniczeniami w odniesieniu do argumentów, które można zastosować.

Przykłady

Niekiedy chcemy uzyskać informację na temat czasu, który upłynął od danego zdarzenia lub spełnienia warunku. Tych informacji może dostarczyć funkcja `@SINCE`, np.:

```
@SINCE(Income > Outgoings)
```

Ta funkcja zwraca przesunięcie ostatniego rekordu, w przypadku którego ten warunek był prawdziwy, czyli liczbę rekordów, przed niniejszym, w przypadku których warunek był prawdziwy. Jeśli warunek nigdy nie był prawdziwy, funkcja `@SINCE` zwraca `@INDEX + 1`.

Niekiedy użytkownik może chcieć się odnieść do wartości bieżącego rekordu w wyrażeniu użytym przez `@SINCE`. Tę operację można wykonać za pomocą funkcji `@THIS`, która określa, że nazwa zmiennej zawsze odnosi się do bieżącego rekordu. Aby określić przesunięcie ostatniego rekordu, którego wartość zmiennej `Koncentracja` była ponad dwukrotnie większa niż w przypadku bieżącego rekordu, można użyć funkcji:

```
@SINCE(Koncentracja > 2 * @THIS(Koncentracja))
```

Niekiedy warunek dodany do `@SINCE` jest z definicji prawdziwy w przypadku bieżącego rekordu, np.:

```
@SINCE(ID == @THIS(ID))
```

Z tego względu funkcja `@SINCE` nie ocenia warunku w przypadku bieżącego rekordu. Należy użyć podobnej funkcji (`@SINCE0`), aby ocenić warunek w przypadku bieżącego rekordu oraz poprzednich rekordów; jeśli warunek jest prawdziwy w przypadku bieżącego rekordu, funkcja `@SINCE0` zwraca wartość `0`.

Tabela 34. Funkcje sekwencji CLEM

Funkcja	Wynik	Opis
`MEAN(ZMIENNA)`	*Liczba rzeczywista*	Zwraca średnią wartości dla określonej *ZMIENNEJ* lub *ZMIENNYCH*.

Tabela 34. Funkcje sekwencji CLEM (kontynuacja)

Funkcja	Wynik	Opis
`@MEAN(ZMIENNA, WYR)`	*Liczba rzeczywista*	Zwraca średnią wartości dla *ZMIENNEJ* wobec ostatnich rekordów *WYR* odebranych przez bieżący węzeł, w tym bieżący rekord. *ZMIENNA* musi być nazwą zmiennej numerycznej. *WYR* może być wyrażeniem rozwijanym do liczby całkowitej większej niż 0. Jeśli *WYR* zostanie pominięte lub jeśli przekracza liczbę odebranych do tej pory rekordów, zwracana jest średnia ze wszystkich rekordów otrzymanych do tej pory.
`@MEAN(ZMIENNA, WYR, LCAŁK)`	*Liczba rzeczywista*	Zwraca średnią wartości dla *ZMIENNEJ* wobec ostatnich rekordów *WYR* odebranych przez bieżący węzeł, w tym bieżący rekord. *ZMIENNA* musi być nazwą zmiennej numerycznej. *WYR* może być wyrażeniem rozwijanym do liczby całkowitej większej niż 0. Jeśli *WYR* zostanie pominięte lub jeśli przekracza liczbę odebranych do tej pory rekordów, zwracana jest średnia ze wszystkich rekordów otrzymanych do tej pory. *LCAŁK* określa maksymalną liczbę wartości sprawdzanych w tył. Takie rozwiązanie jest znacznie bardziej wydajne niż korzystanie z dwóch argumentów.
`@DIFF1(ZMIENNA)`	*Liczba rzeczywista*	Zwraca pierwszą różniczkę *ZMIENNEJ*. Jednoargumentowa postać zwraca różnicę pomiędzy bieżącą wartością a poprzednią wartością zmiennej. Zwraca wartość **$null$**, jeśli prawidłowe, poprzednie rekordy nie istnieją.
`@DIFF1(ZMIENNA1, ZMIENNA2)`	*Liczba rzeczywista*	Postać dwuargumentowa zwraca pierwszą różniczkę *ZMIENNEJ1* względem *ZMIENNEJ2*. Zwraca wartość **$null$**, jeśli prawidłowe, poprzednie rekordy nie istnieją. Obliczenie jest wykonywane jako `@DIFF1(ZMIENNA1)/@DIFF1(ZMIENNA2)`.
`@DIFF2(ZMIENNA)`	*Liczba rzeczywista*	Zwraca drugą różniczkę *ZMIENNEJ*. Jednoargumentowa postać zwraca różnicę pomiędzy bieżącą wartością a poprzednią wartością zmiennej. Zwraca wartość **$null$**, jeśli prawidłowe, poprzednie rekordy nie istnieją. `@DIFF2` jest obliczana jako `@DIFF(@DIFF(ZMIENNA))`.
`@DIFF2(ZMIENNA1, ZMIENNA2)`	*Liczba rzeczywista*	Postać dwuargumentowa zwraca drugą różniczkę *ZMIENNEJ1* względem *ZMIENNEJ2*. Zwraca wartość **$null$**, jeśli prawidłowe, poprzednie rekordy nie istnieją. Jest to obliczenie złożone — `@DIFF1(ZMIENNA1)/@DIFF1(ZMIENNA2) - @OFFSET(@DIFF1(ZMIENNA1),1)/ @OFFSET(@DIFF1(ZMIENNA2))) / @DIFF1(ZMIENNA2)`.
`@INDEX`	*Liczba całkowita*	Zwraca indeks bieżącego rekordu. Indeksy są alokowane w rekordach zgodnie z ich przybyciem w bieżącym węźle. Pierwszemu rekordowi jest nadawany indeks 1, a numery indeksów w przypadku wszystkich kolejnych rekordów zwiększają się o 1.

Tabela 34. Funkcje sekwencji CLEM (kontynuacja)

Funkcja	Wynik	Opis
`@LAST_NON_BLANK(ZMIENNA)`	*Any*	Zwraca ostatnią wartość *ZMIENNEJ*, która nie była pusta, zgodnie z węzłem źródłowym lub wprowadzania danych. Jeśli nie istnieją niepuste wartości *ZMIENNEJ* w odczytanych do tej pory rekordach, zwracana jest wartość `$null$`. Wartości puste (nazywane również brakującymi wartościami użytkownika) można definiować oddzielnie dla każdej zmiennej.
`@MAX(ZMIENNA)`	*Liczba*	Zwraca wartość maksymalną określonej *ZMIENNEJ*.
`@MAX(ZMIENNA, WYR)`	*Liczba*	Zwraca maksymalną wartość *ZMIENNEJ* w ostatnich rekordach *WYR* odebranych do tej pory z uwzględnieniem bieżącego rekordu. *ZMIENNA* musi być nazwą zmiennej numerycznej. *WYR* może być dowolnym wyrażeniem rozwijanym do liczby całkowitej większej niż 0.
`@MAX(ZMIENNA, WYR, LCAŁK)`	*Liczba*	Zwraca maksymalną wartość *ZMIENNEJ* w ostatnich rekordach *WYR* odebranych do tej pory z uwzględnieniem bieżącego rekordu. *ZMIENNA* musi być nazwą zmiennej numerycznej. *WYR* może być wyrażeniem rozwijanym do liczby całkowitej większej niż 0. Jeśli *WYR* zostanie pominięte lub jeśli przekracza liczbę odebranych do tej pory rekordów, zwracana jest wartość maksymalna ze wszystkich rekordów otrzymanych do tej pory. *LCAŁK* określa maksymalną liczbę wartości sprawdzanych w tył. Takie rozwiązanie jest znacznie bardziej wydajne niż korzystanie z dwóch argumentów.
`@MIN(ZMIENNA)`	*Liczba*	Zwraca wartość minimalną określonej *ZMIENNEJ*.
`@MIN(ZMIENNA, WYR)`	*Liczba*	Zwraca minimalną wartość dla *ZMIENNEJ* wobec ostatnich rekordów *WYR* odebranych do tej pory z uwzględnieniem bieżącego rekordu. *ZMIENNA* musi być nazwą zmiennej numerycznej. *WYR* może być dowolnym wyrażeniem rozwijanym do liczby całkowitej większej niż 0.
`@MIN(ZMIENNA, WYR, LCAŁK)`	*Liczba*	Zwraca minimalną wartość dla *ZMIENNEJ* wobec ostatnich rekordów *WYR* odebranych do tej pory z uwzględnieniem bieżącego rekordu. *ZMIENNA* musi być nazwą zmiennej numerycznej. *WYR* może być wyrażeniem rozwijanym do liczby całkowitej większej niż 0. Jeśli *WYR* zostanie pominięte lub jeśli przekracza liczbę odebranych do tej pory rekordów, zwracana jest wartość minimalna ze wszystkich rekordów otrzymanych do tej pory. *LCAŁK* określa maksymalną liczbę wartości sprawdzanych w tył. Takie rozwiązanie jest znacznie bardziej wydajne niż korzystanie z dwóch argumentów.

Tabela 34. Funkcje sekwencji CLEM (kontynuacja)

Funkcja	Wynik	Opis
`@OFFSET(ZMIENNA, WYR)`	*Any*	Zwraca wartość *ZMIENNEJ* w rekordzie przesuniętym względem bieżącego rekordu o wartość *WYR*. Przesunięcie dodatnie odnosi się do rekordu, który już przeszedł ("spojrzenie w tył"), a przesunięcie ujemne określa "spojrzenie w przód" względem rekordu, który dopiero nadejdzie. Przykładowo: funkcja `@OFFSET(Status, 1)` zwraca wartość zmiennej `Status` w poprzednim rekordzie, natomiast `@OFFSET(Status, –4)` "spogląda w przód" w zakresie czterech rekordów w sekwencji (czyli rekordów, które jeszcze nie przeszły przez ten węzeł). *Przesunięcie ujemne ("spojrzenie w przód") należy określić jako stałą.* Tylko w przypadku przesunięć dodatnich *WYR* może również być dowolnym wyrażeniem CLEM, które jest obliczane dla bieżącego rekordu w celu określenia przesunięcia. W takim przypadku trzyargumentowa wersja tej funkcji powinna zwiększyć wydajność działania (patrz następna funkcja). Jeśli wyrażenie zwraca wynik inny niż nieujemna liczba całkowita, wyświetlany jest błąd. Oznacza to, że nie można uzyskać obliczenia przesunięcia w przód. *Uwag*: w funkcji `@OFFSET` odnoszącej się do samej siebie nie można używać literalnego sprawdzania w przód. Przykładowo: w węźle wypełniania nie można zastąpić wartości `zmienna1` za pomocą wyrażenia, takiego jak `@OFFSET(zmienna1,-2)`. *Uwaga*: w węźle wypełniania podczas wypełniania zmiennej istnieją dwie różne wartości tej zmiennej: wartość wstępnie wypełniona i wartość po wypełnieniu. Gdy funkcja `@OFFSET` odnosi się do samej siebie, odnosi się ona do wartości po wypełnieniu. Wartość po wypełnieniu istnieje tylko w przypadku minionych wierszy, więc odnosząca się do samej siebie funkcja `@OFFSET` może się odnosić tylko do minionych wierszy. Odnosząca się do samej siebie funkcja `@OFFSET` nie może się odnosić do przyszłości, więc wykonywane są następujące kontrole przesunięcia: • Jeśli przesunięcie jest literalne i wykonane w kierunku przyszłości, przed rozpoczęciem wykonywania zgłaszany jest błąd. • Jeśli przesunięcie jest wyrażeniem i rozwija się w kierunku przyszłości w środowisku wykonawczym, funkcja `@OFFSET` zwraca wartość `$null$`. *Uwaga*: jednoczesne używanie metod "spojrzenia w przód" i "spojrzenia w tył" w jednym węźle nie jest obsługiwane.

Tabela 34. Funkcje sekwencji CLEM (kontynuacja)

Funkcja	Wynik	Opis
`@OFFSET(ZMIENNA, WYR, LCAŁK)`	*Any*	Wykonuje taką samą operację jak funkcja `@OFFSET`, jednak dodawany jest trzeci argument (*LCAŁK*), który określa maksymalną liczbę wartości do przeszukania w tył. W sytuacjach kiedy przesunięcie jest obliczane na podstawie wyrażenia, dodanie trzeciego argumentu powinno zwiększyć wydajność. Przykładowo: w wyrażeniu, takim jak `@OFFSET(Foo, Month, 12)` system wie, że należy zachować tylko ostatnich dwanaście wartości Foo; w przeciwnym razie na wszelki wypadek przechowywane są wszystkie wartości. W sytuacjach kiedy wartość przesunięcia jest stałą (dotyczy to też ujemnych przesunięć w przód, które muszą być wartościami stałymi), stosowanie trzeciego argumentu nie ma uzasadnienia i należy stosować dwuargumentową wersję tej funkcji. Użytkownik powinien również zapoznać się z funkcjami odnoszącymi się do samych siebie w wersji dwuargumentowej opisanej wcześniej. *Uwaga*: jednoczesne używanie metod "spojrzenia w przód" i "spojrzenia w tył" w jednym węźle nie jest obsługiwane.
`@SDEV(ZMIENNA)`	*Liczba rzeczywista*	Zwraca odchylenie standardowe wartości dla określonej *ZMIENNEJ* lub *ZMIENNYCH*.
`@SDEV(ZMIENNA, WYR)`	*Liczba rzeczywista*	Zwraca odchylenie standardowe wartości *ZMIENNEJ* z ostatnich rekordów *WYR* odebranych przez bieżący węzeł z uwzględnieniem bieżącego rekordu. *ZMIENNA* musi być nazwą zmiennej numerycznej. *WYR* może być wyrażeniem rozwijanym do liczby całkowitej większej niż 0. Jeśli *WYR* zostanie pominięte lub jeśli przekracza liczbę odebranych do tej pory rekordów, zwracane jest odchylenie standardowe ze wszystkich rekordów otrzymanych do tej pory.
`@SDEV(ZMIENNA, WYR, LCAŁK)`	*Liczba rzeczywista*	Zwraca odchylenie standardowe wartości *ZMIENNEJ* z ostatnich rekordów *WYR* odebranych przez bieżący węzeł z uwzględnieniem bieżącego rekordu. *ZMIENNA* musi być nazwą zmiennej numerycznej. *WYR* może być wyrażeniem rozwijanym do liczby całkowitej większej niż 0. Jeśli *WYR* zostanie pominięte lub jeśli przekracza liczbę odebranych do tej pory rekordów, zwracane jest odchylenie standardowe ze wszystkich rekordów otrzymanych do tej pory. *LCAŁK* określa maksymalną liczbę wartości sprawdzanych w tył. Takie rozwiązanie jest znacznie bardziej wydajne niż korzystanie z dwóch argumentów.
`@SINCE(WYR)`	*Any*	Zwraca liczbę rekordów minionych od czasu, kiedy *WYR*, wyrażenie dowolne CLEM, było prawdziwe.

Tabela 34. Funkcje sekwencji CLEM (kontynuacja)		
Funkcja	**Wynik**	**Opis**
`@SINCE(WYR, LCAŁK)`	*Any*	Dodanie drugiego argumentu (*LCAŁK*) określa maksymalną liczbę rekordów, które są badane w tył. Jeśli *WYR* nigdy nie było prawdziwe, funkcją *LCAŁK* jest `@INDEX+1`.
`@SINCE0(WYR)`	*Any*	Uwzględnia bieżący rekord, natomiast funkcja `@SINCE` nie uwzględnia go. Funkcja `@SINCE0` zwraca 0, jeśli *WYR* jest prawdziwe dla bieżącego rekordu.
`@SINCE0(WYR, LCAŁK)`	*Any*	Dodanie drugiego argumentu (*LCAŁK*) określa maksymalną liczbę rekordów, które są badane w tył.
`@SUM(ZMIENNA)`	*Liczba*	Zwraca sumę wartości dla określonej *ZMIENNEJ* lub *ZMIENNYCH*.
`@SUM(ZMIENNA, WYR)`	*Liczba*	Zwraca sumę wartości dla *ZMIENNEJ* wobec ostatnich rekordów *WYR* odebranych przez bieżący węzeł z uwzględnieniem bieżącego rekordu. *ZMIENNA* musi być nazwą zmiennej numerycznej. *WYR* może być wyrażeniem rozwijanym do liczby całkowitej większej niż 0. Jeśli *WYR* zostanie pominięte lub jeśli przekracza liczbę odebranych do tej pory rekordów, zwracana jest suma wszystkich rekordów otrzymanych do tej pory.
`@SUM(ZMIENNA, WYR, LCAŁK)`	*Liczba*	Zwraca sumę wartości dla *ZMIENNEJ* wobec ostatnich rekordów *WYR* odebranych przez bieżący węzeł z uwzględnieniem bieżącego rekordu. *ZMIENNA* musi być nazwą zmiennej numerycznej. *WYR* może być wyrażeniem rozwijanym do liczby całkowitej większej niż 0. Jeśli *WYR* zostanie pominięte lub jeśli przekracza liczbę odebranych do tej pory rekordów, zwracana jest suma wszystkich rekordów otrzymanych do tej pory. *LCAŁK* określa maksymalną liczbę wartości sprawdzanych w tył. Takie rozwiązanie jest znacznie bardziej wydajne niż korzystanie z dwóch argumentów.
`@THIS(ZMIENNA)`	*Any*	Zwraca wartość zmiennej o nazwie *ZMIENNA* w bieżącym rekordzie. Używane tylko w wyrażeniach `@SINCE`.

Funkcje globalne

Funkcje `@MEAN`, `@SUM`, `@MIN`, `@MAX` i `@SDEV` operują na (maksymalnie) wszystkich rekordach odczytanych do bieżącego rekordu włącznie. Jednak niekiedy warto mieć możliwość określenia, w jaki sposób wartości w bieżącym rekordzie prezentują się względem całego zbioru danych. Używając węzła ustawień globalnych w celu tworzenia wartości w całym zbiorze danych, można uzyskać dostęp do tych wartości w wyrażeniu CLEM dzięki zastosowaniu funkcji globalnych.

Na przykład

```
@GLOBAL_MAX(Wiek)
```

zwraca najwyższą wartość `Wiek` dostępną w zbiorze danych, natomiast wyrażenie

```
(Value - @GLOBAL_MEAN(Wartość)) / @GLOBAL_SDEV(Wartość)
```

wyraża różnicę między `Wartością` rekordu i średnią globalną jako krotność odchyleń standardowych. Wartości globalnych można używać dopiero po obliczeniu ich za pomocą węzła ustawień globalnych. Wszystkie bieżące wartości globalne można anulować, klikając przycisk **Wyczyść wartości globalne** dostępny na karcie Globalne w oknie dialogowym właściwości strumienia.

Tabela 35. Funkcje globalne CLEM

Funkcja	Wynik	Opis
`@GLOBAL_MAX(ZMIENNA)`	*Liczba*	Zwraca maksymalną wartość *ZMIENNEJ* w całym zbiorze danych uprzednio utworzonym przez węzeł ustawień globalnych. *ZMIENNA* musi być nazwą zmiennej numerycznej, daty/czasu lub łańcuchową. Jeśli nie określono odpowiedniej wartości globalnej, zostanie wyświetlony błąd.
`@GLOBAL_MIN(ZMIENNA)`	*Liczba*	Zwraca minimalną wartość *ZMIENNEJ* w całym zbiorze danych, uprzednio utworzonym przez węzeł ustawień globalnych. *ZMIENNA* musi być nazwą zmiennej numerycznej, daty/czasu lub łańcuchową. Jeśli nie określono odpowiedniej wartości globalnej, zostanie wyświetlony błąd.
`@GLOBAL_SDEV(ZMIENNA)`	*Liczba*	Zwraca odchylenie standardowe wartości dla *ZMIENNEJ* z całego zbioru danych, uprzednio utworzone przez węzeł ustawień globalnych. *ZMIENNA* musi być nazwą zmiennej numerycznej. Jeśli nie określono odpowiedniej wartości globalnej, zostanie wyświetlony błąd.
`@GLOBAL_MEAN(ZMIENNA)`	*Liczba*	Zwraca średnią wartości dla *ZMIENNEJ* z całego zbioru danych, uprzednio utworzoną przez węzeł ustawień globalnych. *ZMIENNA* musi być nazwą zmiennej numerycznej. Jeśli nie określono odpowiedniej wartości globalnej, zostanie wyświetlony błąd.
`@GLOBAL_SUM(ZMIENNA)`	*Liczba*	Zwraca sumę wartości dla *ZMIENNEJ* z całego zbioru danych, uprzednio utworzoną przez węzeł ustawień globalnych. *ZMIENNA* musi być nazwą zmiennej numerycznej. Jeśli nie określono odpowiedniej wartości globalnej, zostanie wyświetlony błąd.

Funkcje obsługujące wartości puste i null

Za pomocą języka CLEM można określić, że dane wartości w zmiennej są uznawane za "puste" lub brakujące. Poniższe funkcje służą do obsługi wartości pustych.

Tabela 36. CLEM - funkcje do obsługi wartości pustych i null

Funkcja	Wynik	Opis
`@BLANK(ZMIENNA)`	*Boolean*	Zwraca wartość true w przypadku wszystkich rekordów, których wartości są puste zgodnie z regułami obsługi wartości pustych określonymi w poprzedzającym węźle Typ węzła źródłowego (karta Typy).

Tabela 36. CLEM - funkcje do obsługi wartości pustych i null (kontynuacja)		
Funkcja	**Wynik**	**Opis**
`@LAST_NON_BLANK(ZMIENNA)`	*Any*	Zwraca ostatnią wartość *ZMIENNEJ*, która nie była pusta, zgodnie z węzłem źródłowym lub wprowadzania danych. Jeśli nie istnieją niepuste wartości *ZMIENNEJ* w odczytanych do tej pory rekordach, zwracana jest wartość `$null$`. Wartości puste (nazywane również brakującymi wartościami użytkownika) można definiować oddzielnie dla każdej zmiennej.
`@NULL(ZMIENNA)`	*Boolean*	Zwraca wartość true, jeżeli wartość *ZMIENNEJ* jest systemowym brakiem danych `$null$`. Zwraca wartość false w przypadku wszystkich pozostałych wartości w tym wartości pustych zdefiniowanych przez użytkownika. Aby skontrolować obydwie opcje, należy użyć funkcji `@BLANK(ZMIENNA)` i `@NULL(ZMIENNA)`.
`undef`	*Any*	Ta funkcja jest używana głównie w CLEM w celu wprowadzania wartości `$null$`, np. w celu wypełniania wartości pustych wartościami null w węźle wypełniania.

Puste zmienne można "wypełnić" w węźle wypełniania. W węzłach wypełniania i wyliczania (tylko wiele trybów) specjalna funkcja CLEM `@FIELD` odnosi się do aktualnie badanych zmiennych.

Funkcje specjalne

Funkcje specjalne służą do określania konkretnych badanych zmiennych lub do generowania listy zmiennych wejściowych. Przykładowo: aby jednocześnie wyliczyć wiele zmiennych, należy użyć funkcji `@FIELD` w celu wydania polecenia "wykonaj tę operację na wybranych zmiennych”. Wyrażenie `log(@FIELD)` powoduje wyznaczenie nowej zmiennej log dla każdej wybranej zmiennej.

Tabela 37. Zmienne specjalne CLEM		
Funkcja	**Wynik**	**Opis**
`@FIELD`	*Any*	Wykonuje operację na wszystkich zmiennych określonych w kontekście wyrażenia.
`@TARGET`	*Any*	Jeśli w funkcji analizy zdefiniowanej przez użytkownika jest używane wyrażenie CLEM, funkcja `@TARGET` odzwierciedla zmienną przewidywaną lub "prawidłową wartość" docelowej/przewidywanej pary poddawanej analizie. Ta funkcja jest często używana w węźle analizy.

Tabela 37. Zmienne specjalne CLEM (kontynuacja)

Funkcja	Wynik	Opis
`@PREDICTED`	*Any*	Jeśli wyrażenie CLEM jest używane w funkcji analizy zdefiniowanej przez użytkownika, funkcja `@PREDICTED` odzwierciedla wartość przewidywaną dla analizowanej pary docelowej/przewidywanej. Ta funkcja jest często używana w węźle analizy.
`@PARTITION_FIELD`	*Any*	Zastępuje nazwę bieżącej zmiennej dzielącej na podzbiory.
`@TRAINING_PARTITION`	*Any*	Zwraca wartość bieżącego podzbioru uczącego. Przykładowo: aby wybrać rekordy uczące za pomocą węzła wyboru, należy użyć wyrażenia CLEM: `@PARTITION_FIELD = @TRAINING_PARTITION` Dzięki temu węzeł wyboru będzie zawsze działał bez względu na wartości używane do reprezentowania podzbiorów w danych.
`@TESTING_PARTITION`	*Any*	Zwraca wartość bieżącego podzbioru testowego.
`@VALIDATION_PARTITION`	*Any*	Zwraca wartość bieżącego podzbioru walidacyjnego.
`@FIELDS_BETWEEN(początek, koniec)`	*Any*	Zwraca listę nazw zmiennych w zakresie pomiędzy określonymi zmiennymi początkowymi i końcowymi (włącznie) w oparciu o naturalny (zgodnie ze wstawianiem) porządek zmiennych w danych.

Tabela 37. Zmienne specjalne CLEM (kontynuacja)

Funkcja	Wynik	Opis
`@FIELDS_MATCHING(wzorzec)`	*Any*	Zwraca listę nazw zmiennych odpowiadającą określonemu wzorcowi. Znak zapytania (?) można uwzględnić we wzorcu. Ten znak odpowiada dokładnie jednemu znakowi. Znak gwiazdki (*) odpowiada zerowej lub dowolnej liczbie znaków. Aby wyszukać rzeczywisty znak zapytania lub gwiazdki (a nie używać tych znaków jako symboli wieloznacznych), należy je poprzedzić znakiem ukośnika odwrotnego (\). **Uwaga:** Argument musi być literałem łańcuchowym; niedozwolone jest użycie zagnieżdżonego wyrażenia do wygenerowania argumentu.
`@MULTI_RESPONSE_SET`	*Any*	Zwraca listę zmiennych w nazwanym zestawie wielokrotnych odpowiedzi.

Rozdział 11. Używanie programu IBM SPSS Modeler razem z repozytorium

Informacje o programie IBM SPSS Collaboration and Deployment Services Repository

Z programu SPSS Modeler można korzystać w połączeniu z repozytorium IBM SPSS Collaboration and Deployment Services, co pozwala zarządzać cyklem życia modeli eksploracji danych i powiązanych obiektów predykcyjnych, dzięki czemu obiekty te mogą być używane przez aplikacje, narzędzia i inne rozwiązania stosowane w przedsiębiorstwie. Do obiektów IBM SPSS Modeler, które można udostępniać w taki sposób, należą strumienie, węzły, dane wyjściowe strumieni, projekty i modele. Obiekty są składowane w repozytorium centralnym, skąd mogą być udostępniane innym aplikacjom i śledzone za pośrednictwem rozszerzonych możliwości zarządzania wersjami, metadanych i funkcji wyszukiwania.

Przed użyciem programu SPSS Modeler wraz z repozytorium konieczne jest zainstalowanie adaptera na hoście repozytorium. W przeciwnym razie podczas próby uzyskania dostępu do obiektów repozytorium z różnych węzłów lub modeli SPSS Modeler może zostać wyświetlony następujący komunikat:

```
Obsługa nowych węzłów, modeli oraz typów danych wynikowych może wymagać aktualizacji
Repozytorium.
```

Instrukcje instalowania adaptera zawiera publikacja *SPSS Modeler — instalacja składnika Deployment Adapter* dostępna jako plik PDF w pobranym pakiecie z produktem. Szczegóły na temat sposobu uzyskania dostępu do obiektów repozytorium IBM SPSS Modeler z aplikacji IBM SPSS Deployment Manager zostały omówione w publikacji *SPSS Modeler — podręcznik wdrażania.*

W poniższych sekcjach przedstawiono informacje na temat uzyskiwania dostępu do repozytorium za pośrednictwem aplikacji SPSS Modeler.

Rozszerzona obsługa wersji i wyszukiwania

Repozytorium udostępnia kompleksowe możliwości zarządzania wersjami i wyszukiwania. Załóżmy na przykład, że utworzono strumień i zapisano go w repozytorium, z którego może być udostępniany badaczom z innych oddziałów. Jeśli strumień w późniejszym czasie zostanie zaktualizowany w aplikacji SPSS Modeler, można dodać zaktualizowaną wersję do repozytorium bez nadpisywania poprzedniej wersji. Wszystkie wersje pozostają dostępne i można je wyszukiwać według użytej nazwy, etykiety, pól lub innych atrybutów. Można na przykład wyszukać wszystkie wersje modelu, w których wartością wejściową jest przychód netto lub wszystkie modele utworzone przez konkretną osobę. (W tradycyjnym systemie plików konieczne byłoby wyszukanie poszczególnych wersji z różną nazwą pliku, a relacje pomiędzy wersjami pozostałyby nieznane dla oprogramowania).

Pojedyncze uwierzytelnianie

Funkcja pojedynczego uwierzytelniania umożliwia użytkownikom połączenie się z repozytorium bez konieczności wprowadzania za każdym razem szczegółów nazwy użytkownika i hasła. Istniejące szczegóły logowania użytkownika do sieci lokalnej zapewniają uwierzytelnianie wymagane przez program IBM SPSS Collaboration and Deployment Services. Funkcja ta zależy od następujących czynników:

- Program IBM SPSS Collaboration and Deployment Services musi być skonfigurowany do korzystania z dostawcy pojedynczego uwierzytelniania.
- Użytkownik musi być zalogowany do hosta, który jest kompatybilny z dostawcą.

Aby uzyskać więcej informacji, zobacz “Łączenie się z repozytorium” na stronie 204.

Przechowywanie i wdrażanie obiektów repozytorium

Strumienie tworzone w oprogramowaniu IBM SPSS Modeler można **przechowywać** w repozytorium w taki sam sposób jak pliki. Rozszerzeniem strumieni jest `.str`. Dzięki takiemu rozwiązaniu dostęp do jednego strumienia może uzyskać wielu użytkowników w całym przedsiębiorstwie. Więcej informacji można znaleźć w temacie “Zapisywanie obiektów w repozytorium” na stronie 205.

Istnieje również możliwość wdrożenia strumienia w repozytorium. Taki strumień jest przechowywany jako plik wraz z dodatkowymi metadanymi. Wdrożony strumień może w pełni korzystać z funkcji biznesowych dostępnych w oprogramowaniu IBM SPSS Collaboration and Deployment Services, takich jak automatyczne ocenianie i odświeżanie modelu. Przykładowo: model może być automatycznie, regularnie aktualizowany, gdy dostępne są nowe dane. Można również wdrożyć zestaw strumieni do analizy Champion Challenger, w ramach której strumienie są porównywane w celu określenia, który z nich zawiera najbardziej efektywny model predykcyjny.

Uwaga: Reguły asocjacyjne, węzły modelowania TCM i STP nie obsługują oceny modelu ani kroków Champion Challenger w oprogramowaniu IBM SPSS Collaboration and Deployment Services.

Strumień (z rozszerzeniem `.str`) można wdrożyć. Wdrożenie jako strumień pozwala na używanie tego strumienia przez uproszczoną aplikację kliencką IBM SPSS Modeler Advantage. Więcej informacji można znaleźć w temacie “Otwieranie strumienia w programie IBM SPSS Modeler Advantage ” na stronie 223.

Aby uzyskać więcej informacji, zobacz “Opcje wdrażania strumieni” na stronie 216.

Inne opcje wdrażania

Oprogramowanie IBM SPSS Collaboration and Deployment Services zapewnia szeroki wachlarz funkcji służących do zarządzania zawartościami dostępnymi w przedsiębiorstwie oraz wiele innych mechanizmów pozwalających na wdrażanie lub eksportowanie strumieni, takich jak na przykład:

- Funkcje eksportowania strumienia i modelu do wykorzystania w przyszłości z użyciem środowiska wykonawczego IBM SPSS Modeler Solution Publisher.
- Funkcje eksportowania modeli w formacie PMML. Jest to format oparty na języku XML i służy do kodowania informacji o modelu. Więcej informacji można znaleźć w temacie “Importowanie i eksportowanie modeli w formacie PMML” na stronie 224.

Łączenie się z repozytorium

1. Aby połączyć się z repozytorium, w menu głównym programu IBM SPSS Modeler, kliknij:

   Narzędzia > **Repozytorium** > **Opcje...**
2. W polu adresu URL repozytorium wprowadź lub wybierz ścieżkę do katalogu lub adres URL instalacji repozytorium, do którego zamierzasz uzyskać dostęp. Można połączyć się tylko z jednym repozytorium naraz.

   Ustawienia są specyficzne dla każdej lokalizacji lub instalacji. Konkretne szczegóły logowania można uzyskać od administratora systemu.

Ustaw dane uwierzytelniające. To pole wyboru powinno być niezaznaczone, aby możliwe było włączenie funkcji *pojedynczego uwierzytelniania*, która próbuje zalogować użytkownika przy użyciu nazwy użytkownika i hasła na lokalnym komputerze. Jeśli pojedyncze logowanie nie jest możliwe, lub w przypadku zaznaczenia tego pola w celu wyłączenia pojedynczego logowania (na przykład, w celu zalogowania się na konto administratora) wyświetlany jest kolejny ekran umożliwiający wprowadzenie danych uwierzytelniających.

Wprowadzanie danych uwierzytelniających do repozytorium

W zależności od ustawień w oknie dialogowym Repozytorium: Dane uwierzytelniające wymagane mogą być następujące pola:

ID użytkownika i hasło. Podaj poprawną nazwę użytkownika i hasło w celu zalogowania się. W razie konieczności skontaktuj się z lokalnym administratorem systemu w celu uzyskania dalszych informacji.

Dostawca. Wybierz dostawcę zabezpieczeń dla uwierzytelniania. Repozytorium można skonfigurować w sposób pozwalający na korzystanie z różnych dostawców zabezpieczeń; w razie potrzeby skontaktuj się z lokalnym administratorem systemu w celu uzyskania dalszych informacji.

Zapamiętaj identyfikator repozytorium i użytkownika. Zapisuje bieżące ustawienia jako domyślne, dzięki czemu nie będzie konieczne ponowne wprowadzanie ich za każdym razem podczas próby połączenia.

Przeglądanie danych uwierzytelniających repozytorium

Podczas łączenia się z repozytorium za pośrednictwem węzła źródłowego Analytic Server, Cognos, ODBC lub TM1 można wybrać wcześniej zapisane dane uwierzytelniające. Dane te są wyświetlone w oknie dialogowym **Wybierz dane uwierzytelniające dla repozytorium**. Aby wybrać to okno dialogowe, kliknij przycisk Przeglądaj obok pola **Dane uwierzytelniające**.

W oknie dialogowym **Wybierz dane uwierzytelniające dla repozytorium** wyróżnij dane uwierzytelniające na udostępnionej liście i kliknij przycisk OK. Jeśli lista jest zbyt duża, użyj pola **Filtrowanie**, aby wprowadzić nazwę lub część nazwy i wyszukać wymagane dane uwierzytelniające.

Przeglądanie zawartości repozytorium

Repozytorium umożliwia przeglądanie zapisanej zawartości w sposób podobny jak w programie Windows Explorer; można również przeglądać *wersje* poszczególnych zapisanych obiektów.

1. Aby otworzyć okno IBM SPSS Collaboration and Deployment Services Repository, w menu SPSS Modeler kliknij:

Narzędzia > **Repozytorium** > **Eksploruj...**

1. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie “Łączenie się z repozytorium” na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.

W oknie przeglądarki początkowo wyświetlany jest widok drzewa hierarchii folderu. Kliknij nazwę folderu, aby wyświetlić jego zawartość.

Obiekty, które są zgodne z bieżącym wyborem lub z kryteriami wyszukiwania, są przedstawione w panelu po prawej stronie, wraz ze szczegółowymi informacjami na temat wybranej wersji wyświetlonymi w dolnym panelu po prawej stronie. Wyświetlane atrybuty mają zastosowanie do najnowszej wersji.

Zapisywanie obiektów w repozytorium

Strumienie, węzły, modele, palety modeli, projekty i obiekty wyjściowe można zapisywać w repozytorium, które umożliwia innym użytkownikom i aplikacjom uzyskiwanie dostępu do tych danych.

Elementy wyjściowe strumienia można również publikować w repozytorium w formacie, który umożliwia innym użytkownikom wyświetlanie ich w Internecie za pośrednictwem portalu IBM SPSS Collaboration and Deployment Services Deployment Portal.

Określanie właściwości obiektów

Gdy użytkownik wybierze polecenie zapisania obiektu, wyświetlane jest okno dialogowe Repozytorium: Składowanie, w którym można określić wartości szeregu właściwości obiektu. Można:

- Wybrać nazwę folderu repozytorium, w którym obiekt ma być składowany.
- Dodać informacje o obiekcie, takie jak etykieta wersji i inne właściwości podlegające przeszukiwaniu.
- Przypisać do obiektu jeden lub więcej tematów klasyfikacji.

- Określić ustawienia zabezpieczeń obiektu.

W poniższych sekcjach opisano właściwości, które można określić.

Wybór lokalizacji zapisywania obiektów

W oknie dialogowym Repozytorium: Składuj wprowadź poniższe informacje.

Zapisz w. Wyświetla bieżący folder — lokalizację, w której zostanie zapisany obiekt. Kliknij dwukrotnie nazwę folderu na liście, aby ustawić ten folder jako folder bieżący. Użyj przycisku Do góry o jeden poziom, aby przejść do folderu nadrzędnego. Użyj przycisku Nowy folder, aby utworzyć folder na bieżącym poziomie.

Nazwa pliku. Nazwa, pod którą obiekt ma zostać zapisany.

Składuj. Zapisuje obiekt w bieżącej lokalizacji.

Dodawanie informacji o zapisanych obiektach

Wszystkie pola na karcie Informacje w oknie dialogowym Repozytorium: Składuj są opcjonalne.

Autor. Nazwa użytkownika tworzącego obiekt w repozytorium. Domyślnie wyświetlana jest tutaj nazwa użytkownika, na potrzeby połączenia z repozytorium można ją jednak zmienić.

Etykieta wersji. Wybierz etykietę z listy, aby określić wersję obiektu, lub kliknij opcję **Dodaj**, aby utworzyć nową etykietę. Nie używaj w etykiecie znaku „[”. Upewnij się, że nie jest zaznaczone żadne okno, o ile nie chcesz przypisać etykiety do tej wersji obiektu. Więcej informacji można znaleźć w temacie “Wyświetlanie i edytowanie właściwości obiektu” na stronie 214.

Description. Opis obiektu. Użytkownicy mogą wyszukiwać obiekty wg opisu (patrz uwaga).

Słowa kluczowe. Jedno lub więcej słów kluczowych związanych z obiektem, które może być używane do wyszukiwania (patrz uwaga).

Data wygaśnięcia. Data, po której obiekt przestaje być widoczny dla większości użytkowników, mimo że nadal jest widoczny dla jego właściciela oraz dla administratora repozytorium. Aby ustawić datę utraty ważności, wybierz opcję **Data** i wprowadź datę lub wybierz datę, korzystając z przycisku kalendarza.

Składuj. Zapisuje obiekt w bieżącej lokalizacji.

Uwaga: Informacje w polach **Opis** i **Słowa kluczowe** są traktowane jako odmienne od tych wprowadzonych na karcie Adnotacje obiektu w programie SPSS Modeler. Wyszukiwanie w repozytorium wg opisu lub słowa kluczowego nie zwraca informacji z karty Adnotacje. Więcej informacji można znaleźć w temacie “Wyszukiwanie obiektów w repozytorium ” na stronie 211.

Przypisywanie tematów do zapisanego obiektu

Tematy składają się na system klasyfikacji hierarchicznej treści składowanych w repozytorium. Tematy można wybrać spośród listy dostępnych podczas zapisywania obiektów. Ponadto użytkownicy mogą również wyszukiwać obiekty wg tematu. Lista dostępnych tematów jest ustalana przez użytkowników repozytorium z odpowiednimi uprawnieniami (więcej informacji zawiera publikacja *Deployment Manager User's Guide*).

Aby przypisać temat do obiektu, na karcie Tematy w oknie dialogowym Repozytorium: Składuj:

1. Kliknij przycisk **Dodaj**.
2. Kliknij nazwę tematu na liście dostępnych tematów.
3. Kliknij przycisk **OK**.

   Aby usunąć przypisanie tematu:
4. Wybierz temat z listy przypisanych tematów:
5. Kliknij przycisk **Usuń**.

Ustawianie opcji zabezpieczeń dla zapisanych obiektów

Liczbę opcji zabezpieczeń dla zapisanego obiektu można zmienić w oknie dialogowym Repozytorium: Składuj. Dla jednej lub więcej **nazw użytkowników** (to jest, użytkowników lub grup użytkowników) można:

- Przypisywać prawa dostępu do obiektu
- Modyfikować prawa dostępu do obiektu
- Usuwać prawa dostępu do obiektu

Nazwa użytkownika. Nazwa użytkownika lub grupy użytkowników repozytorium mających prawa dostępu do tego obiektu.

Uprawnienia. Prawa dostępu, które ten użytkownik lub grupa użytkowników ma do obiektu.

Dodaj. Umożliwia dodawanie jednego lub więcej użytkowników lub grup użytkowników do listy tych z prawami dostępu do tego obiektu. Więcej informacji można znaleźć w temacie “Dodawanie użytkowników do listy uprawnień” na stronie 207.

Zmodyfikuj. Umożliwia modyfikowanie praw dostępu wybranego użytkownika lub grupy do tego obiektu. Domyślnie przyznawany jest dostęp do odczytu. Opcja ta pozwala przyznać dodatkowe prawa dostępu, takie jak Właściciel, Zapis, Usuń i Zmodyfikuj uprawnienia.

Usuń. Usuwa wybranego użytkownika lub grupę z listy uprawnień dla tego obiektu.

Dodawanie użytkowników do listy uprawnień

Poniższe pola są dostępne po wybraniu opcji **Dodaj** na karcie Zabezpieczenia w oknie dialogowym Repozytorium: Składuj:

Wybierz dostawcę. Wybierz dostawcę zabezpieczeń dla uwierzytelniania. Repozytorium można skonfigurować w sposób pozwalający na korzystanie z różnych dostawców zabezpieczeń; w razie potrzeby skontaktuj się z lokalnym administratorem systemu w celu uzyskania dalszych informacji.

Znajdź. Wprowadź nazwę użytkownika użytkownika lub grupy użytkowników repozytorium, którą chcesz dodać, a następnie kliknij opcję **Wyszukaj**, aby wyświetlić tę nazwę na liście użytkowników. Aby dodać więcej niż jedną nazwę użytkownika naraz, pozostaw to pole puste i po prostu kliknij opcję **Wyszukaj**, aby wyświetlić listę wszystkich nazw użytkowników repozytorium.

Lista użytkowników. Wybierz jedną lub więcej nazw użytkowników z listy i kliknij przycisk OK, aby dodać je do listy uprawnień.

Modyfikowanie praw dostępu do obiektu

Poniższe pola są dostępne po wybraniu opcji **Zmodyfikuj** na karcie Zabezpieczenia w oknie dialogowym Repozytorium: Składuj:

Właściciel. Wybierz tę opcję, aby nadać temu użytkownikowi lub grupie użytkowników prawa dostępu do obiektu. Właściciel ma pełną kontrolę nad obiektem, w tym możliwość wykonywania operacji Usuń i Zmodyfikuj.

Odczyt. Domyślnie użytkownik lub grupa niebędąca właścicielem obiektu ma tylko prawa dostępu do odczytu danego obiektu. Zaznaczenie odpowiednich pól wyboru pozwala na dodanie temu użytkownikowi lub grupie praw dostępu do zapisu, usuwania i modyfikowania.

Składowanie strumieni

Strumień można zapisać w repozytorium jako plik *.str*. Składowany w ten sposób strumień będzie dostępny dla innych użytkowników.

Uwaga: Aby uzyskać informacje o wdrażaniu strumienia w taki sposób, by możliwe było skorzystanie z dodatkowych funkcji repozytorium, patrz “Wdrażanie strumieni” na stronie 215.

Aby zapisać bieżący strumień:

1. W menu głównym kliknij opcje:

   Plik > **Składuj** > **Składuj jako strumień...**

2. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie “Łączenie się z repozytorium” na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.
3. W oknie dialogowym Repozytorium: Składuj wybierz folder, w którym chcesz składować obiekt, podaj pozostałe informacje, które chcesz zapisać, a następnie kliknij przycisk **Składuj**. Więcej informacji można znaleźć w temacie “Określanie właściwości obiektów” na stronie 205.

Składowanie projektów

Kompletny projekt programu IBM SPSS Modeler można zapisać w repozytorium jako plik *.cpj*. Składowany w ten sposób projekt będzie dostępny dla innych użytkowników.

Ponieważ plik projektu jest kontenerem dla innych obiektów programu IBM SPSS Modeler, należy poinstruować program IBM SPSS Modeler, aby składował obiekty projektu w repozytorium. W tym celu należy wybrać odpowiednie ustawienie w oknie dialogowym Właściwości projektu. Więcej informacji można znaleźć w temacie “Ustawianie właściwości projektu” na stronie 230.

Gdy już skonfigurujemy projekt tak, aby jego obiekty były składowane w repozytorium, każdorazowo przy dodawaniu nowego obiektu do projektu program IBM SPSS Modeler będzie wyświetlał monit dotyczący składowania obiektu.

Po zakończeniu sesji pracy w programie IBM SPSS Modeler należy zapisać w repozytorium nową wersję pliku projektu, aby dodane obiekty zostały zapamiętane. W pliku projektu automatycznie umieszczane są (i są z niego pobierane) najnowsze wersje obiektów projektu. Jeśli w trakcie sesji pracy z programem IBM SPSS Modeler nie dodano do projektu żadnych obiektów, to nie trzeba ponownie zapisywać pliku projektu w repozytorium. Należy jednak zapisać w repozytorium nowe wersje tych obiektów projektu (strumieni, wyników itd.), które zostały zmienione.

Aby zapisać (składować) projekt w repozytorium:

1. Wybierz projekt na karcie CRISP-DM lub Klasy na panelu zarządzania w programie IBM SPSS Modeler i w menu głównym kliknij:

   Plik > **Projekt** > **Składuj projekt...**

2. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie “Łączenie się z repozytorium” na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.
3. W oknie dialogowym Repozytorium: Składuj wybierz folder, w którym chcesz składować obiekt, podaj pozostałe informacje, które chcesz zapisać, a następnie kliknij przycisk **Składuj**. Więcej informacji można znaleźć w temacie “Określanie właściwości obiektów” na stronie 205.

Składowanie węzłów

Definicję określonego węzła z bieżącego strumienia można zapisać w repozytorium jako plik *.nod*. Składowany w ten sposób węzeł będzie dostępny dla innych użytkowników.

Aby zapisać (składować) węzeł w repozytorium:

1. Prawym przyciskiem myszy kliknij węzeł w obszarze roboczym strumienia i kliknij polecenie **Składuj węzeł**.
2. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie “Łączenie się z repozytorium” na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.

3. W oknie dialogowym Repozytorium: Składuj wybierz folder, w którym chcesz składować obiekt, podaj pozostałe informacje, które chcesz zapisać, a następnie kliknij przycisk **Składuj**. Więcej informacji można znaleźć w temacie “Określanie właściwości obiektów” na stronie 205.

Składowanie obiektów wynikowych

Obiekt wynikowy z bieżącego strumienia można zapisać w repozytorium jako plik *.cpu*. Składowany w ten sposób obiekt będzie dostępny dla innych użytkowników.

Aby zapisać (składować) w repozytorium obiekt wynikowy:

1. Kliknij obiekt na karcie Wyniki panelu zarządzania w programie SPSS Modeler, a następnie w menu głównym kliknij kolejno:

   Plik > **Wyniki** > **Składuj wynik...**
2. Zamiast tego można kliknąć obiekt prawym przyciskiem myszy na karcie Wyniki, a następnie kliknąć opcję **Składuj**.
3. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie “Łączenie się z repozytorium” na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.
4. W oknie dialogowym Repozytorium: Składuj wybierz folder, w którym chcesz składować obiekt, podaj pozostałe informacje, które chcesz zapisać, a następnie kliknij przycisk **Składuj**. Więcej informacji można znaleźć w temacie “Określanie właściwości obiektów” na stronie 205.

Składowanie modeli i palet modeli

Określony model można zapisać w repozytorium jako plik *.gm*. Składowany w ten sposób model będzie dostępny dla innych użytkowników. Można też zapisać całą zawartość palety Modele w repozytorium jako plik *.gen*.

Składowanie modelu:

1. Kliknij obiekt na palecie Modele w programie SPSS Modeler, a następnie w menu głównym kliknij kolejno:

   Plik > **Modele** > **Składuj model...**
2. Zamiast tego można kliknąć obiekt prawym przyciskiem myszy na palecie Modele, a następnie kliknąć opcję **Składuj model**.
3. Kontynuuj od punktu „Kończenie procedury składowania”.

Składowanie palety modeli:

1. Prawym przyciskiem myszy kliknij tło palety modeli.
2. Z menu podręcznego wybierz polecenie **Składuj paletę**.
3. Kontynuuj od punktu „Kończenie procedury składowania”.

Kończenie procedury składowania:

1. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie “Łączenie się z repozytorium” na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.
2. W oknie dialogowym Repozytorium: Składuj wybierz folder, w którym chcesz składować obiekt, podaj pozostałe informacje, które chcesz zapisać, a następnie kliknij przycisk **Składuj**. Więcej informacji można znaleźć w temacie “Określanie właściwości obiektów” na stronie 205.

Pobieranie obiektów z repozytorium

Istnieje możliwość pobierania strumieni, modeli, palet modeli, węzłów, projektów i obiektów wyjściowych zapisanych w repozytorium.

Uwaga: Poza korzystaniem z opcji menu zgodnie z opisem tutaj można również pobierać strumienie, obiekty wyjściowe, modele i palety modeli, klikając prawym przyciskiem myszy na odpowiedniej karcie w panelu zarządzania w górnej części okna SPSS Modeler po prawej stronie.

1. Aby pobrać strumień, w menu głównym programu IBM SPSS Modeler kliknij:

   Plik > **Pobierz strumień...**
2. Aby pobrać model, paletę modeli, projekt lub obiekt wyjściowy, w menu głównym programu IBM SPSS Modeler kliknij:

   Plik > **Modele** > **Pobierz model...**

   or

   Plik > **Modele** > **Pobierz paletę modeli...**

   or

   Plik > **Projekty** > **Pobierz projekt...**

   or

   Plik > **Wyniki** > **Pobierz wynik...**
3. Alternatywnie, kliknij prawym przyciskiem myszy w panelu zarządzania lub projektów i kliknij opcję **Pobierz** w menu podręcznym.
4. Aby pobrać węzeł, w menu głównym programu IBM SPSS Modeler kliknij:

   Wstaw > **Węzeł (lub Superwęzeł) z repozytorium...**

   a. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie “Łączenie się z repozytorium” na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.
5. W oknie dialogowym Repozytorium: Otwórz przejdź do obiektu, zaznacz go i kliknij przycisk **Pobierz**. Więcej informacji można znaleźć w temacie .

Wybieranie obiektu do pobrania

W oknie dialogowym Repozytorium: Pobierz/Wyszukaj dostępne są następujące pola:

Szukaj w. Zawiera hierarchię folderów wewnątrz bieżącego folderu. Aby przejść do innego folderu, wybierz folder z listy, co spowoduje przejście bezpośrednio do wybranego folderu, albo nawiguj za pomocą listy obiektów pod tym polem.

Przycisk Do góry o jeden poziom. Powoduje przejście o jeden poziom wyżej w hierarchii folderów.

Przycisk Nowy folder. Tworzy nowy folder na bieżącym poziomie w hierarchii.

Nazwa pliku. Nazwa pliku wybranego obiektu w repozytorium. Aby pobrać ten obiekt, kliknij przycisk **Pobierz**.

Pliki typu. Typ obiektu wybranego do pobrania. Na liście obiektów wyświetlane będą tylko obiekty tego typu wraz z folderami. Aby wyświetlić obiekty innego typu, wybierz typ obiektów z listy.

Otwórz jako zablokowany. Domyślnie obiekt, który został pobrany, pozostaje w repozytorium jako zablokowany, tak aby inni użytkownicy nie mogli go zmienić. Jeśli nie chcesz, by obiekty były blokowane po pobraniu, usuń zaznaczenie tego pola wyboru.

Opis, słowa kluczowe. Jeśli podczas zapisywania obiektu w repozytorium określono dodatkowe dotyczące go informacje, to będą one tutaj wyświetlane. Więcej informacji można znaleźć w temacie “Dodawanie informacji o zapisanych obiektach” na stronie 206.

Wersja. Aby pobrać wersję obiektu inną niż najnowsza, kliknij ten przycisk. Zostaną wyświetlone szczegółowe informacje na temat wszystkich wersji obiektu i możliwe będzie wybranie żądanej wersji.

Wybieranie wersji obiektu

Aby wybrać konkretną wersję obiektu z repozytorium, w oknie dialogowym Repozytorium: Wybierz wersję:

1. (Opcjonalnie) Posortuj listę według wersji, etykiety, rozmiaru, daty utworzenia lub osoby, która utworzyła obiekt, klikając dwukrotnie nagłówek odpowiedniej kolumny.
2. Wybierz wersję obiektu, z którą chcesz pracować.
3. Kliknij przycisk Kontynuuj.

Wyszukiwanie obiektów w repozytorium

Obiekty można wyszukiwać wg nazwy, folderu, typu, etykiety, daty lub innych kryteriów.

Wyszukiwanie obiektów według nazwy

1. W menu głównym programu IBM SPSS Modeler kliknij:

   Narzędzia > **Repozytorium** > **Eksploruj...**

   a. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie “Łączenie się z repozytorium” na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.
2. Kliknij zakładkę **Wyszukaj**.
3. W polu **Szukaj obiektów o nazwie** podaj nazwę obiektu, jaki ma zostać znaleziony.

Podczas wyszukiwania obiektów wg nazwy gwiazdka (*) może być używana jako symbol wieloznaczny pozwalający na dopasowanie łańcucha znaków, a znak zapytania (?) oznacza dowolny pojedynczy znak. Na przykład zapis `*cluster*` pozwala dopasować wszystkie obiekty zawierające łańcuch `cluster` w dowolnym miejscu w nazwie. Łańcuch wyszukiwania `m0?_*` pozwala dopasować zapis `M01_cluster.str` oraz `M02_cluster.str`, ale zapis `M01a_cluster.str` już nie jest dopasowany. W wyszukiwaniu nie jest uwzględniana wielkość liter (zapis `cluster` jest zgodny z zapisem `Cluster` oraz z zapisem `CLUSTER`).

Uwaga: Jeśli obiektów jest bardzo dużo, wyszukiwanie może chwilę potrwać.

Wyszukiwanie według innych kryteriów

Wyszukiwanie można przeprowadzić na podstawie tytułu, etykiety, dat, autora, słów kluczowych, zaindeksowanej zawartości lub opisu. Wyszukane zostaną tylko obiekty, które są zgodne ze *wszystkimi* określonymi kryteriami wyszukiwania. Na przykład można zlokalizować wszystkie strumienie zawierające co najmniej jeden model grupowania, do którego zastosowano konkretną etykietę i który został zmodyfikowany po określonej dacie.

Typy obiektów. Wyszukiwanie można ograniczyć do modeli, strumieni, wyników, węzłów, superwęzłów, projektów, palet modeli lub innego typu obiektów.

- **Modele.** Modele można wyszukiwać według kategorii (klasyfikacja, przybliżenie, grupowanie itd.) lub według określonego algorytmu modelowania, takiego jak modelowanie Kohonena.

  Można również wyszukiwać według użytych zmiennych — na przykład wszystkie modele, w których jako zmienna wejściowa lub wyjściowa (przewidywana) została użyta zmienna o nazwie *dochód*.
- **Strumienie.** W przypadku strumieni wyszukiwanie można ograniczyć według użytych zmiennych lub typu modelu (kategoria lub algorytm) zawartych w strumieniu.

Tematy. Wyszukiwanie można przeprowadzić w modelach powiązanych z określonymi tematami z listy opracowanej według użytkowników repozytorium z odpowiednimi uprawnieniami (więcej informacji

zawiera publikacja *Deployment Manager User's Guide*). Aby uzyskać dostęp do listy, zaznacz to pole, a następnie kliknij wyświetlony przycisk Dodaj tematy, wybierz co najmniej jeden temat z listy i kliknij przycisk OK.

Etykieta. Ogranicza wyszukiwanie do konkretnych etykiet wersji obiektu.

Daty. Można określić datę utworzenia lub modyfikacji i przeprowadzić wyszukiwanie w obiektach w określonym przedziale dat, przed tym przedziałem lub po nim.

Autor. Ogranicza wyszukiwanie do obiektów utworzonych przez konkretnego użytkownika.

Słowa kluczowe. Umożliwia wyszukiwanie według określonych słów kluczowych. W programie IBM SPSS Modeler słowa kluczowe są określane na karcie Adnotacja dla strumienia, modelu, lub obiektu wyjściowego.

Opis. Umożliwia wyszukiwanie według określonych składników w polu opisowym. W programie IBM SPSS Modeler opis jest wprowadzany na karcie Adnotacja dla strumienia, modelu, lub obiektu wyjściowego. Wiele fraz wyszukiwania można oddzielać średnikami — na przykład `income; crop type; claim value`. (Należy zwrócić uwagę, że w frazie wyszukiwania spacje mają znaczenie. Na przykład, zapisy `crop type` z jedną spacją i `crop  type` z dwiema spacjami różnią się od siebie).

Modyfikowanie obiektów w repozytorium

Bezpośrednio z programu SPSS Modeler można modyfikować istniejące obiekty w repozytorium. Można:

- Tworzyć foldery, zmieniać nazwy folderów i usuwać foldery.
- Blokować i odblokowywać obiekty.
- Usuwać obiekty.

Tworzenie, zmiana nazwy i usuwanie folderów

1. Aby wykonywać operacje na folderach w repozytorium, w menu głównym programu SPSS Modeler kliknij:

   Narzędzia > **Repozytorium** > **Eksploruj...**

   a. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie "Łączenie się z repozytorium" na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.
2. Upewnij się, że karta **Foldery** jest aktywna.
3. Aby utworzyć nowy folder, kliknij prawym przyciskiem myszy folder nadrzędny, a następnie kliknij przycisk **Nowy folder**.
4. Aby zmienić nazwę folderu, kliknij go prawym przyciskiem myszy, a następnie kliknij przycisk **Zmień nazwę folderu**.
5. Aby usunąć folder, kliknij go prawym przyciskiem myszy, a następnie kliknij przycisk **Usuń folder**.

Blokowanie i usuwanie blokady obiektów repozytorium

Obiekt można zablokować, aby uniemożliwić innym użytkownikom aktualizowanie jego istniejących wersji lub tworzenie nowych. Zablokowany obiekt jest oznaczony symbolem kłódki umieszczonym nad ikoną obiektu.

Rysunek 17. Zablokowany obiekt

Aby zablokować obiekt

1. W oknie przeglądarki repozytorium kliknij wymagany obiekt prawym przyciskiem myszy.

2. Kliknij przycisk **Zablokuj**.

Aby odblokować obiekt

1. W oknie przeglądarki repozytorium kliknij wymagany obiekt prawym przyciskiem myszy.
2. Kliknij przycisk **Odblokuj**.

Usuwanie obiektów repozytorium

Przed usunięciem obiektu z repozytorium należy zdecydować, czy usunięte mają zostać wszystkie wersje tego obiektu czy tylko konkretna jego wersja.

Aby usunąć wszystkie wersje obiektu

1. W oknie przeglądarki repozytorium kliknij wymagany obiekt prawym przyciskiem myszy.
2. Kliknij przycisk **Usuń obiekty**.

Aby usunąć najnowszą wersję obiektu

1. W oknie przeglądarki repozytorium kliknij wymagany obiekt prawym przyciskiem myszy.
2. Kliknij przycisk **Usuń**.

Aby usunąć poprzednią wersję obiektu

1. W oknie przeglądarki repozytorium kliknij wymagany obiekt prawym przyciskiem myszy.
2. Kliknij przycisk **Usuń wersje**.
3. Wybierz wersje do usunięcia i kliknij przycisk **OK**.

Zarządzanie właściwościami obiektów w repozytorium

Z programu SPSS Modeler można modyfikować różne właściwości obiektów. Można:

- Wyświetlać właściwości folderu.
- Wyświetlać i edytować właściwości obiektu.
- Tworzyć, stosować i usuwać etykiety wersji obiektu.

Wyświetlanie właściwości folderu

Aby wyświetlić właściwości dowolnego folderu w oknie repozytorium, kliknij wymagany folder prawym przyciskiem myszy. Kliknij **Właściwości folderu**.

Karta Opcje ogólne

Na tej karcie wyświetlana jest nazwa folderu, daty utworzenia go i modyfikacji.

Karta Uprawnienia

Na tej karcie można określić uprawnienia do odczytu i zapisu dla określonego folderu. Wyświetlani są wszyscy użytkownicy i grupy z prawem dostępu do folderu nadrzędnego. Uprawnienia są nadawane według hierarchii. Na przykład, użytkownik, który ma uprawnienia do odczytu, nie może mieć uprawnień do zapisu. Jeśli użytkownik nie ma uprawnień do zapisu, nie może mieć uprawnień do usuwania.

Użytkownicy i grupy. Wyświetla listę użytkowników i grup repozytorium posiadających co najmniej prawo do odczytu danego folderu. Zaznacz pola wyboru Zapisz i Usuń, aby dodać te prawa dostępu dla danego folderu do określonego użytkownika lub grupy. Kliknij ikonę **Dodaj użytkowników/grupy** po prawej stronie karty Uprawnienia, aby przypisać prawo dostępu do kolejnych użytkowników i grup. Listą dostępnych użytkowników i grup steruje administrator.

Dziedziczenie uprawnień. Ta opcja pozwala kontrolować, w jaki sposób zmiany dokonane w bieżącym folderze będą stosowane w jego folderach podrzędnych, o ile wystąpią.

- **Dziedziczenie wszystkich uprawnień.** Umożliwia dziedziczenie ustawień uprawnień z bieżącego folderu we wszystkich folderach podrzędnych i potomnych. Jest to szybki sposób ustawienia uprawnień

dla kilku folderów jednocześnie. Ustaw uprawnienia zgodnie z wymaganiami dla folderu nadrzędnego, a następnie przeprowadź ich dziedziczenie odpowiednio do wymagań.

- **Dziedziczenie wyłącznie zmian.** Umożliwia dziedziczenie tylko zmian wprowadzonych od czasu ostatniego ich zastosowania. Na przykład, jeśli została dodana nowa grupa i ma uzyskać prawo dostępu do wszystkich folderów w gałęzi Sprzedaż, można nadać tej grupie prawo dostępu do folderu głównego Sprzedaż i przeprowadzić dziedziczenie tej zmiany we wszystkich folderach podrzędnych. Wszystkie uprawnienia do istniejących folderów podrzędnych pozostaną bez zmian.
- **Bez dziedziczenia.** Wszystkie wprowadzone zmiany będą miały zastosowanie tylko do bieżącego folderu i nie będą dziedziczone w folderach podrzędnych.

Wyświetlanie i edytowanie właściwości obiektu

W oknie dialogowym Właściwości obiektu można wyświetlać i edytować właściwości. Choć niektórych właściwości nie można zmienić, zawsze można zaktualizować obiekt poprzez dodanie nowej wersji.

1. W oknie repozytorium kliknij wymagany obiekt prawym przyciskiem myszy.
2. Kliknij przycisk **Właściwości obiektu**.

Karta Opcje ogólne

Nazwa. Nazwa obiektu, zgodna z nazwą wyświetlaną w repozytorium.

Utworzono dnia. Data utworzenia obiektu (nie wersji).

Ostatnia modyfikacja. Data ostatniej modyfikacji wersji.

Autor. Nazwa logowania użytkownika.

Description. Domyślnie zawiera opis określony na karcie Adnotacja obiektu w programie SPSS Modeler.

Połączone tematy. O ile jest to konieczne, repozytorium umożliwia organizowanie modeli i powiązanych obiektów według tematów. Lista dostępnych tematów jest ustalana przez użytkowników repozytorium z odpowiednimi uprawnieniami (więcej informacji zawiera publikacja *Deployment Manager User's Guide*).

Słowa kluczowe. Słowa kluczowe są określane dla strumienia, modelu lub obiektu wyjściowego na karcie Adnotacja. Wiele słów kluczowych należy oddzielić spacjami; maksymalna liczba znaków to 255. (Jeśli słowa kluczowe zawierają spacje, należy je rozdzielić znakiem cudzysłowu.)

Karta Wersje

Obiekty zapisane w repozytorium mogą mieć kilka wersji. Na karcie Wersje wyświetlane są informacje na temat poszczególnych wersji.

Poniżej przedstawiono właściwości, jakie można określić lub zmodyfikować dla określonych wersji zapisanego obiektu:

Wersja. Unikalny identyfikator wersji wygenerowany na podstawie czasu zapisania wersji.

Etykieta. Aktualna etykieta dla wersji, o ile jest dostępna. W odróżnieniu od identyfikatora wersji etykiety można przenosić pomiędzy wersjami obiektów.

Dla każdej wersji wyświetlana jest również wielkość pliku, data utworzenia i autor.

Edytuj etykiety. Kliknij ikonę **Edytuj etykiety** u góry po prawej stronie karty Wersje, aby zdefiniować, zastosować lub usunąć etykiety dla zapisanych obiektów. Więcej informacji można znaleźć w temacie "Zarządzanie etykietami wersji obiektu" na stronie 215.

Karta Uprawnienia

Na karcie Uprawnienia można ustawić uprawnienia do odczytu i zapisu dla obiektu. Wyświetlani są wszyscy użytkownicy i grupy z prawem dostępu do bieżącego obiektu. Uprawnienia są nadawane według hierarchii. Na przykład, użytkownik, który ma uprawnienia do odczytu, nie może mieć uprawnień do zapisu. Jeśli użytkownik nie ma uprawnień do zapisu, nie może mieć uprawnień do usuwania.

Użytkownicy i grupy. Wyświetla listę użytkowników i grup repozytorium posiadających co najmniej prawo do odczytu danego obiektu. Zaznacz pola wyboru Zapis i Usuń, aby dodać te prawa dostępu

dla danego obiektu do określonego użytkownika lub grupy. Kliknij ikonę **Dodaj użytkowników/grupy** po prawej stronie karty Uprawnienia, aby przypisać prawo dostępu do kolejnych użytkowników i grup. Listą dostępnych użytkowników i grup steruje administrator.

Zarządzanie etykietami wersji obiektu

Okno dialogowe Edytuj etykiety wersji umożliwia:

- Zastosowanie etykiet do wybranego obiektu
- Usuwanie etykiet z wybranego obiektu
- Zdefiniowanie nowej etykiety i zastosowanie jej w obiekcie

Aby zastosować etykiety do obiektu

1. Wybierz co najmniej jedną etykietę z listy **Dostępne etykiety**.
2. Kliknij przycisk strzałki w prawo, aby przenieść wybrane etykiety na listę **Zastosowane etykiety**.
3. Kliknij przycisk **OK**.

Aby usunąć etykiety z obiektu

1. Wybierz co najmniej jedną etykietę z listy **Zastosowane etykiety**.
2. Kliknij przycisk strzałki w lewo, aby przenieść wybrane etykiety na listę **Dostępne etykiety**.
3. Kliknij przycisk **OK**.

Aby zdefiniować nową etykietę i zastosować ją do obiektu

1. Wpisz nazwę etykiety w polu **Nowa etykieta**.
2. Kliknij przycisk strzałki w prawo, aby przenieść nową etykietę na listę **Zastosowane etykiety**.
3. Kliknij przycisk **OK**.

Wdrażanie strumieni

Aby możliwe było używane strumienia z uproszczoną aplikacją kliencką IBM SPSS Modeler Advantage, należy wdrożyć go jako strumień (plik `.str`) w repozytorium.

Uwaga: Nie można wdrożyć strumienia, który w gałęzi oceniania ma więcej niż jeden węzeł źródłowy.

Wdrożony strumień może w pełni korzystać z funkcji klasy korporacyjnej dostępnych w oprogramowaniu IBM SPSS Collaboration and Deployment Services. Więcej informacji można znaleźć w temacie “Przechowywanie i wdrażanie obiektów repozytorium” na stronie 204.

Aby wdrożyć bieżący strumień (przy użyciu menu Plik):

1. W menu głównym kliknij opcje:

   Plik > **Składuj** > **Dokonaj wdrożenia**
2. Wybierz typ wdrożenia i odpowiednio do potrzeb wypełnij pozostałe pola w oknie dialogowym.
3. Kliknij opcję **Wdrażaj jako strumień**, aby wdrożyć strumień do użytku z programem IBM SPSS Modeler Advantage lub IBM SPSS Collaboration and Deployment Services.
4. Kliknij opcję **Składuj**. Aby uzyskać więcej informacji, kliknij przycisk **Pomoc**.
5. Kontynuuj od punktu „Kończenie procesu wdrażania”.

Aby wdrożyć bieżący strumień (przy użyciu menu Narzędzia):

1. W menu głównym kliknij opcje:

   Narzędzia > **Właściwości strumienia** > **Wdrożenie**

2. Wybierz typ wdrożenia, odpowiednio do potrzeb wypełnij pozostałe pola na karcie Wdrożenie i kliknij przycisk **Składuj**. Więcej informacji można znaleźć w temacie "Opcje wdrażania strumieni" na stronie 216.

Kończenie procesu wdrażania

1. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie "Łączenie się z repozytorium" na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.
2. W oknie dialogowym Repozytorium: Składuj wybierz folder, w którym chcesz składować obiekt, podaj pozostałe informacje, które chcesz zapisać, a następnie kliknij przycisk **Składuj**. Więcej informacji można znaleźć w temacie "Określanie właściwości obiektów" na stronie 205.

Opcje wdrażania strumieni

Karta Wdrożenie dostępna w oknie dialogowym opcji strumienia umożliwia określenie opcji wdrożenia strumienia.

Wdrożony strumień jest przechowywany w repozytorium w postaci pliku z rozszerzeniem `.str`.

Wdrożenie strumienia umożliwia wykorzystanie dodatkowych funkcji oprogramowania IBM SPSS Collaboration and Deployment Services, takich jak dostęp dla wielu użytkowników, automatyczne ocenianie, odświeżanie modelu i analiza Champion Challenger.

Uwaga: Reguły asocjacyjne, węzły modelowania TCM i STP nie obsługują oceny modelu ani kroków Champion Challenger w oprogramowaniu IBM SPSS Collaboration and Deployment Services.

Z poziomu karty Wdrożenie można także przeglądać opisy strumieni tworzone w oprogramowaniu IBM SPSS Modeler z przeznaczeniem dla strumienia. Więcej informacji można znaleźć w temacie "Opisy strumieni" na stronie 53.

Typ wdrożenia. Zdecyduj, jak chcesz wdrożyć strumień. W przypadku wszystkich strumieni przed ich wdrożeniem należy zastosować określony węzeł oceniający. Opcje oraz dodatkowe wymagania są uzależnione od typu wdrożenia.

Uwaga: Reguły asocjacyjne, węzły modelowania TCM i STP nie obsługują oceny modelu ani kroków Champion Challenger w oprogramowaniu IBM SPSS Collaboration and Deployment Services.

- **<brak>**. Strumień nie zostanie wdrożony w repozytorium. Wszystkie opcje są wyłączone, z wyjątkiem podglądu opisu strumienia.
- **Tylko scoring**. Strumień jest wdrażany w repozytorium po kliknięciu przycisku **Składuj**. Dane mogą być oceniane za pomocą węzła wskazanego w zmiennej **Węzeł oceniający**.
- **Odświeżenie modelu**. Podobnie jak dla opcji Tylko scoring. Ponadto, model może zostać zaktualizowany w repozytorium za pomocą obiektów wskazanych w przez użytkownika w **węźle modelowania** oraz w **modelu użytkowym**. Automatyczne odświeżenie modelu nie jest obsługiwane domyślnie w programie IBM SPSS Collaboration and Deployment Services, dlatego konieczne jest wybranie tego typu wdrożenia w razie potrzeby użycia tej funkcji podczas uruchamiania strumienia z repozytorium. Więcej informacji można znaleźć w temacie "Odświeżanie modelu" na stronie 218.

Węzeł oceniający. Wybierz wykres, dane wynikowe lub węzeł eksportowania, aby zidentyfikować gałąź strumienia, która ma być używana do oceny danych. Choć strumień może zawierać dowolną liczbę poprawnych gałęzi, modeli i węzłów końcowych, do celów wdrożenia musi zostać wybrana jedna i tylko jedna gałąź oceniania. Jest to najbardziej podstawowy warunek wdrożenia jakiegokolwiek strumienia.

Parametry oceniania. Umożliwiają określenie parametrów, które można zmodyfikować podczas uruchamiania węzła modelowania. Więcej informacji można znaleźć w temacie "Parametry oceniania i modelowania" na stronie 217.

Węzeł modelowania. W przypadku odświeżenia modelu, określa węzeł modelowania służący do ponownego tworzenia lub aktualizowania modelu w repozytorium. Musi być to węzeł modelowania tego samego typu, co określony dla opcji **Model użytkowy**.

Parametry budowania modelu. Umożliwiają określenie parametrów, które można zmodyfikować podczas uruchamiania węzła modelowania. Więcej informacji można znaleźć w temacie “Parametry oceniania i modelowania” na stronie 217.

Model użytkowy. Na potrzeby odświeżenia modelu określa model użytkowy, który będzie aktualizowany lub ponownie tworzony przy każdej aktualizacji strumienia w repozytorium (zazwyczaj jako część zaplanowanego zadania). Model musi znajdować się na gałęzi oceniania. Choć na gałęzi oceniania może występować wiele modeli, tylko jeden z nich może zostać oznaczony. Należy zwrócić uwagę, że przy tworzeniu strumienia po raz pierwszy może być to w rzeczywistości model zastępczy, aktualizowany lub tworzony ponownie dopiero w miarę dostępności danych.

Sprawdź. Kliknij ten przycisk, aby sprawdzić, czy jest to właściwy strumień do wdrożenia. Wszystkie strumienie muszą mieć wyznaczony węzeł oceniający, zanim będzie możliwe ich wdrożenie. Jeśli te wymagania nie są spełnione, wyświetlane są komunikaty o błędach.

Składuj. Wdraża strumień, o ile jest on poprawny. W przeciwnym wypadku wyświetlany jest komunikat o błędzie. Kliknij przycisk **Napraw**, skoryguj błąd i spróbuj ponownie.

Podgląd opisu strumienia. Umożliwia wyświetlanie treści opisu strumienia, który program IBM SPSS Modeler tworzy dla strumienia. Więcej informacji można znaleźć w temacie “Opisy strumieni” na stronie 53.

Uwaga: Reguły asocjacyjne, węzły modelowania TCM i STP nie obsługują oceny modelu ani kroków Champion Challenger w oprogramowaniu IBM SPSS Collaboration and Deployment Services.

Parametry oceniania i modelowania

Podczas wdrażania strumienia w oprogramowaniu IBM SPSS Collaboration and Deployment Services można wybrać parametry, które będzie można wyświetlać lub edytować podczas aktualizacji lub oceniania modelu. Przykładowo: można określić wartości minimalne i maksymalne lub inne wartości, które mogą ulegać zmianie podczas uruchamiania zadania.

1. Aby zapewnić widoczność parametru w celu jego wyświetlenia lub edytowania po wdrożeniu strumienia, należy go wybrać na liście w oknie dialogowym Parametry oceniania.

Lista dostępnych parametrów jest zdefiniowana na karcie Parametry w oknie dialogowym właściwości strumienia. Więcej informacji można znaleźć w temacie “Konfigurowanie parametrów strumienia i sesji” na stronie 49.

Gałąź oceniania

Aby możliwe było wdrożenie strumienia, jedna gałąź strumienia musi być wskazana jako **gałąź oceniania** (tj. gałąź zawierająca węzeł oceniający). Gałąź wskazana jako gałąź oceniania jest wyróżniona w obszarze roboczym strumienia, podobnie jak łącze modelu prowadzące do modelu użytkowego w gałęzi oceniania. Ta forma reprezentacji wizualnej jest szczególnie przydatna w złożonych strumieniach z wieloma gałęziami, gdzie nie zawsze od razu wiadomo, która gałąź jest gałęzią oceniania.

Uwaga: Tylko jedną gałąź strumienia można wskazać jako gałąź oceniania.

Jeśli dla strumienia jest już zdefiniowana gałąź oceniania, nowo wskazana gałąź zastąpi dotychczasową. Korzystając z opcji Niestandardowe kolory, można określić kolor wyróżnienia gałęzi oceniania. Więcej informacji można znaleźć w temacie “Określanie opcji wyświetlania” na stronie 238.

Wyróżnienie gałęzi oceniania można uwidaczniać lub ukrywać, stosując przycisk Pokaż/Ukryj adnotacje strumienia na pasku narzędzi.

Rysunek 18. Przycisk Pokaż/Ukryj adnotacje strumienia na pasku narzędzi

Wybór gałęzi oceniania do wdrożenia

Gałąź oceniania wyznacza się albo z menu podręcznego węzła końcowego, albo z menu Narzędzia. W wypadku użycia menu podręcznego węzeł oceniania jest automatycznie ustawiany na karcie Wdrożenie we właściwościach strumienia.

Aby wyznaczyć gałąź na gałąź oceniania (za pomocą menu podręcznego):

1. Połącz model użytkowy z węzłem końcowym (węzłem przetwarzającym lub wynikowym za modelem użytkowym).
2. Kliknij węzeł końcowy prawym przyciskiem myszy.
3. W menu kliknij opcję **Użyj jako gałęzi oceniania**.

Aby wyznaczyć gałąź na gałąź oceniania (za pomocą menu Narzędzia):

1. Połącz model użytkowy z węzłem końcowym (węzłem przetwarzającym lub wynikowym za modelem użytkowym).
2. W menu głównym kliknij opcje:

   Narzędzia > **Właściwości strumienia** > **Wdrożenie**
3. Na liście **Typ wdrożenia** kliknij opcję **Tylko scoring** albo **Odświeżenie modelu**, odpowiednio do potrzeb. Więcej informacji można znaleźć w temacie “Opcje wdrażania strumieni” na stronie 216.
4. Kliknij pole **Węzeł oceniający** i wybierz węzeł końcowy z listy.
5. Kliknij przycisk **OK**.

Odświeżanie modelu

Odświeżanie modelu to proces odbudowywania istniejącego modelu w strumieniu na podstawie nowszych danych. Sam strumień w repozytorium nie ulega zmianie. Na przykład nie zmienia się typ algorytmu ani ustawienia charakterystyczne dla strumienia, jednak model szkolony jest ponownie na nowych danych i przekazywany do repozytorium, jeśli nowa wersja modelu działa lepiej od starej.

Do odświeżenia można wyznaczyć tylko jednej model użytkowy w strumieniu — jest to tak zwany **model odświeżający**. Kliknięcie opcji **Odświeżenie modelu** na karcie Wdrożenie w właściwościach strumienia (patrz “Opcje wdrażania strumieni” na stronie 216) powoduje, że wyznaczony w danym momencie model użytkowy stanie się modelem odświeżającym. Model można także wyznaczyć do roli modelu odświeżającego, korzystając z menu podręcznego modelu użytkowego. Aby było to możliwe, model użytkowy musi już znajdować się w gałęzi oceniania.

Odebranie modelowi użytkowemu statusu „modelu odświeżającego” jest równoważne wybraniu typu Tylko scoring dla wdrożenia strumienia i powoduje odpowiednią aktualizację karty Wdrożenie w oknie dialogowym Właściwości strumienia. Można nadawać i odbierać ten status, korzystając z opcji **Użyj jako modelu odświeżającego** w menu podręcznym modelu użytkowego w bieżącej gałęzi ocenienia.

Usunięcie łącza modelu użytkowego w gałęzi oceniania również powoduje odebranie modelowi użytkowemu statusu „modelu odświeżającego”. Można cofnąć usunięcie łącza modelu za pomocą menu Edycja lub paska narzędzi; operacja ta powoduje także przywrócenie modelowi użytkowemu statusu „modelu odświeżającego”.

Sposób wyboru modelu odświeżającego

Podobnie jak gałąź oceniania łącze do modelu odświeżającego również jest wyróżnione w strumieniu. Jako model odświeżający wybrano model użytkowy i dlatego wyróżnienie łącza zależy od liczby modeli użytkowych w danym strumieniu.

Jeden model w strumieniu

Jeśli w momencie wskazania gałęzi jako gałęzi oceniania znajduje się w niej jeden powiązany model użytkowy, model ten stanie się modelem odświeżającym dla strumienia.

Wiele modeli w strumieniu

Jeśli w strumieniu znajduje się więcej niż jeden powiązany model użytkowy, model odświeżający jest wybierany w następujący sposób.

Jeśli model użytkowy został zdefiniowany na karcie Wdrożenie w oknie dialogowym właściwości strumienia i znajduje się również w strumieniu, wówczas ten model użytkowy staje się modelem odświeżającym.

Jeśli na karcie Wdrożenie nie zdefiniowano żadnego modelu użytkowego lub jeśli zdefiniowano jeden, ale nie znajduje się on w gałęzi oceniania, wówczas model użytkowy znajdujący się najbliżej węzła końcowego staje się modelem odświeżającym.

Jeśli kolejno usunięte zostanie zaznaczenie wszystkich łączy modelu stanowiących łącza odświeżające, wyróżniona będzie tylko gałąź oceniania (bez łączy). Typ wdrożenia jest ustawiony na Tylko scoring.

Uwaga: Dla jednego z łączy można ustawić status Zamień, ale nie żaden inny. W takim przypadku podczas wyznaczania gałęzi oceniania wybranym modelem użytkowym stanowiącym model odświeżający jest model zawierający łącze odświeżające i znajdujący się najbliżej węzła końcowego.

Brak modeli w strumieniu

Jeśli w strumieniu nie ma żadnych modeli lub istnieją tylko modele bez łączy, typ wdrożenia jest ustawiany na Tylko scoring.

Sprawdzanie, czy w gałęzi oceniania nie ma błędów

Gałąź wskazana jako gałąź oceniania jest sprawdzana w celu wykrycia ewentualnych błędów.

W razie wykrycia błędów gałąź oceniania jest wyróżniana kolorem błędu gałęzi oceniania i wyświetlany jest komunikat o błędzie. Korzystając z opcji Niestandardowe kolory, można określić kolor błędu gałęzi oceniania. Więcej informacji można znaleźć w temacie “Określanie opcji wyświetlania” na stronie 238.

W razie wykrycia błędów:

1. Wyeliminuj błąd, kierując się treścią komunikatu o błędzie.
2. W menu głównym kliknij opcje:

   Narzędzia > **Właściwości strumienia** > **Wdrożenie**

   i kliknij opcję **Sprawdź**.
3. W razie potrzeby powtarzaj ten proces aż do wyeliminowania wszystkich błędów.

Rozdział 12. Zapisywanie strumieni na serwerze IBM Cloud Pak for Data

Można zapisywać strumienie na serwerze IBM Cloud Pak for Data, na którym następnie można je otwierać jako przepływy i korzystać z wielu funkcji dostępnych w oprogramowaniu Cloud Pak for Data.

Cloud Pak for Data upraszcza zbieranie, organizowanie i analizowanie danych oraz zastosowanie sztucznej inteligencji w przedsiębiorstwie. Pozwala na bezproblemową integrację szerokiej gamy oprogramowania do pracy z danymi i sztuczną inteligencją, czołowych mechanizmów nadzoru i bezpieczeństwa oraz zunifikowanego środowiska pracy zespołowej. Aby uzyskać więcej informacji, zobacz https://www.ibm.com/products/cloud-pak-for-data. Użytkownicy zainteresowani rozwiązaniem Cloud Pak for Data mogą kontaktować się z przedstawicielem IBM.

Aby skonfigurować połączenie z serwerem

W tej instrukcji przyjęto założenie, że używana jest przeglądarka Firefox. W innej przeglądarce może być konieczne wykonanie innych czynności.

1. Otwórz program Firefox, wklej adres URL rozwiązania Cloud Pak for Data do paska adresu i naciśnij klawisz Enter.

   Adres URL można uzyskać, otwierając klienta WWW Cloud Pak for Data i odczytując adres URL z paska adresu przeglądarki (jest to cały adres aż do łańcucha `/zen`, bez tego łańcucha). Domyślnie adres URL przyjmuje postać `https://<przestrzeń nazw>-cpd-<przestrzeń nazw>.apps.<poddomena klastra>`

   gdzie:

   - `<przestrzeń nazw>` jest przestrzenią nazw (np. `ab123`).
   - `<poddomena klastra>` jest nazwą poddomeny klastra, na przykład `ocp454x86-aqt-mycompany.com`)

   W tym przykładzie adres URL to `https://ab123-cpd-ab123.apps.ocp454x86-aqt-mycompany.com`

   Więcej informacji o adresie URL rozwiązania Cloud Pak for Data zawiera dokumentacja produktu Cloud Pak for Data.
2. Kliknij żółtą ikonę kłódki po lewej stronie paska adresu, a następnie kliknij strzałkę obok tekstu **Niezabezpieczone połączenie**.
3. Kliknij opcję **Więcej informacji**. Zostanie otwarte okno dialogowe Informacje o stronie.
4. W oknie dialogowym Informacje o stronie przejdź na kartę **Bezpieczeństwo** i kliknij przycisk **Wyświetl certyfikat**.
5. Przewiń do sekcji **Różne** i pobierz plik PEM certyfikatu.
6. Z wiersza komend zaimportuj plik certyfikatu do katalogu klienta IBM SPSS Modeler w następujący sposób:

   W systemie Windows:

   ```
   "%JAVA_HOME%"\bin\keytool.exe -importcert -trustcacerts -file <Ścieżka bezwzględna
   do certyfikatu> -alias <alias> -keystore "<Ścieżka instalacji klienta
   Modeler>"\jre\lib\security\cacerts
   ```

   W systemie macOS:

   ```
   "%JAVA_HOME%"/bin/keytool -importcert -trustcacerts -file <Ścieżka bezwzględna do
   certyfikatu> -alias <alias> -keystore "<Ścieżka instalacji klienta Modeler>"/jre/lib/
   security/cacerts
   ```

Gdy pojawi się odpowiedni monit, wprowadź `changeit` jako hasło magazynu kluczy i wprowadź `yes`, aby zaufać certyfikatowi.

7. W menu głównym kliknij IBM SPSS Modeler:

   Narzędzia > **IBM Cloud Pak for Data Server** > **Połączenia z serwerem IBM Cloud Pak for Data...**

8. Określ ustawienia połączenia z serwerem. Użyj tego samego adresu URL, co w kroku 1. Jeśli nie znasz nazwy użytkownika lub hasła, skontaktuj się z lokalnym administratorem systemu. W przypadku wybrania logowania za pomocą klucza API należy użyć następującej metody w celu uzyskania adresu URL:

   - Zaloguj się do klastra Cloud Pak for Data i uruchom komendę `oc get routes -n ibm-common-services`, a następnie skopiuj wartość HOST/PORT dla `cp-console`. W poniższym przykładzie adres URL to `https://cp-console.apps.wml46x04.cp.myserver.com`:

     ```
     [root@wml46x04-inf ocp-scripts]# oc get routes -n ibm-common-services
     NAME         HOST/PORT                                   PATH   SERVICES
     PORT    TERMINATION           WILDCARD
     cp-console   cp-console.apps.wml46x04.cp.myserver.com          icp-management-ingress
     https   reencrypt/Redirect     None
     cp-proxy     cp-proxy.apps.wml46x04.cp.myserver.com            nginx-ingress-controller
     https   passthrough/Redirect   None
     ```

9. Kliknij opcje **Połącz**, aby połączyć się z serwerem.

Aby zapisać strumienie na serwerze Cloud Pak for Data

1. W menu głównym kliknij opcje:

   Plik > **Zapisz jako przepływ...**. *Strumienie* po migracji do Cloud Pak for Data nazywane są *przepływami*.

   Można też kliknąć prawym przyciskiem w menedżerze strumieni i wybrać opcję **Zapisz jako przepływ...**

2. Wybierz skonfigurowany wcześniej serwer Cloud Pak for Data. Następnie wybierz projekt, do którego chcesz zapisać przepływ.

Wskazówka: Można wybrać do zapisania więcej niż jeden strumień naraz. W menu głównym kliknij kolejno opcje **Narzędzia** > **Serwer IBM Cloud Pak for Data...** > **Masowe przesyłanie strumieni do serwera IBM Cloud Pak for Data**. Pojawi się monit o wybranie strumieni w lokalnym systemie plików, wybranie serwera i wybranie projektu.

Rozdział 13. Eksportowanie do aplikacji zewnętrznych

Informacje dotyczące eksportowania do aplikacji zewnętrznych

IBM SPSS Modeler oferuje pewną liczbę mechanizmów umożliwiających wyeksportowanie całego procesu eksploracji danych do aplikacji zewnętrznych, tak że wysiłek włożony w przygotowanie danych i utworzenie modeli może zostać wykorzystany również poza programem IBM SPSS Modeler.

W poprzedniej sekcji przedstawiono możliwości wdrażania strumieni w repozytorium IBM SPSS Collaboration and Deployment Services z myślą o umożliwieniu dostępu dla wielu użytkowników, planowaniu pracy i innych funkcjach. W podobny sposób strumienie IBM SPSS Modeler mogą być również używane w:

- IBM SPSS Modeler Advantage
- aplikacjach umożliwiających import i eksport plików w formacie PMML

Więcej informacji na temat korzystania ze strumieni w programie IBM SPSS Modeler Advantage zawiera temat “Otwieranie strumienia w programie IBM SPSS Modeler Advantage ” na stronie 223.

Informacje dotyczące eksportowania i importowania modelu jako plików PMML, co umożliwia współużytkowanie modelu z dowolną inną aplikacją obsługującą ten format, zawiera temat “Importowanie i eksportowanie modeli w formacie PMML” na stronie 224.

Otwieranie strumienia w programie IBM SPSS Modeler Advantage

Strumienie IBM SPSS Modeler mogą być używane w połączeniu z aplikacją cienkiego klienta IBM SPSS Modeler Advantage. Choć istnieje możliwość utworzenia dostosowanych aplikacji w całości w ramach programu IBM SPSS Modeler Advantage, można także użyć strumienia już utworzonego w programie IBM SPSS Modeler jako podstawy dla przepływu pracy aplikacji.

Aby otworzyć strumień w programie IBM SPSS Modeler Advantage:

1. Wdróż strumień w repozytorium IBM SPSS Collaboration and Deployment Services, koniecznie wybierając wcześniej opcję **Wdrażaj jako strumień**. Więcej informacji można znaleźć w temacie “Wdrażanie strumieni” na stronie 215.
2. Kliknij przycisk Otwórz na pasku narzędzi programu IBM SPSS Modeler Advantage lub z menu głównego, klikając opcje:

Plik > **Otwórz w IBM® SPSS® Modeler Advantage**

1. W razie potrzeby należy zdefiniować połączenia z repozytorium. Więcej informacji można znaleźć w temacie “Łączenie się z repozytorium” na stronie 204. W celu uzyskania danych konkretnego portu, hasła i innych szczegółowych danych połączenia należy skontaktować się z lokalnym administratorem systemu.

Uwaga: Na serwerze repozytorium musi być także zainstalowane oprogramowanie IBM SPSS Modeler Advantage.

1. W oknie dialogowym Repozytorium: Składuj wybierz folder, w którym chcesz składować obiekt, podaj pozostałe informacje, które chcesz zapisać, a następnie kliknij przycisk **Składuj**. Więcej informacji można znaleźć w temacie “Określanie właściwości obiektów” na stronie 205.

Postępowanie w ten sposób powoduje uruchomienie programu IBM SPSS Modeler Advantage z już otwartym strumieniem. Następnie strumień jest zamykany w programie IBM SPSS Modeler.

Importowanie i eksportowanie modeli w formacie PMML

PMML, czyli Predictive Model Markup Language, to format XML służący do opisu modeli eksploracji danych i modeli statystycznych, a w szczególności danych wejściowych modeli, transformacji używanych w celu przygotowania danych do eksploracji oraz parametrów definiujących same modele. IBM SPSS Modeler może importować i eksportować modele opisane w języku PMML, przez co pozwala na współużytkowanie modeli z innymi aplikacjami obsługującymi ten format, na przykład IBM SPSS Statistics.

Więcej informacji na temat języka PMML można znaleźć w serwisie WWW Data Mining Group (*http://www.dmg.org*).

Aby wyeksportować model

Większość typów modeli wygenerowanych w programie IBM SPSS Modeler można eksportować w języku PMML. Więcej informacji można znaleźć w temacie "Typy modeli obsługujące język PMML" na stronie 224.

1. Prawym przyciskiem myszy kliknij model użytkowy na palecie modeli. (Zamiast tego można kliknąć model użytkowy w obszarze roboczym i wybrać menu Plik).
2. W menu kliknij opcję **Eksportuj PMML**.
3. W oknie dialogowym eksportu (lub zapisywania) określ katalog docelowy i unikalną nazwę modelu.

Uwaga:

Opcje eksportu PMML można zmienić w oknie dialogowym Opcje użytkownika. W menu głównym kliknij opcje:

Narzędzia > **Opcje** > **Opcje użytkownika**

i kliknij kartę PMML.

Więcej informacji można znaleźć w temacie "Określanie opcji eksportowania PMML" na stronie 239.

Aby zaimportować model zapisany w formacie PMML

Modele wyeksportowane w formacie PMML z programu IBM SPSS Modeler lub innych aplikacji można importować do palety modeli. Więcej informacji można znaleźć w temacie "Typy modeli obsługujące język PMML" na stronie 224.

1. Kliknij paletę modeli prawym przyciskiem myszy i z menu wybierz polecenie **Importuj PMML**.
2. Wybierz plik do zaimportowania i określ opcje dotyczące etykiet zmiennych, odpowiednio do potrzeb.
3. Kliknij przycisk **Otwórz**.

Stosuj etykiety zmiennych. W pliku PMML mogą być określone zarówno nazwy, jak i etykiety zmiennych ze słownika danych (na przykład Identyfikator kierującego jako etykieta zmiennej *RefID*). Wybranie tej opcji spowoduje, że używane będą etykiety zmiennych, jeśli będą obecne w wyeksportowanym pliku PMML.

Jeśli wybrana jest opcja stosowania etykiet zmiennych, ale plik PMML nie zawiera etykiet zmiennych, to używane będą jak zwykle nazwy zmiennych.

Typy modeli obsługujące język PMML

Eksport PMML

Modele IBM SPSS Modeler. Następujące modele utworzone w programie IBM SPSS Modeler można eksportować w języku PMML 4.0:

- C&RT
- QUEST
- CHAID

- Sieć neuronowa
- C5.0
- Regresja logistyczna
- Genlin
- SVM
- Apriori
- Carma
- K-średnie
- Sieć Kohonena
- Dwustopniowa
- Dwustopniowa-AS
- GLMM (PMML jest eksportowany dla wszystkich modeli GLMM, ale PMML ma jedynie efekty stałe)
- Lista decyzyjna
- Coxa
- Sekwencja (ocenianie modeli PMML typu Sekwencja nie jest obsługiwane)
- Drzewa losowe
- Drzewo-AS
- Liniowy
- Liniowy-AS
- Regresja
- Regresja logistyczna
- GLE
- LSVM
- Wykrywanie anomalii
- KNN
- reguły asocjacyjne

Modele rodzime bazy danych. Spośród modeli wygenerowanych przez algorytmy rodzime bazy danych, eksport PMML jest niedostępny. Nie można eksportować modeli utworzonych za pomocą programu Analysis Services firmy Microsoft lub programu Oracle Data Miner.

Import modeli PMML

IBM SPSS Modeler może importować i oceniać modele PMML wygenerowane przez aktualne wersje wszystkich produktów IBM SPSS Statistics, w tym modele wyeksportowane z programu IBM SPSS Modeler oraz kod PMML modelu lub transformacji wygenerowany przez IBM SPSS Statistics w wersji 17.0 lub nowszej. Zasadniczo kryteria te spełnia każdy model PMML, który mechanizm oceniania jest w stanie ocenić, z następującymi wyjątkami:

- Apriori, CARMA, Wykrywanie anomalii, Sekwencje i Reguły asocjacyjne — tych modeli nie można importować.
- Modeli PMML nie można przeglądać po zaimportowaniu do programu IBM SPSS Modeler, mimo że mogą być używane w ocenianiu. (Uwaga: dotyczy to także modeli, które pierwotnie wyeksportowano z programu IBM SPSS Modeler. Aby uniknąć tego ograniczenia, należy wyeksportować model jako wygenerowany plik modelu [`*.gm`], a nie w języku PMML).
- Przy importowaniu poprawność jest sprawdzana w ograniczonym zakresie, ale przy próbie oceny modelu przeprowadzane jest pełne sprawdzanie poprawności. Dlatego może się zdarzyć, że mimo udanego importu ocenianie nie powiedzie się lub przyniesie nieprawidłowe wyniki.

Uwaga: W przypadku modeli PMML z innych programów zaimportowanych do programu IBM SPSS ModelerIBM SPSS Modeler podejmowana jest próba oceny poprawnego modelu PMML, który da się rozpoznać i ocenić. Nie ma jednak gwarancji, że wszystkie modele PMML dadzą się ocenić i że będą oceniane tak samo, jak w aplikacji, która je wygenerowała.

Rozdział 14. Projekty i raporty

Wprowadzenie do projektów

Projekt jest to grupa plików związanych z zadaniem eksploracji danych. Projekty zawierają strumienie danych, wykresy, wygenerowane modele, raporty i wszelkie inne elementy utworzone przez użytkownika w programie IBM SPSS Modeler. Na pierwszy rzut oka może się wydawać, że projekty IBM SPSS Modeler są po prostu środkiem organizacji wyników, jednak w rzeczywistości oferują o wiele większe możliwości. Korzystając z projektów, można:

- Dodawać adnotacje do poszczególnych obiektów w pliku projektu.
- Prowadzić eksplorację danych zgodnie z metodologią CRISP-DM. Projekty zawierają także system pomocy Metodologia CRISP-DM, który zawiera szczegółowe informacje i praktyczne przykłady eksploracji danych z wykorzystaniem modelu procesu CRISP-DM.
- Dodawać do projektu obiekty spoza programu IBM SPSS Modeler, takie jak pokazy slajdów PowerPoint służące do prezentowania celów eksploracji danych lub artykuły opisujące algorytmy, których zamierzamy użyć.
- Generować rozbudowane i proste raporty o zmianach na podstawie adnotacji użytkownika. Raporty te mogą być generowane w formacie HTML, tak by dało się je łatwo opublikować w intranecie organizacji użytkownika.

Uwaga: Jeśli panel projektów nie jest widoczny w oknie programu IBM SPSS Modeler, kliknij opcję **Projekt** w menu Widok.

Obiekty dodane do projektu można przeglądać na dwa sposoby: **w widoku Klasy** i w **widoku CRISP-DM**. Wszystkie obiekty dodawane do projektu pojawiają się w obu widokach i można przełączać się między widokami tak, by optymalnie zorganizować sobie pracę.

Widok CRISP-DM

Dzięki zgodności z metodologią Cross-Industry Standard Process for Data Mining (CRISP-DM), projekty w programie IBM SPSS Modeler umożliwiają organizację komponentów procesu eksploracji danych w sposób zgodny ze sprawdzonymi procedurami branżowymi i nieprzypisany do konkretnego narzędzia. W metodologii CRISP-DM proces podzielony jest na sześć faz: od określenia wymagań biznesowych aż po wdrożenie wyników. Mimo że program IBM SPSS Modeler zwykle jest używany tylko w niektórych z tych faz, na panelu projektu uwzględniono wszystkie sześć faz, dzięki czemu mamy do dyspozycji jedno, centralne środowisko przechowywania i dokumentowania wszystkich materiałów związanych z projektem. Na przykład w fazie Zapoznanie z zadaniem zwykle odbywa się szereg spotkań w celu sformułowania wymagań i celów. Program IBM SPSS Modeler raczej nie jest używany w tej fazie do pracy z danymi. Jednak na panelu projektu można zapisywać notatki z takich spotkań w folderze *Zapoznanie z zadaniem* i w przyszłości wracać do nich i uwzględniać je w raportach.

Widok CRISP-DM w panelu projektu wyposażony jest także w osobny system pomocy dotyczący kolejnych faz eksploracji danych. W programie IBM SPSS Modeler dostęp do tego systemu pomocy można uzyskać, wybierając opcje **Metodologia CRISP-DM** z menu Pomoc.

Uwaga: Jeśli panel projektów nie jest widoczny w oknie, kliknij opcję **Projekt** w menu Widok.

Ustawianie domyślnej fazy projektu

Obiekty dodawane do projektu są dodawane do domyślnej fazy CRISP-DM. Oznacza to, że konieczne jest uporządkowanie obiektów ręcznie, zgodnie z fazą eksploracji danych, w której były one używane. Przydatne jest ustawienie folderu domyślnego na fazę, w której użytkownik obecnie pracuje.

Aby wybrać fazę, która ma być używana jako domyślna:

1. W widoku CRISP-DM kliknij prawym przyciskiem folder fazy, aby ustawić go jako domyślny.

2. Kliknij pozycję **Ustaw jako domyślny** w menu.

Folder domyślny jest wyświetlany jako pogrubiony.

Widok Klasy

Widok Klasy w panelu projektu umożliwia organizowanie pracy wg kategorii w oprogramowaniu IBM SPSS Modeler, czyli wg typów tworzonych obiektów. Zapisywane obiekty można dodawać do dowolnych spośród następujących kategorii:

- Strumienie
- Węzły
- Modele
- Tabele, wykresy i raporty
- Inne pliki (spoza programu IBM SPSS Modeler), takie jak pokazy slajdów lub artykuły związane z eksploracją danych

Dodanie obiektu do widoku Klasy powoduje również dodanie go do domyślnego folderu danej fazy w widoku CRISP-DM.

Uwaga: Jeśli panel projektów nie jest widoczny w oknie, kliknij opcję **Projekt** w menu Widok.

Tworzenie projektu

Projekt to zasadniczo plik zawierający odwołania do wszystkich plików powiązanych z projektem. Oznacza to, że elementy projektu są zapisywane zarówno indywidualnie, jak i w postaci odwołania w pliku projektu (*.cpj*). Ze względu na tę strukturę należy pamiętać o następujących ograniczeniach:

- Przede wszystkim, każdy z elementów projektu wymaga zapisania przed dodaniem go do bieżącego projektu. W przypadku niezapisania elementu zostanie wyświetlony monit o zapisanie go przed dodaniem do bieżącego projektu.
- Obiekty aktualizowane z osobna, takie jak strumienie, są również aktualizowane w pliku projektu.
- Ręczne przenoszenie i usuwanie obiektów (takich jak strumienie, węzły i obiekty wynikowe) z systemu plików spowoduje powstanie nieprawidłowych odsyłaczy w pliku projektu.

Tworzenie nowego projektu

W oknie IBM SPSS Modeler można w łatwy sposób tworzyć nowe projekty. Można albo rozpocząć tworzenie projektu (jeśli żaden nie jest otwarty), lub można zamknąć istniejący projekt i rozpocząć od zera.

W menu głównym kliknij opcje:

Plik > **Projekt** > **Nowy projekt...**

Dodawanie do projektu

Po utworzeniu lub otwarciu projektu można, korzystając z kilku metod, dodać obiekty takie jak strumienie danych, węzły czy raporty.

Dodawanie obiektów z obszarów Zarządzanie

Korzystając z obszarów Zarządzanie w prawym górnym rogu okna IBM SPSS Modeler, można dodawać strumienie lub dane wynikowe.

1. Wybierz obiekt taki jak tabela lub strumień z jednej z kart Zarządzanie.
2. Kliknij prawym przyciskiem myszy, a następnie kliknij opcję **Dodaj do projektu**.

   Jeśli obiekt został wcześniej zapisany, zostanie on automatycznie dodany do odpowiedniego folderu obiektów (w widoku Klasy) lub do domyślnego folderu obiektów (w widoku CRISP-DM).

3. Można też też przeciągnąć i upuścić obiekty z obszaru Zarządzanie do panelu projektu.

Uwaga: może zostać wyświetlony komunikat z prośbą o zapisanie obiektu. Podczas zapisywania należy upewnić się, że opcja **Dodaj plik do projektu** w oknie dialogowym Zapisz jest zaznaczona. Spowoduje to automatyczne dodanie obiektu do projektu po jego zapisaniu.

Dodawanie węzłów z obszarów roboczych

Istnieje możliwość dodania poszczególnych węzłów z obszaru roboczego strumienia za pomocą okna dialogowego Zapisz.

1. Wybierz węzeł w obszarze roboczym.
2. Kliknij prawym przyciskiem myszy, a następnie kliknij opcję **Zapisz węzeł**. Inny sposób: w menu głównym kliknij:

   Edycja > **Węzeł** > **Zapisz węzeł...**
3. W oknie dialogowym Zapisz wybierz pozycję **Dodaj plik do projektu**.
4. Wprowadź nazwę węzła, a następnie kliknij przycisk **Zapisz**.

Spowoduje to zapisanie pliku i dodanie go do projektu. Węzły są dodawane do folderu *Węzły* w widoku Klasy oraz do folderu domyślnego fazy w widoku CRISP-DM.

Dodawanie plików zewnętrznych

Istnieje możliwość dodania do projektu szeregu obiektów innych niż pochodzące z programu IBM SPSS Modeler. Jest to szczególnie przydatne w przypadku zarządzania całym procesem eksploracji danych z poziomu programu IBM SPSS Modeler. Możliwe jest na przykład przechowywanie w projekcie łączy do danych, uwag, prezentacji i grafiki. W widoku CRISP-DM pliki zewnętrzne można dodawać do dowolnego folderu. W widoku Klasy pliki zewnętrzne można zapisywać wyłącznie w folderze *Inne*.

Aby dodać pliki zewnętrzne do projektu:

1. Przeciągnij pliki z pulpitu do projektu.

   lub
2. Kliknij prawym przyciskiem myszy folder w widoku CRISP-DM lub Klasy.
3. W menu kliknij opcję **Dodaj do folderu**.
4. Wybierz plik w oknie dialogowym i kliknij przycisk **Otwórz**.

Spowoduje to dodanie odwołania do wybranego obiektu w projekcie IBM SPSS Modeler.

Przenoszenie projektów do repozytorium IBM SPSS Collaboration and Deployment Services Repository

Istnieje możliwość przeniesienia w jednym kroku całego projektu, wraz z plikami wszystkich składników, do repozytorium IBM SPSS Collaboration and Deployment Services Repository. Obiekty znajdujące się już w lokalizacji docelowej nie zostaną przeniesione. Funkcja ta działa również kierunku odwrotnym: można przenieść całe projekty z repozytorium IBM SPSS Collaboration and Deployment Services Repository do lokalnego systemu plików.

Przenoszenie projektu

Upewnij się, że projekt, który ma zostać przeniesiony, jest otwarty w panelu projektu.

Aby przenieść projekt:

1. Kliknij prawym przyciskiem myszy główny folder projektu, a następnie kliknij opcję **Przenieś projekt**.
2. Po wyświetleniu monitu zaloguj się do repozytorium IBM SPSS Collaboration and Deployment Services Repository.
3. Wskaż nową lokalizację projektu i kliknij przycisk **OK**.

Ustawianie właściwości projektu

Zawartość i dokumentację projektu można dostosować do swoich potrzeb, korzystając z okna dialogowego właściwości projektu. Aby uzyskać dostęp do właściwości projektu:

1. Kliknij prawym przyciskiem myszy obiekt lub folder w panelu projektu, a następnie kliknij opcję **Właściwości projektu**.
2. Kliknij kartę **Projekt**, aby wskazać podstawowe informacje o projekcie.

Utworzono. Wyświetla datę utworzenia projektu (dane nieedytowalne).

Podsumowanie. Możliwe jest wprowadzenie podsumowania do projektu eksploracji danych, które mają zostać wyświetlone w raporcie projektu.

Zawartość. Typy i liczby komponentów, do których istnieją odwołania z pliku projektu (dane nieedytowalne).

Zapisz niezapisane obiekty jako. Określa, czy niezapisane obiekty powinny być zapisywane w lokalnym systemie plików, czy też przechowywane w repozytorium. Więcej informacji można znaleźć w temacie “Informacje o programie IBM SPSS Collaboration and Deployment Services Repository ” na stronie 203.

Aktualizuj odwołania obiektów podczas ładowania projektu. Wybierz tę opcję, aby zaktualizować odwołania projektu odpowiednio do jego składników. *Uwaga*: pliki dodane do projektu nie są zapisywane w samym pliku projektu. W projekcie zapisywane jest odwołanie do plików. Oznacza to, że przeniesienie lub usunięcie pliku spowoduje usunięcie tego obiektu z projektu.

Dodawanie adnotacji do projektu

Panel projektu umożliwia dodawanie adnotacji do wyników eksploracji danych na kilka sposobów. Adnotacje na poziomie projektu są często używane do śledzenia celów i decyzji w ujęciu ogólnym, podczas gdy adnotacje folderów lub plików zawierają więcej szczegółów. Na karcie Adnotacje znajduje się wystarczająca ilość miejsca na szczegóły na poziomie projektu, takie jak informacje na temat wyłączenia danych zawierających niemożliwe do odzyskania braki danych czy informacje o pomyślnie rokujących hipotezach utworzonych podczas eksploracji danych.

Aby dodać adnotację do projektu:

1. Wybierz folder projektu w widoku CRISP-DM lub w widoku Klasy.
2. Kliknij folder prawym przyciskiem myszy i kliknij opcję **Właściwości projektu**.
3. Kliknij kartę **Adnotacje**.
4. Wprowadź słowa kluczowe i tekst opisu projektu.

Właściwości folderu i adnotacje

Nie jest możliwe wprowadzenie adnotacji dla pojedynczych folderów projektów (zarówno w widoku CRISP-DM jak i w widoku Klasy). W widoku CRISP-DM szczególnie efektywne może okazać się dokumentowanie celów organizacji dla każdej fazy eksploracji danych. Na przykład, korzystając z narzędzia do adnotacji w folderze *Zapoznanie z zadaniem*, można wprowadzić elementy opisowe, takie jak „Celem biznesowym tego badania jest zredukowanie poziomu odejścia wśród najbardziej wartościowych klientów.” Tekst ten może następnie zostać automatycznie umieszczony w raporcie projektu po wybraniu opcji **Uwzględnij w raporcie**.

Aby wprowadzić adnotacje do folderu:

1. Wybierz folder w panelu projektu.
2. Kliknij folder prawym przyciskiem myszy i kliknij opcję **Właściwości folderu**.

W widoku CRISP-DM adnotacje do folderów są wprowadzane wraz z Podsumowaniem celów dla każdej fazy, a także wytycznymi co do ukończenia odpowiednich zadań eksploracji danych. Adnotacje można usuwać i edytować.

Nazwa. W tym obszarze wyświetlana jest nazwa wybranej zmiennej.

Tekst podpowiedzi. Istnieje możliwość tworzenia własnych podpowiedzi, które są następnie wyświetlane po wskazaniu myszą folderu projektu. Jest to szczególnie użyteczne w przypadku widoku CRISP-DM, na przykład w celu zapewnienia szybkiego przeglądu celów każdej z faz lub w celu oznaczenia statusu danej fazy, takiego tak „W toku” lub „Zakończona”.

Zmienna Adnotacja. Tej zmiennej można użyć w przypadku dłuższych adnotacji, które mogą być następnie zestawiane w raporcie projektu. Widok CRISP-DM zawiera opisy każdej z faz eksploracji danych umieszczone w adnotacjach, lecz użytkownik może je swobodnie dostosowywać odpowiednio do potrzeb projektu.

Uwzględnij w raporcie. W celu uwzględnienia adnotacji w raportach należy wybrać opcję **Uwzględnij w raporcie**.

Właściwości obiektu

Istnieje możliwość wyświetlenia właściwości obiektu i decydowania, czy poszczególne obiekty mają być uwzględniane w raporcie projektu. Aby uzyskać dostęp do właściwości obiektu:

1. Kliknij prawym przyciskiem obiekt w panelu projektu.
2. W menu kliknij opcję **Właściwości obiektu**.

Nazwa. W tym obszarze wyświetlana jest nazwa zapisanego obiektu.

Ścieżka. W tym obszarze wyświetlana jest lokalizacja zapisanego obiektu.

Uwzględnij w raporcie. Wybierz tę opcję, aby uwzględnić w wygenerowanym raporcie szczegóły obiektu.

Zamykanie projektu

Z chwilą zamknięcia programu IBM SPSS Modeler lub otwarcia nowego projektu plik obecnego projektu (*.cpj*) jest zamykany.

Niektóre pliki skojarzone z projektem (takie jak strumienie, węzły czy wykresy) mogą być nadal otwarte. Aby pozostawić te pliki otwarte, odpowiedz `Nie` na komunikat **... Czy chcesz zapisać i zamknąć te pliki?**

W przypadku zmodyfikowania i zapisania wszelkich powiązanych plików po zamknięciu projektu ich zaktualizowane wersje zostaną ujęte w projekcie przy jego następnym otwarciu. Aby wyeliminować tę domyślne działanie, usuń plik z projektu lub zapisz go pod inną nazwą pliku.

Generowanie raportu

Jedną z najbardziej przydatnych funkcji projektów jest możliwość generowania raportów na podstawie elementów projektu i adnotacji. Jest to element krytyczny dla wyników eksploracja danych, a jego rolę opisano dokładnie w sekcji dotyczącej metodologii CRISP-DM. Raport można wygenerować bezpośrednio do jednego z kilku typów plików lub do okna wyników na ekranie, gdzie można się z nim od razu zapoznać. Z tego miejsca można wydrukować, zapisać lub wyświetlić raport w przeglądarce WWW. Zapisane raporty można przesyłać do innych użytkowników w organizacji.

Raporty na podstawie plików projektów są często generowane kilkakrotnie w trakcie procesu eksploracji danych w celu ich przesłania do użytkowników zaangażowanych w projekt. Raport zawiera wybrane informacje na temat obiektów, do których istnieją odwołania z pliku projektu, a także wszelkie utworzone adnotacje. Istnieje możliwość tworzenia raportów w oparciu o widok Klasy lub widok CRISP-DM.

Aby wygenerować raport:

1. Wybierz folder projektu w widoku CRISP-DM lub w widoku Klasy.
2. Kliknij folder prawym przyciskiem myszy i kliknij opcję **Raport projektu**.
3. Określ opcje raportu i kliknij przycisk **Generuj raport**.

Opcje w oknie dialogowym raportu oferują szereg sposobów na wygenerowanie potrzebnego typu raportu:

Nazwa wyniku. Podaj nazwę okna wyników, jeśli chcesz wysłać wynikowy raport na ekran. Możesz też określić nazwę niestandardową lub pozwolić, aby program IBM SPSS Modeler nadał nazwę temu oknu.

Wynik na ekran. Wybierz tę opcję, aby wygenerować i wyświetlić raport w oknie wyników. Należy przy tym pamiętać, że istnieje możliwość wyeksportowania raportu z poziomu okna wyników do pliku jednego z kilku typów.

Wynik do pliku. Wybierz tę opcję, aby wygenerować i zapisać raport jako plik typu wskazanego na liście Typ pliku.

Nazwa pliku. Podaj nazwę pliku wygenerowanego raportu. Pliki są zapisywane domyślnie w katalogu IBM SPSS Modeler *bin*. Skorzystaj z przycisku z wielokropkiem **(...)** w celu wskazania innej lokalizacji.

Typ pliku. Dostępne typy plików to:

- **Dokument HTML.** Raport jest zapisywany jako pojedynczy plik HTML. Jeśli raport zawiera wykresy, są one zapisywane jako pliki PNG i są to nich tworzone odwołania z pliku HTML. Podczas publikowania raportu w Internecie należy zwrócić uwagę, aby przesłane zostały zarówno plik HTML jak i wszelkie obrazy, do których zawiera on odwołania.
- **Dokument tekstowy.** Raport jest zapisywany jako pojedynczy plik tekstowy. Jeśli raport zawiera wykresy, są w nim zawarte wyłącznie nazwy plików i ścieżki.
- **Dokument programu Microsoft Word.** Raport jest zapisywany jako pojedynczy plik, zaś wykresy są osadzone bezpośrednio w dokumencie.
- **Dokument programu Microsoft Excel.** Raport jest zapisywany jako pojedynczy arkusz kalkulacyjny, zaś wykresy są osadzone bezpośrednio w arkuszu.
- **Dokument programu Microsoft PowerPoint.** Każda faza jest przedstawiana na osobnym slajdzie. Wszelkie wykresy są osadzone bezpośrednio na slajdach dokumentu PowerPoint.
- **Obiekt wynikowy.** Po otwarciu w programie IBM SPSS Modeler ten plik (*.cou*) odpowiada opcji **Wynik na ekran** w grupie **Format raportu**.

Uwaga: Aby możliwe było eksportowanie danych do pliku programu Microsoft Office, konieczne jest dysponowanie odpowiednią aplikacją.

Tytuł. Podaj tytuł raportu.

Struktura raportu. Wybierz opcję **CRISP-DM** lub **Klasy**. W widoku CRISP-DM dostępny jest raport o statusie zawierające dane podsumowujące w ujęciu ogólnym, a także szczegółowe informacje na temat poszczególnych faz eksploracji danych. Widok Klasy to widok w oparciu o obiekty znacznie bardziej odpowiedni do wewnętrznej pracy z danymi i strumieniami.

Autor. Wyświetlana jest domyślna nazwa użytkownika, można ją jednak zmienić.

Raport obejmuje. Wybierz metodę uwzględnienia obiektów w raporcie. Wybierz **wszystkie foldery i obiekty**, aby uwzględnić wszystkie elementy dodane do pliku projektu. Możesz także uwzględnić elementy w oparciu o to, czy we właściwościach obiektu zaznaczono opcję **Uwzględnij w raporcie**. Aby sprawdzić, które elementy nie zostaną uwzględnione w raporcie, można też wybrać uwzględnienie tylko elementów oznaczonych do wykluczenia (tych, dla których nie zaznaczono opcji **Uwzględnij w raporcie**).

Selekcja. Ta opcja pozwala na udostępnianie aktualizacji projektu po wybraniu w raporcie tylko pola **ostatnie elementy**. Można też śledzić starsze i prawdopodobnie nierozwiązane problemy, ustawiając parametry dla pola **stare elementy**. Wybierz **wszystkie elementy**, aby odrzucić czas jako parametr raportu.

Porządkuj według. Możliwy jest wybór kombinacji następujących cech obiektów w celu posortowania ich w ramach folderu:

- **Typ.** Grupuje obiekty wg typu.
- **Nazwa.** Pozwala porządkować obiekty alfabetycznie.
- **Data dodania.** Umożliwia sortowanie obiektów wg daty dodania ich do projektu.

Zapisywanie i eksportowanie wygenerowanych raportów

Raport wygenerowany na ekran jest wyświetlany w nowym oknie wyników. Wszelkie wykresy uwzględnione w raporcie są wyświetlane jako obrazy wstawiane

Terminologia używana w raporcie

W raporcie podawana jest łączna liczba węzłów zawartych w każdym strumieniu. Liczby są wyświetlane pod nagłówkami, w których stosowana jest terminologia IBM SPSS Modeler, nie zaś terminologia CRISP-DM:

- **Data readers.** Węzły źródłowe.
- **#Data writers.** Węzły eksportu.
- **Model builders.** Węzły budowy lub modelowania.
- **Model appliers.** Wygenerowane modele, zwane również modelami użytkowymi.
- **Output builders.** Węzły Wykres lub Dane wynikowe
- **Inne.** Wszelkie inne węzły powiązane z projektem. Na przykład te dostępne na karcie Zmienne lub Rekordy na Palecie węzłów.

Aby zapisać raport:

1. W menu Plik kliknij pozycję **Zapisz**.
2. Podaj nazwę pliku.

   Raport jest zapisywany jako obiekt wynikowy.

   Aby wyeksportować raport:
3. W menu Plik kliknij pozycję **Eksportuj** i podaj typ pliku, do jakiego mają zostać wyeksportowane dane.
4. Podaj nazwę pliku.

Raport zostanie zapisany w wybranym formacie.

Istnieje możliwość wyeksportowania następujących typów plików:

- HTML
- Tekst
- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint

Uwaga: Aby możliwe było eksportowanie danych do pliku programu Microsoft Office, konieczne jest dysponowanie odpowiednią aplikacją.

Użyj przycisków w górnej części okna, aby:

- Wydrukować raport.
- Wyświetlić raport jako plik HTML w zewnętrznej przeglądarce WWW.

Rozdział 15. Dostosowywanie produktu IBM SPSS Modeler

Dostosowywanie opcji produktu IBM SPSS Modeler

Istnieje kilka operacji, których przeprowadzenie pozwala dostosować produkt IBM SPSS Modeler do własnych potrzeb. Przede wszystkim taka operacja dostosowywania polega na ustawieniu wybranych opcji użytkownika, takich jak alokacja pamięci, katalogi domyślne oraz użycie dźwięków i kolorów. Można także dostosować paletę węzłów znajdującą się w dolnej części okna IBM SPSS Modeler.

Ustawianie opcji programu IBM SPSS Modeler

Opcje IBM SPSS Modeler można dostosować i skonfigurować na kilka sposobów:

- Opcje systemowe, takie jak zużycie pamięci i ustawienia narodowe, można skonfigurować, klikając pozycję **Opcje systemowe** dostępną w menu **Narzędzia > Opcje**.
- Opcje użytkownika, takie jak czcionki i kolory, można skonfigurować, klikając pozycję **Opcje użytkownika** dostępną w menu **Narzędzia > Opcje**.
- Lokalizację aplikacji współpracujących z oprogramowaniem IBM SPSS Modeler można określić, klikając pozycję **Aplikacje pomocnicze** w menu **Narzędzia > Opcje**.
- Katalogi domyślne używane w ramach oprogramowania IBM SPSS Modeler można określić, klikając pozycję **Katalog klienta** lub **Katalog serwera** w menu Plik.

Istnieje również możliwość ustawienia opcji dotyczących niektórych lub wszystkich strumieni. Więcej informacji można znaleźć w temacie “Ustawianie opcji strumieni” na stronie 41.

Opcje systemowe

W języku IBM SPSS Modeler można wybrać język lub ustawienia regionalne. W tym celu należy kliknąć pozycję **Opcje systemowe** dostępną w menu **Narzędzia > Opcje**. Można tutaj określić także limit wykorzystania pamięci wewnętrznej przez program SPSS Modeler i częstotliwość automatycznego zapisywania strumieni. Zmiany wprowadzone w tym oknie dialogowym zostaną zastosowane dopiero po ponownym uruchomieniu oprogramowania SPSS Modeler.

Maksimum pamięci. Wybierz, aby określić limit wykorzystania pamięci przez program IBM SPSS Modeler (w MB). Na niektórych platformach oprogramowanie SPSS Modeler ogranicza zasoby zużywane przez proces w celu zmniejszenia obciążenia komputerów dysponujących ograniczoną wydajnością. Jeśli użytkownik obsługuje duże ilości danych, to może to być przyczyną wystąpienia błędu braku pamięci. Można ograniczyć zużycie pamięci, określając nową wartość progową.

Przykładowo: próba wyświetlenia bardzo dużego drzewa decyzyjnego może spowodować wystąpienie błędu dotyczącego braku pamięci. W takim przypadku zaleca się zwiększenie rozmiaru pamięci do wielkości maksymalnej: 4096 MB. W sytuacjach kiedy przetwarzane będą bardzo duże ilości danych, po zwiększeniu ilości dostępnej pamięci i wyłączeniu oprogramowania SPSS Modeler program ten należy uruchomić ponownie z poziomu wiersza komend w celu zapewnienia, że maksymalna ilość pamięci będzie używana podczas przetwarzania danych.

Aby uruchomić oprogramowanie z poziomu wiersza komend (zakładając, że oprogramowanie SPSS Modeler jest zainstalowane w położeniu domyślnym), w wierszu komend należy wpisać następujący łańcuch:

```
C:\Program Files\IBM\SPSS\Modeler\18.3.0\bin\modelerclient.exe" -J-Xss4096M
```

Użyj ustawień regionalnych systemu. Ta opcja jest wybrana domyślnie, a domyślnym ustawieniem jest English (United States). Usuń zaznaczenie tej opcji, aby na liście wybrać inny język i ustawienia regionalne.

Interwał automatycznego zapisywania strumienia (minuty). Określ, jak często SPSS Modeler ma automatycznie zapisywać strumienie. Wartość maksymalna to 60 minut, minimalna to 1 minuta, a domyślna to 5 minut.

Zarządzanie pamięcią

Oprócz ustawienia **Maksimum pamięci** w oknie dialogowym Opcje systemowe dostępnych jest kilka innych sposobów optymalizacji wykorzystania pamięci:

- Można zmodyfikować wartość **Maksimum członków dla zmiennych nominalnych** w oknie dialogowym właściwości strumienia. Opcja ta określa maksymalną liczbę członków dla zmiennych nominalnych, po przekroczeniu której poziom pomiaru zmiennej zmieniony zostanie na *Bez typu*. Więcej informacji można znaleźć w temacie "Określanie opcji ogólnych strumienia" na stronie 42.
- Można nakazać programowi IBM SPSS Modeler zwolnienie pamięci, klikając w prawym dolnym rogu okna, w miejscu gdzie wyświetlana jest ilość pamięci używana przez program IBM SPSS Modeler oraz ilość przydzielona (*xx*MB / *xx*MB). Kliknięcie tego obszaru spowoduje, że przyjmie on ciemniejszy odcień, a następnie wyświetlone ilości przydzielonej pamięci spadną. Powrót do zwykłego koloru oznacza, że program IBM SPSS Modeler zwolnił tyle pamięci, ile tylko było można.

Ustawianie katalogów domyślnych

Można określić katalog domyślny używany dla przeglądarek plików i danych wynikowych, wybierając z menu Plik opcje **Katalog klienta** lub **Katalog serwera**.

- **Katalog klienta.** Tej opcji można użyć do ustawienia katalogu roboczego. Domyślny katalog roboczy zależy od ścieżki instalacji wersji programu IBM SPSS Modeler użytkownika lub od ścieżki w wierszu komend użytej do uruchomienia programu IBM SPSS Modeler. W trybie lokalnym katalog roboczy to jednocześnie ścieżka używana dla wszystkich operacji po stronie klienta i plików wynikowych (jeśli utworzono odwołania do nich za pomocą ścieżek względnych).
- **Katalog serwera.** Opcja Katalog serwera w menu Plik jest włączona zawsze wtedy, gdy dostępne jest zdalne połączenie z serwerem. Tej opcji można użyć do określenia katalogu domyślnego dla wszystkich plików serwera i plików danych określonych dla danych wejściowych lub wyjściowych. Domyślny katalog serwera to *$CLEO/data*, gdzie *$CLEO* to katalog, w którym zainstalowana jest wersja serwera programu IBM SPSS Modeler. Korzystając z wiersza komend, można także zastąpić tę wartość domyślną flagą `-server_directory` z argumentem wiersza komend `modelerclient`.

Określanie opcji użytkownika

Opcje ogólne IBM SPSS Modeler można określić, wybierając pozycję **Opcje użytkownika** w menu **Narzędzia > Opcje**. Te opcje są stosowane do wszystkich strumieni używanych w oprogramowaniu IBM SPSS Modeler.

Następujące typy opcji można skonfigurować, klikając odpowiednie karty:

- Opcje powiadomień, np. dotyczące nadpisywania i komunikatów o błędach.
- Opcje wyświetlania, np, dotyczące wykresów i kolorów tła.
- Opcje wyświetlania kolorów składni.
- Opcje eksportowania PMML używane podczas eksportowania modeli do formatu PMML (Predictive Model Markup Language).
- Informacje o użytkowniku, np. nazwisko, inicjały i adres e-mail. Te informacje można wyświetlić na karcie Adnotacje w przypadku węzłów i innych tworzonych obiektów.
- Przełączanie między trybem tradycyjnym a trybem serwera Analytic Server.

Aby określić opcje strumienia, np. separatory dziesiętne, formaty daty i czasu, opcje optymalizacji, układu strumienia i skryptów, użyj okna dialogowego właściwości strumienia dostępnego z poziomu menu Plik i Narzędzia.

Określanie opcji powiadomień

Za pomocą karty Powiadomienia dostępnej w oknie dialogowym Opcje użytkownika można określić różnorodne opcje dotyczące wystąpień i typów ostrzeżeń oraz okien potwierdzeń w oprogramowaniu IBM SPSS Modeler. Użytkownik może także określić zachowanie kart Wyniki i Modele w panelu zarządzania podczas tworzenia nowych wyników i modeli.

Pokaż okno dialogowe informacji zwrotnych z wykonywania strumienia Wybranie tej opcji powoduje wyświetlenie okna dialogowego zawierającego wskaźnik postępu, gdy strumień jest wykonywany przez ponad trzy sekundy. W tym oknie dialogowym są również widoczne obiekty wynikowe utworzone przez strumień.

- **Zamknij to okno po zakończeniu działania** Domyślnie to okno dialogowe jest zamykane po zakończeniu wykonywania strumienia. Odznacz to pole wyboru, aby po zakończeniu wykonywania strumienia okno dialogowe pozostało widoczne.

Ostrzegaj, gdy węzeł nadpisuje plik Zaznacz to pole wyboru, aby wyświetlać komunikaty o błędach, gdy operacje wykonywane na węźle powodują zastąpienie istniejących plików.

Ostrzegaj przed nadpisywaniem tabeli w bazie danych Zaznacz to pole wyboru, aby wyświetlać komunikaty o błędach, gdy operacje wykonywane na węźle powodują zastąpienie istniejącej tabeli bazy danych.

Powiadomienia dźwiękiem

Za pomocą tej listy można określić, czy błędy lub zdarzenia są sygnalizowane dźwiękowo. Dostępnych jest kilka rodzajów dźwięków. Aby odtworzyć wybrany dźwięk, należy kliknąć przycisk odtwarzania (ikona głośnika). Kliknięcie przycisku wielokropka **(...)** pozwala na wyszukanie dźwięku.

Uwaga: Pliki `.wav`, na podstawie których utworzono dźwięki IBM SPSS Modeler, znajdują się w katalogu `/media/sounds` w folderze instalacyjnym oprogramowania.

- **Wycisz wszystkie dźwięki** Wybierz tę opcję, aby wyłączyć wszystkie powiadomienia dźwiękowe.

Powiadomienia wizualne

Opcje w tej grupie służą do określania zachowania kart Wyniki i Modele w panelu zarządzania (prawa górna część ekranu) podczas tworzenia nowych elementów. Na liście wybierz pozycje **Nowy model** lub **Nowy wynik**, aby określić zachowanie odpowiedniej karty.

Ta opcja jest dostępna dla ustawienia **Nowy model**:

Zastąp poprzedni model Jeśli ta opcja jest wybrana (domyślnie), model istniejący w tym strumieniu jest zastępowany na karcie Modele i w obszarze roboczym strumienia. Usunięcie zaznaczenia tej opcji powoduje dodanie modelu do istniejącego modelu na karcie i w obszarze roboczym. To ustawienie jest przesłaniane przez ustawienie zastępowania modelu w łączu modelu.

Ta opcja jest dostępna dla ustawienia **Nowy wynik**:

Ostrzegaj o przekroczeniu liczby wyników [n] Określ, czy wyświetlać ostrzeżenie po przekroczeniu wstępnie określonej liczby elementów na karcie wyników. Ustawienie domyślne wynosi 20, jednak użytkownik może jest zmienić.

Następujące opcje są dostępne we wszystkich przypadkach:

Wybierz zakładkę Wybierz, która karta zostanie wyświetlona (Wyniki czy Modele), gdy odpowiedni obiekt zostanie utworzony podczas wykonywania strumienia.

- Wybierz opcję **Zawsze**, aby przełączyć kartę w panelu zarządzania.
- Wybierz opcję **Jeżeli wygenerowany przez bieżący strumień**, aby wybrać kartę odpowiadającą tylko obiektom utworzonym przez strumień aktualnie widoczny w obszarze roboczym.
- Wybierz opcję **Nigdy**, aby program nie powodował przechodzenia na inne karty w celu powiadomienia użytkownika o wygenerowanych wynikach i modelach.

Migaj zakładką Wybierz tę opcję, aby karta Wyniki lub Modele migała w panelu zarządzania po utworzeniu nowych wyników lub modeli.

- Wybierz opcję **Jeśli nie wybrany**, aby odpowiednia karta migała (jeśli nie jest jeszcze wybrana) po utworzeniu nowego obiektu w panelu zarządzania.
- Wybierz opcję **Nigdy**, aby program nie migał zakładką w celu powiadomienia użytkownika utworzonych wynikach lub modelach.

Przewiń paletę, aby pokazać element (tylko **Nowy model**). Określ, czy automatycznie przewijać kartę Modele (dostępną w panelu zarządzania) w celu przedstawienia najnowszych modeli.

- Wybierz opcję **Zawsze**, aby włączyć przewijanie.
- Wybierz opcję **Jeżeli wygenerowany przez bieżący strumień**, aby przewijać tylko obiekty utworzone przez strumień aktualnie widoczny w obszarze roboczym.
- Wybierz opcję **Nigdy**, aby w oprogramowaniu nie przewijać automatycznie karty Modele.

Otwórz okno wynikowe (tylko **Gdy pojawia się nowy wynik**). Wybierz tę opcję, aby określić, czy po utworzeniu okno wyników jest otwierane automatycznie.

- Wybierz opcję **Zawsze**, aby zawsze otwierać nowe okno wyników.
- Wybierz opcję **Jeżeli wygenerowany przez bieżący strumień**, aby otwierać nowe okno dla wyników utworzonych przez strumień aktualnie widoczny w obszarze roboczym.
- Wybierz opcję **Nigdy**, aby w oprogramowaniu nie otwierać automatycznie nowych okien wygenerowanych wyników.

Aby w przypadku tej karty przywrócić ustawienia domyślne, kliknij pozycję **Domyślne wartości**.

Określanie opcji wyświetlania

Za pomocą karty Wyświetl w oknie dialogowym Opcje użytkownika można skonfigurować opcje wyświetlania czcionek i kolorów w oprogramowaniu IBM SPSS Modeler.

Wyświetl okno powitalne podczas uruchamiania programu. Wybierz tę opcję, aby podczas uruchamiania oprogramowania wyświetlać okno powitalne. Opcje zawarte w powitalnym oknie dialogowym pozwalają na uruchomienie samouczka, otwarcie przykładowego strumienia, istniejącego strumienia, projektu bądź utworzenie nowego strumienia.

Pokaż komentarze strumienia i superwęzłów. Wybranie tej opcji powoduje, że znaczniki (o ile istnieją) w strumieniach i Superwęzłach są wyświetlane domyślnie. Znacznik zawiera komentarze dotyczące strumienia, łącza modelu i wyróżnione gałęzie oceniania.

Standardowe czcionki i kolory (dostępne po ponownym uruchomieniu). Opcje w tym polu sterowania służą do konfiguracji wyglądu ekranu programu IBM SPSS Modeler, schematu kolorów i rozmiaru wyświetlanych czcionek. Określone w tym położeniu opcje są stosowane dopiero po zamknięciu i ponownym uruchomieniu oprogramowania IBM SPSS Modeler.

- **Wygląd.** Wybierz standardowy schemat kolorów i wygląd ekranu. Dostępne opcje:
  - **SPSS Standard**, wygląd domyślny.
  - **SPSS Classic** - wygląd znany użytkownikom wcześniejszych wersji oprogramowania SPSS Modeler.
  - **Windows** - wygląd systemu Windows zapewniający wyższy kontrast w przypadku obszaru roboczego strumienia i palet.
  - **Analytics Carbon**, nowoczesny wygląd z minimalistycznymi ikonami i kolorystyką.
- **Domyślna wielkość czcionek dla węzłów.** Określ rozmiar czcionki używanej w paletach węzłów wyświetlanych w obszarze roboczym strumienia.
- **Określ stałą szerokość czcionki.** Zaznacz to pole wyboru, aby wybrać czcionkę o stałej szerokości oraz **Rozmiar** czcionki używanej podczas tworzenia skryptów i wyrażeń CLEM. Czcionką domyślną jest `Monospace plain`. Kliknij przycisk **Zmień...**, aby wyświetlić listę dostępnych czcionek.

Uwaga: Można określić rozmiar ikon węzłów strumienia. W tym celu należy wybrać panel Układ dostępny na karcie Opcje w oknie dialogowym właściwości strumienia. Z menu głównego wybierz pozycje **Narzędzia** > **Właściwości strumienia** > **Opcje** > **Układ**.

Niestandardowe kolory. W tej tabeli znajdują się aktualnie wybrane kolory używane w celu wyświetlania różnych elementów. W przypadku każdego elementu na liście bieżący kolor można zmienić, klikając dwukrotnie odpowiedni wiersz w kolumnie **Kolor** i wybierając na liście żądany kolor. Aby określić kolor niestandardowy, należy przewinąć listę do końca i kliknąć wpis **Kolor...**.

Porządek kolorów kategorii na wykresie. W tej tabeli znajdują się aktualnie wybrane kolory używane w celu wyświetlania nowych wykresów. Kolejność kolorów odzwierciedla kolejność ich używania na wykresie. Przykładowo: jeśli zmienna nominalna używana jako kolorowa nakładka zawiera cztery niepowtarzalne wartości, to używane są tylko pierwsze cztery kolory zawarte na liście. W przypadku każdego elementu na liście bieżący kolor można zmienić, klikając dwukrotnie odpowiedni wiersz w kolumnie **Kolor** i wybierając na liście żądany kolor. Aby określić kolor niestandardowy, należy przewinąć listę do końca i kliknąć wpis **Kolor...**. Zmiany wprowadzone w tym obszarze nie wpływają na wcześniej utworzone wykresy.

Aby w przypadku tej karty przywrócić ustawienia domyślne, kliknij pozycję **Domyślne wartości**.

Określanie opcji wyświetlania składni

Za pomocą karty Składnia w oknie dialogowym Opcje użytkownika można skonfigurować opcje atrybutów czcionek i kolorów wyświetlanych w skryptach tworzonych w oprogramowaniu IBM SPSS Modeler.

Wyróżnianie składni. W tej tabeli znajdują się aktualnie wybrane kolory używane w celu wyświetlania różnych elementów składni oraz czcionki i okna, w którym składnia jest wyświetlana. W przypadku każdego elementu na liście można zmienić kolor, klikając odpowiednie menu rozwijane w wierszu i wybierając na liście żądany kolor. Ponadto czcionki można pogrubić lub wybrać opcję kursywy.

Podgląd. W tej tabeli przedstawiono przykładową składnię korzystającą z kolorów i atrybutów tekstowych wybranych w tabeli **Wyróżnianie składni**. Podgląd jest aktualizowany po zmianie wybranego elementu.

Aby w przypadku tej karty przywrócić ustawienia domyślne, kliknij pozycję **Domyślne wartości**.

Określanie opcji eksportowania PMML

Na karcie PMML można sterować sposobem eksportowania modeli w oprogramowaniu IBM SPSS Modeler do formatu PMML (Predictive Model Markup Language). Więcej informacji można znaleźć w temacie “Importowanie i eksportowanie modeli w formacie PMML” na stronie 224.

Eksportuj PMML. W tym obszarze można skonfigurować odmianę formatu PMML najlepiej odpowiadającą aplikacji docelowej.

- Wybierz opcję **Z rozszerzeniami**, aby umożliwić stosowanie rozszerzeń PMML w wyjątkowych przypadkach, gdy nie istnieją standardowe ekwiwalenty PMML. Należy pamiętać, że w większości sytuacji wynik będzie zgodny ze standardowym PMML.
- Wybierz opcję **Jako standard PMML...**, aby wyeksportować PMML jak najbardziej zgodny ze standardowym PMML.

Opcje standardu PMML. Po określeniu opcji **Jako standard PMML...** można wybrać jeden z dwóch prawidłowych sposobów eksportowania modeli regresji logistycznej i liniowej:

- W postaci modeli **PMML < GeneralRegression>**
- W postaci modeli **PMML < Regression>**

Dodatkowe informacje na temat języka PMML zawarto w serwisie Data Mining Group pod adresem http://www.dmg.org.

Określanie informacji o użytkowniku

Informacje o użytkowniku. Informacje wprowadzone w tym obszarze można wyświetlić na karcie Adnotacje w przypadku węzłów i innych tworzonych obiektów.

Określanie trybu

Ustawienia trybu Modeler. Na karcie **Tryb** można wybierać spośród następujących trybów:

- **Tryb tradycyjny SPSS Modeler** — w tym trybie w interfejsie użytkownika wyświetlane są wszystkie dostępne węzły i wyrażenia.
- **Tryb Analytic Server** — w tym trybie wyświetlane są tylko węzły i wyrażenia obsługiwane przez Analytic Server. Należy jednak zauważyć, że niektóre węzły i wyrażenia CLEM nadal będą wyświetlane, mimo że nie są *w pełni* obsługiwane przezAnalytic Server. Poniższa tabela zawiera ogólne informacje o tym, które węzły są obsługiwane, częściowo obsługiwane i nieobsługiwane przez Analytic Server.

Więcej informacji o serwerze Analytic Server zawiera dokumentacja serwera Analytic Server.

Jeśli program SPSS Modeler jest skonfigurowany w taki sposób, by węzły bazy danych były widoczne na palecie Modelowanie w bazie, to zmiana trybu nie będzie miała na nie wpływu. Węzły bazy danych zawsze będą wyświetlane. W przypadku korzystania z integracji z programem IBM Db2 for z/OS lub IBM Netezza niekiedy węzły te mogą zniknąć z palety Modelowanie w bazie po przejściu do trybu Analytic Server. W takim przypadku należy wybrać kolejno opcje **Narzędzia** > **Opcje** > **Aplikacje pomocnicze** i ponownie ustawić pola wyboru.

Tabela 38. Obsługa węzłów

Typ węzła (nazwa palety)	Obsługiwany przez Analytic Server	Częściowo obsługiwany przez Analytic Server	Nieobsługiwany przez Analytic Server
Źródła	• Węzeł źródłowy Analytic Server		• Baza danych • Var. File • Plik kolumnowy • Plik Statistics • Gromadzenie danych • IBM Cognos • Import z TWC • Import z TM1 • Plik SAS • Excel • Dane niestandardowe XML • Symulacje Generowanie • Widok danych • Geoprzestrzenny • Pamięć obiektowa • Importowanie przez rozszerzenie • Import R • SNA (Analiza grupy i Analiza dyfuzji)

Tabela 38. Obsługa węzłów (kontynuacja)

Typ węzła (nazwa palety)	Obsługiwany przez Analytic Server	Częściowo obsługiwany przez Analytic Server	Nieobsługiwany przez Analytic Server
Rekordy	• Selekcja • Sortowanie • Zrównoważenie • Powtórzenia • Agregacja RFM • Dołączanie • Szeregi czasowe • Transformacja przez rozszerzenie • Strumień TCM	• Próba (obsługuje tylko opcję Losowo % dla metody prostej. Metoda złożona nie jest obsługiwana.) • Łączenie (obsługuje tylko łączenia tylko za pomocą kluczy i warunków) • Agregacja (Analytic Server nie obsługuje 1. kwantyla, 3. kwantyla i mediany)	• Siatka czasoprzestrzeni • Optymalizacja CPLEX
Zmienne	• Typ • Filtrowanie • Wyliczanie • Wypełnianie • Rekodowanie • Zespół • Flagowanie • Restrukturyzacja • Reorganizacja • Odwzorowanie • Przedziały czasowe	• Auto Przygotowanie (obsługuje tylko transformację) • Kategoryzacja (metoda kategoryzacji `equalFreq` nie jest obsługiwana, gdy wybrane jest ustawienie Wiązania **Pozostaw w bieżącej**) • Analiza RFM(metoda kategoryzacji `N-tyl` nie jest obsługiwana, gdy wybrane jest ustawienie Wiązania **Pozostaw w bieżącej**) • Podział (Analytic Server obsługuje ten węzeł tylko wtedy, gdy do powtarzalnego przydzielania wierszy do podzbiorów używana jest unikalna zmienna)	• Anonimizacja • Historia • Transpozycja

Tabela 38. Obsługa węzłów (kontynuacja)

Typ węzła (nazwa palety)	**Obsługiwany przez Analytic Server**	**Częściowo obsługiwany przez Analytic Server**	**Nieobsługiwany przez Analytic Server**
Wykresy	• Wykresy • Wykres wielokrotny • Wykres sekwencyjny • Rozkład • Histogram • Zbiór • Sieciowy • Ewaluacja • Wizualizacja na mapie • Wykres E-Plot (Beta) • t-SNE	• Wizualizacja (obsługuje tylko funkcję trybu agregacji dla zmiennych o poziomie pomiaru dyskretnym, nominalnym, liczby porządkowej lub flagi)	

Tabela 38. Obsługa węzłów (kontynuacja)

Typ węzła (nazwa palety)	Obsługiwany przez Analytic Server	Częściowo obsługiwany przez Analytic Server	Nieobsługiwany przez Analytic Server
Modelowanie	• Szereg czasowy • TCM • Izotoniczna-AS • Drzewa losowe • Drzewo-AS • Liniowy-AS • GLE • LSVM • STP • Dwustopniowa-AS • Reguły asocjacyjne • XGBoost-AS • K-średnie-AS	• Auto Klasyfikacja • Auto Predykcja (dwa węzły obsługują tylko rozdział, a zmienna z rolą rozdziału należy dostarczyć podczas używania opcji Auto Klasyfikacja **Uruchom na Analytic Server (rozdziały włączone)**) • Rozszerzenie (polecenia R tworzące model nie są obsługiwane przez Analytic Server) • Następujące węzły obsługują tylko podział i PSM: – Drzewo C&R – Liniowy – Sieć neuronowa – CHAID – Quest	• Auto Grupowanie • Lista decyzyjna • C5.0 • Regresja • Redukcja wymiarów • Dobór predyktorów • Analiza dyskryminacyjna • Regresja logistyczna • Modele uogólnione • GLMM • Sieć Bayesa • Apriori • Carma • Sekwencje • K-średnie • Sieć Kohonena • Dwustopniowa • Anomalie • KNN • R • Las losowy • Następujące węzły uwzględniają ASL, ale tylko w operacjach odczytu/zapisu: – Coxa – SVM – SLRM
Wynik	• Tabela • Macierz • Analiza • Audyt danych • Transformacja • Statystyki • Średnie • Raport • Globalne	• Wyniki przez rozszerzenie • Polecenia R tworzące wynik	• Symulacje Wynik • Symulacje Dopasowanie • Wynik R

Tabela 38. Obsługa węzłów (kontynuacja)

Typ węzła (nazwa palety)	Obsługiwany przez Analytic Server	Częściowo obsługiwany przez Analytic Server	Nieobsługiwany przez Analytic Server
Eksport	• Węzeł eksportu Analytic Server	• Eksportowanie przez rozszerzenie • Polecenia R eksportu	• Baza danych • Plik płaski • Plik Statistics • Gromadzenie danych • Excel • Eksport IBM Cognos • Eksport do TM1 • SAS • XML • Pamięć obiektowa • Eksport R
IBM SPSS Statistics			• Plik Statistics • Przekształcenia Statistics • Model Statistics • Wynik Statistics • Plik Statistics
IBM SPSS Text Analytics	• Analiza powiązań w tekście • Eksploracja tekstu • Węzeł Język		• Lista plików • Kanał informacyjny WWW • Analiza powiązań w tekście • Tłumacz • Eksploracja tekstu • Przeglądarka plików
Python			• SMOTE • SVM z jedną klasą • Drzewo XGBoost • Liniowy XGBoost • t-SNE • Las losowy • HDBSCAN
Spark	Wszystkie obsługiwane		

Dostosowywanie palety węzłów

Strumienie tworzy się, korzystając z węzłów. Paleta węzłów w dolnej części okna IBM SPSS Modeler zawiera wszystkie węzły, których można użyć podczas tworzenia strumienia. Więcej informacji można znaleźć w temacie “Paleta węzłów” na stronie 15.

Paletę węzłów można zreorganizować na dwa sposoby:

- Dostosowując ustawienia opcji Zarządzanie paletami. Więcej informacji można znaleźć w temacie “Dostosowywanie ustawień opcji Zarządzanie paletami” na stronie 245.
- Zmieniając sposób wyświetlania kart palety zawierających podpalety w palecie węzłów. Więcej informacji można znaleźć w temacie “Tworzenie podpalety” na stronie 246.

Dostosowywanie ustawień opcji Zarządzanie paletami

Opcję Zarządzanie paletami można dostosować odpowiednio do sposobów korzystania z programu IBM SPSS Modeler. Na przykład, jeśli często analizujemy dane szeregów czasowych z bazy danych, możemy umieścić węzeł źródłowy bazy danych, węzeł przedziałów czasowych, węzeł szeregów czasowych oraz węzeł wykresu sekwencyjnego na osobnej karcie palety. Panel Zarządzanie paletami ułatwia wprowadzanie powyższych modyfikacji dzięki możliwości tworzenia niestandardowych kart palet na Palecie węzłów.

Zarządzanie paletami umożliwia wykonywanie różnych zadań:

- Sterowanie widocznością kart palety na karcie Paleta węzłów poniżej obszaru roboczego strumienia.
- Zmianę kolejności, w jakiej karty palety są wyświetlane na Palecie węzłów.
- Tworzenie i edycję własnych kart palet oraz dowolnych powiązanych z nimi podpalet.
- Edycję domyślnych wyborów węzłów na kartach.

W celu uzyskania dostępu do karty Zarządzanie paletami w menu Narzędzia kliknij opcję **Zarządzaj paletami**.

Nazwa palety. Na liście wymieniona jest każda dostępna karta palety, niezależnie od tego, czy jest ona wyświetlana na Palecie węzłów, czy nie. Obejmuje to wszelkie utworzone przez użytkownika karty palet. Więcej informacji można znaleźć w temacie “Tworzenie karty Paleta” na stronie 245.

Liczba węzłów. Liczba węzłów wyświetlanych na każdej karcie palety. Wyższa liczba tutaj oznacza, że znacznie wygodniejsze może okazać się utworzenie podpalet, a następnie podzielenie węzłów na karcie. Więcej informacji można znaleźć w temacie “Tworzenie podpalety” na stronie 246.

Pokaż Zaznacz to pole, aby wyświetlać kartę palety na Palecie węzłów. Więcej informacji można znaleźć w temacie “Wyświetlanie kart palety na Palecie węzłów” na stronie 246.

Podpalety. Aby wybrać podpalety do wyświetlenia na karcie palety, podświetl właściwy wpis **Nazwa palety** i kliknij ten przycisk; zostanie wyświetlone okno dialogowe Podpalety. Więcej informacji można znaleźć w temacie “Tworzenie podpalety” na stronie 246.

Przywróć domyślne. Kliknij ten przycisk, aby całkowicie usunąć wszystkie zmiany i uzupełnienia wprowadzone na paletę, a także podpalety, i powrócić do domyślnych ustawień palety.

Tworzenie karty Paleta

Aby utworzyć niestandardową kartę palety:

1. Korzystając z menu Narzędzia, otwórz okno Zarządzanie paletami.
2. Po prawej stronie kolumny *Pokaż* kliknij przycisk Dodaj paletę; zostanie wyświetlone okno dialogowe Utwórz/Zmień paletę.
3. Wpisz unikalną nazwę w polu **Nazwa palety**.
4. W obszarze **Węzły dostępne** wybierz węzeł, który ma zostać dodany na karcie palety.
5. Kliknij przycisk ze strzałką w prawo (Dodaj węzeł), aby przenieść wyróżniony węzeł do obszaru **Wybrane węzły**. Powtarzaj tę czynność aż do chwili dodania wszystkich węzłów.

   Po dodaniu wszystkich wymaganych węzłów można zmienić kolejność, w której są one wyświetlane na karcie palety:
6. Użyj przycisków ze strzałką, aby przenieść węzeł o jeden wiersz w górę lub w dół.
7. Użyj przycisków z linią i strzałką, aby przenieść węzeł na górę lub na dół listy.

8. Aby usunąć węzeł z palety, podświetl węzeł, a następnie kliknij przycisk Usuń po prawej stronie obszaru **Wybrane węzły**.

Wyświetlanie kart palety na Palecie węzłów

W programie IBM SPSS Modeler mogą być dostępne opcje, z których użytkownik nigdy nie korzysta; w takim przypadku można za pomocą funkcji Zarządzanie paletami ukryć karty zawierające te węzły.

Aby wybrać karty, które mają być wyświetlane na Palecie węzłów:

1. Korzystając z menu Narzędzia, otwórz okno Zarządzanie paletami.
2. Korzystając z pól wyboru w kolumnie *Pokaż*, zdecyduj osobno dla każdej z kart palety, czy ma być ona wyświetlana, czy ukrywana.

Aby trwale usunąć kartę palety z palety węzłów, podświetl węzeł i kliknij przycisk Usuń po prawej stronie kolumny *Pokaż* . Karty palety po jej usunięciu nie można już odzyskać.

Uwaga: Nie można usunąć kart palety domyślnej dołączonej do programu IBM SPSS Modeler, z wyjątkiem karty Ulubione.

Zmiana kolejności wyświetlania na Palecie węzłów

Po wybraniu kart palety, które mają być wyświetlane, można zmienić kolejność, w której są one wyświetlane na Palecie węzłów:

1. Użyj przycisków z samą strzałką, aby przenieść węzeł o jeden wiersz w górę lub w dół. Przeniesienie w górę oznacza przeniesienie ich na lewą stronę Palety węzłów, i odwrotnie.
2. Użyj przycisków z linią i strzałką, aby przenieść kartę palety na górę lub na dół listy. Pozycje u góry listy zostaną wyświetlone po lewej stronie Palety węzłów.

Wyświetlanie podpalet na karcie Paleta

W ten sam sposób, w jaki można zdecydować o tym, które karty palety są wyświetlane na Palecie węzłów, można także zarządzać dostępnością podpalet na karcie palety nadrzędnej.

Aby zaznaczyć podpalety do wyświetlania na karcie palety:

1. Korzystając z menu Narzędzia, otwórz okno Zarządzanie paletami.
2. Wybierz wymaganą paletę.
3. Kliknij przycisk Podpalety; wyświetlane jest okno dialogowe Podpalety.
4. Korzystając z pól wyboru w kolumnie *Pokaż*, zdecyduj, czy uwzględniać każdą z podpalet na karcie palety. Podpaleta **Wszystkie** jest zawsze wyświetlana i nie można jej usunąć.
5. Aby trwale usunąć podpaletę z karty palet, podświetl podpaletę i kliknij przycisk Usuń po prawej stronie kolumny *Pokaż* .

Uwaga: Nie można usuwać podpalet domyślnych dołączonych do karty palety Modelowanie.

Zmiana kolejności wyświetlania na karcie Paleta

Po wybraniu podpalet, które mają być wyświetlane, można zmienić kolejność, w której są one wyświetlane na karcie palety nadrzędnej:

1. Użyj przycisków ze strzałką, aby przenieść podpaletę o jeden wiersz w górę lub w dół.
2. Użyj przycisków z linią i strzałką, aby przenieść podpaletę na górę lub na dół listy.

Tworzone podpalety są wyświetlane na Palecie węzłów po wybraniu karty ich palety nadrzędnej. Więcej informacji można znaleźć w temacie “Zmiana widoku karty palety” na stronie 247.

Tworzenie podpalety

Z uwagi na fakt, że istniejące węzły można dodać do kart palet niestandardowych, możliwe, że będzie ich więcej, niż można wyświetlić na ekranie bez przewijania. Aby uniknąć konieczności przewijania, można utworzyć podpalety, w których mają zostać umieszczone węzły wybrane na karcie palety. Na przykład

w przypadku utworzenia karty palety zawierającej węzły najczęściej używane do tworzenia strumieni można utworzyć cztery podpalety umożliwiające podział na węzły źródłowe, operacje na zmiennych, modelowanie i dane wynikowe.

Uwaga: Węzły podpalety można wybrać wyłącznie spośród tych dodanych na karcie palety nadrzędnej.

Aby utworzyć podpaletę:

1. Korzystając z menu Narzędzia, otwórz okno Zarządzanie paletami.
2. Wybierz paletę, na którą chcesz dodać podpalety.
3. Kliknij przycisk Podpalety; wyświetlane jest okno dialogowe Podpalety.
4. Po prawej stronie kolumny *Pokaż* kliknij przycisk Dodaj podpaletę; zostanie wyświetlone okno dialogowe Utwórz/Zmień podpaletę.
5. Wpisz unikalną nazwę w polu **Podpaleta**.
6. W obszarze **Węzły dostępne** wybierz węzeł, który ma zostać dodany do podpalety.
7. Kliknij przycisk strzałki w prawo Dodaj węzeł, aby przenieść wyróżniony węzeł do obszaru **Wybrane węzły**.
8. Po dodaniu potrzebnych węzłów kliknij przycisk **OK**, aby powrócić do okna dialogowego Podpalety.

Tworzone podpalety są wyświetlane na Palecie węzłów po wybraniu karty ich palety nadrzędnej. Więcej informacji można znaleźć w temacie “Zmiana widoku karty palety” na stronie 247.

Zmiana widoku karty palety

Z uwagi na dostępność w programie IBM SPSS Modeler dużej liczby węzłów mogą być one niewidoczne na mniejszych ekranach bez przewijania na lewą lub prawą stronę palety węzłów; jest to szczególnie widoczne na karcie palety Modelowanie. W celu zredukowania potrzeby przewijania można wybrać wyświetlanie tylko węzłów zawartych w podpalecie (tam gdzie jest ona dostępna). Więcej informacji można znaleźć w temacie “Tworzenie podpalety” na stronie 246.

W celu zmiany węzłów wyświetlanych na karcie palety wybierz kartę palety, a następnie, z menu po lewej stronie wybierz opcję wyświetlania wszystkich węzłów lub tylko węzłów z konkretnej podpalety.

Rozdział 16. Czynniki wpływające na wydajność dotyczące strumieni i węzłów

Istnieje możliwość zaprojektowania strumieni w sposób umożliwiający maksymalizację wydajności dzięki ustawieniu węzłów w najbardziej wydajnej konfiguracji, dzięki włączeniu, tam gdzie jest taka możliwość, pamięci podręcznych węzłów, a także dzięki uwzględnieniu innych czynników wymienionych w niniejszej sekcji.

Obok omawianych tu środków możliwe jest wprowadzenie modyfikacji w znacznie większym stopniu zwiększających wydajność, polegających na efektywnym wykorzystaniu baz danych, w szczególności dzięki optymalizacji kodu SQL.

Kolejność węzłów

Nawet w przypadku rezygnacji z optymalizacji kodu SQL, już sama kolejność węzłów w strumieniu może mieć wpływ na wydajność. Ogólny cel to minimalizacja przetwarzania w dół strumienia; stąd, w przypadku dysponowania węzłami zmniejszającymi ilość danych należy je umieścić w pobliżu początku strumienia. Program IBM SPSS Modeler Server pozwala zastosować niektóre reguły zmiany kolejności automatycznie, podczas kompilacji, umieszczając niektóre węzły na przedzie, o ile tylko istnieje gwarancja, że nie wprowadzi to zakłóceń. (Ta funkcja jest domyślnie włączona. Należy upewnić się u administratora systemu, że jest ona włączona w instalacji użytkownika.)

W przypadku korzystania z optymalizacji kodu SQL może okazać się niezbędna maksymalizacja jego dostępności i skuteczności. Ponieważ optymalizacja jest wstrzymywana, jeśli strumień zawiera operację, której nie można przeprowadzić w bazie danych, najlepiej jest grupować operacje ze zoptymalizowanym kodem SQL na początku strumienia. Strategia ta pozwala zachować przetwarzanie w większym zakresie w bazie danych — w efekcie do programu IBM SPSS Modeler wprowadzana jest mniejsza ilość danych.

W przypadku większości baz danych możliwe jest przeprowadzenie wymienionych poniżej operacji. Należy w miarę możliwości grupować je na *początku* strumienia:

- Łączenie wg klucza (złączenie)
- Selekcja
- Agregacja
- Sortowanie
- Próbkowanie
- Dołączanie
- Powtórzenia w trybie *Uwzględnij*, w którym wybrano wszystkie zmienne
- Wypełnianie
- Podstawowe operacje wyliczania (arytmetyczne lub polegające na manipulacji na łańcuchach — w zależności od tego, które z tych operacji są obsługiwane przez bazę danych)
- Ustawienie do oznaczenia flagą

W przypadku większości baz danych przeprowadzenie wymienionych poniżej operacji nie jest możliwe. Powinny one zostać umieszczone w strumieniu *za* operacjami z poprzedniej listy:

- Operacje na dowolnych danych niebazodanowych, na przykład pliki płaskie
- Łączenie wg kolejności
- Zrównoważenie
- Powtórzenia w trybie *Odrzuć* lub operacje, w których tylko podzbiór zmiennych oznaczono jako powtórzenia

- Wszelkie operacje wymagające uzyskania dostępu do danych z rekordów innych niż ten obecnie przetwarzany.
- Wyliczenia dla zmiennych Stan i Liczebność
- Operacje na węźle Historia
- Operacje obejmujące funkcje "@" (szereg czasowy)
- Tryby sprawdzania typu: *Ostrzegaj* i *Przerwij*
- Tworzenie modelu, zastosowanie i analizy

  Uwaga: Drzewa decyzyjne, zestawy reguł, regresja liniowa i modele generowane na podstawie czynnika umożliwiają generowanie kodu SQL, co pozwala na wstawienie ich z powrotem do bazy danych.
- Generowanie wyników w dowolne miejsce inne niż baza danych przetwarzająca dane

Pamięci podręczne węzłów

Aby zoptymalizować wykonywanie strumienia, można ustawić *pamięć podręczną* w dowolnym węźle niebędącym węzłem końcowym. Po skonfigurowaniu pamięci podręcznej dla węzła pamięć podręczna jest wypełniana danymi przechodzącymi przez węzeł przy następnym uruchomieniu strumienia danych. Od tego momentu dane są odczytywane z pamięci podręcznej (która jest przechowywana w katalogu tymczasowym na dysku), nie zaś ze źródła danych.

Buforowanie jest szczególnie przydatne w przypadku przeprowadzenia wcześniej czasochłonnej operacji takiej jak sortowanie, scalanie czy agregacja. Załóżmy na przykład, że mamy węzeł źródłowy z ustawieniem odczytu danych sprzedażowych z bazy danych oraz węzeł Agregacja podsumowujący sprzedaż wg lokalizacji. Istnieje możliwość skonfigurowania pamięci podręcznej w węźle Agregacja zamiast w węźle źródłowym, tak aby w pamięci podręcznej były przechowywane dane zagregowane, nie zaś cały zbiór danych.

Uwaga: Buforowanie w węzłach źródłowych, którego istotą jest po prostu przechowywanie kopii oryginalnych danych wczytywanych do IBM SPSS Modeler, w większości przypadków nie wpływa dodatnio na wydajność.

Węzły z włączonym buforowaniem są wyświetlane z niewielką ikoną dokumentu w prawym górnym rogu. W przypadku zbuforowania danych w węźle ikona dokumentu ma kolor zielony.

Aby włączyć pamięć podręczną

1. W obszarze roboczym strumienia kliknij węzeł prawym przyciskiem myszy i w menu wybierz pozycję **Pamięć podręczna**.
2. W podmenu buforowania kliknij opcję **Włącz**.
3. Pamięć podręczną można wyłączyć, klikając węzeł prawym przyciskiem myszy i klikając opcję **Wyłącz** w podmenu buforowania.

Buforowanie węzłów w bazie danych

W przypadku strumieni uruchamianych w bazie danych dane mogą być buforowane na bieżąco w tabeli tymczasowej w bazie danych zamiast w systemie plików. W przypadku połączenia z optymalizacją SQL może to skutkować znaczącymi korzyściami, jeśli chodzi o wydajność. Na przykład dane wynikowe ze strumienia scalającego wiele tabel z myślą o stworzeniu widoku eksploracji bazy danych mogą zostać zbuforowane i wykorzystane ponownie w razie potrzeby. Wydajność można dodatkowo podnieść, automatycznie generując kod SQL dla wszystkich kolejnych węzłów.

Aby korzystać z buforowania w bazie danych, należy włączyć zarówno optymalizację SQL, jak i buforowanie w bazie danych. Należy pamiętać, że ustawienia optymalizacji serwera zastępują odpowiadające im ustawienia klienta. Więcej informacji można znaleźć w temacie “Ustawianie opcji optymalizacji strumieni” na stronie 44.

Jeśli włączono buforowanie w bazie danych, należy po prostu kliknąć prawym przyciskiem myszy dowolny węzeł niebędący węzłem końcowym celu zbuforowania danych w tym punkcie. Spowoduje to automatyczne utworzenie pamięci podręcznej bezpośrednio w bazie danych przy następnym

uruchomieniu strumienia. Jeśli nie włączono buforowania bazy danych lub optymalizacji SQL, wówczas pamięć podręczna zostanie zapisana w systemie plików.

Uwaga: Następujące bazy danych obsługują tymczasowe tabele do celów buforowania: Db2, Oracle, SQL Server i Teradata. W przypadku innych baz danych, takich jak Netezza, na potrzeby buforowania w bazie danych będzie używana normalna tabela. Kod SQL można dostosować do konkretnych baz danych — w celu uzyskania pomocy należy skontaktować się z działem usług.

Wydajność: węzły procesowe

Sortowanie. Węzeł Sortowanie musi odczytywać cały zbiór danych, zanim będzie możliwe ich posortowanie. Dane są przechowywane w pamięci do pewnego limitu, a po jego przekroczeniu są zapisywane na dysku. Algorytm sortowania jest algorytmem łączonym; dane są wczytywane do pamięci do określonego limitu i sortowane za pośrednictwem szybkiego hybrydowego algorytmu szybkiego sortowania. Po umieszczeniu wszystkich danych w pamięci sortowanie jest ukończone. W przeciwnym wypadku stosowany jest łączony algorytm sortowania. Posortowane dane są zapisywane w pliku, a następnie kolejna porcja danych jest wczytywana do pamięci, sortowana i zapisywana na dysku. Jest to powtarzane aż do chwili wczytania wszystkich danych; następnie posortowane porcje są łączone. Łączenie może wymagać cyklicznych przejść przez dane przechowywane na dysku. W okresie szczytowego użycia węzeł Sortowanie będzie dysponował dwoma kompletnymi egzemplarzami zbioru danych na dysku: posortowanym i nieposortowanym.

Łączny czas działania algorytmu jest rzędu `N*log(N)`, gdzie *N* jest liczbą rekordów. Sortowanie w pamięci przebiega szybciej niż łączenie z dysku, tak że rzeczywisty czas uruchomienia może zostać zredukowany przez alokację większej ilości miejsca dla operacji sortowania. Algorytm alokuje sobie ułamek pamięci fizycznej RAM zgodnie z opcją konfiguracji programu IBM SPSS Modeler Server *Mnożnik użycia pamięci*. W celu zwiększenia ilości miejsca dostępnej na potrzeby sortowania, należy zwiększyć ilość fizycznej pamięci RAM lub zwiększyć tę wartość. Należy zwrócić uwagę, że jeśli proporcja wykorzystania pamięci przekracza zestaw roboczy procesu tak, że tylko ta część pamięci jest stronicowana na dysku, wydajność obniża się, ponieważ wzorzec dostępu do pamięci w algorytmie sortowania w pamięci jest losowy i może powodować nadmiarowe stronicowanie. Algorytm sortowania jest używany przez kilka innych węzłów poza węzłem Sortowanie, lecz mają wówczas zastosowanie te same reguły wydajności.

Kategoryzacja. Węzeł Kategoryzacja pozwala na odczytywanie całego zbioru danych wejściowych z myślą o obliczaniu granic kategorii, zanim przeprowadzi alokację rekordów do kategorii. Zbiór danych jest zapisywany w pamięci podręcznej, gdy obliczane są granice; następnie jest ponownie skanowany z myślą o alokacji. Jeśli metoda kategoryzacji to *ustalona szerokość przedziałów* lub *średnia+odchylenie standardowe*, zbiór danych jest zapisywany w pamięci podręcznej bezpośrednio na dysku. Metody te oferują liniowy czas uruchamiania i wymagają wystarczającej ilości miejsca na dysku do zapisania całego zbioru danych. Jeśli metoda kategoryzacji to *rangi* lub *N-tyle*, zbiór danych jest sortowany z użyciem opisanego wcześniej algorytmu sortowania, zaś posortowany zbiór danych jest używany jako pamięć podręczna. Sortowanie powoduje, że czas wykonywania dla tych metod wynosi `M*N*log(N)`, gdzie *M* jest liczbą pól poddanych kategoryzacji, zaś *N* jest liczbą rekordów; wymaga to dwukrotnie więcej miejsca na dysku, niż zajmuje zbiór danych.

Generowanie węzła Wyliczanie w oparciu o wygenerowane kategorie poprawi wydajność w kolejnych przejściach. Operacje wyliczania przebiegają znacznie szybciej niż kategoryzacja.

Łączenie wg klucza (złączenie). Węzeł Łączenie, gdzie metoda łączenia to *klucze* (równoważna złączaniu bazy danych), sortuje każdy z wejściowych zestawów danych wg zmiennych kluczowych. W przypadku tej części procedury czas wykonywania to `M*N*log(N)`, gdzie *M* jest liczbą danych wejściowych, zaś *N* jest liczbą rekordów najliczniejszych danych wejściowych; wymaga ona miejsca na dysku wystarczającego do przechowywania wszystkich zbiorów danych wejściowych, a także drugiej kopii najliczniejszych danych wejściowych. Czas wykonywania samego łączenia jest proporcjonalny do rozmiaru wynikowego zbioru danych, przy czym zależy to od częstotliwości wzajemnych dopasowań kluczy. W najgorszym przypadku, tam gdzie dane wynikowe stanowią iloczyn kartezjański danych wejściowych, czas uruchamiania może sięgnąć `NM`. Jest to rzadkość — większość złączeń ma o wiele mniej pasujących do siebie kluczy. Jeśli tylko jeden zbiór danych jest relatywnie większy niż inne, lub jeśli dane wejściowe są już posortowane według zmiennej kluczowej, wówczas możliwe jest podniesienie wydajności tego węzła za pomocą karty Optymalizacja.

Agregacja. Jeśli nie ustawiono opcji *Zmienne grupujące są posortowane*, ten węzeł wczytuje (lecz nie zapisuje) cały zbiór danych wejściowych, zanim wygeneruje jakiekolwiek zagregowane dane wyjściowe. W sytuacjach bardziej skrajnych, jeśli wielkość zagregowanych danych osiągnie limit (określony przez opcję konfiguracji *Mnożnik użycia pamięci* programu IBM SPSS Modeler Server, wówczas pozostała część zbioru danych jest sortowana i przetwarzana tak, jak gdyby ustawiono opcję *Zmienne grupujące są posortowane*. Po ustawieniu tej opcji żadne dane nie są zapisywane, ponieważ zagregowane rekordy wynikowe są generowane w miarę wczytywania danych wejściowych.

Powtórzenia. Węzeł Powtórzenia przechowuje wszystkie unikalne zmienne kluczowe w wejściowym zbiorze danych; w przypadkach, gdzie wszystkie zmienne są zmiennymi kluczowymi, zaś wszystkie rekordy są unikalne, przechowuje on cały zbiór danych. Domyślnie węzeł Powtórzenia sortuje dane wg zmiennych kluczowych, a następnie wybiera (lub odrzuca) pierwszy rekord odmienny z każdej grupy. W przypadku mniejszych zbiorów danych o mniejszej liczbie odmiennych kluczy, lub w przypadku tych, które zostały wstępnie posortowane, można wybrać opcje pozwalające poprawić prędkość i efektywność przetwarzania.

Typ. W niektórych przypadkach węzeł Typ zapisuje w pamięci podręcznej dane wejściowe podczas odczytu wartości; pamięć podręczna jest następnie używana do przetwarzania w dół strumienia. Pamięć podręczna wymaga wystarczającej ilości miejsca na dysku w celu przechowania całego zbioru danych, lecz jednocześnie podnosi prędkość przetwarzania.

Ewaluacja. Węzeł Ewaluacja musi sortować dane wejściowe, aby możliwe było obliczenie N-tyli. Sortowanie jest powtarzane dla każdego ocenianego modelu, ponieważ oceny i w efekcie kolejność rekordów różni się w zależności od sytuacji. Czas wykonywania wynosi `M*N*log(N)`, gdzie *M* jest liczbą modeli, zaś *N* jest liczbą rekordów.

Wydajność: węzły modelowania

Sieci neuronowa i Kohonena. Algorytmy uczące sieci neuronowej (w tym algorytm Kohonena) wykonują wiele przebiegów przez dane uczące. Dane są przechowywane w pamięci do pewnego limitu, a po jego przekroczeniu są zapisywane na dysku. Uzyskiwanie dostępu do danych uczących w istotnym stopniu pochłania zasoby, ponieważ metoda dostępu jest losowa, co może doprowadzić do nadmiernej aktywności dysku. Istnieje możliwość wyłączenia wykorzystania pamięci na dysku na potrzeby tych algorytmów; w tym celu należy wybrać opcję **Optymalizuj ze względu na szybkość** na karcie Model w oknie dialogowym węzła. Należy zwrócić uwagę, że jeśli ilość pamięci wymaganej do przechowywania danych jest większa niż uruchomiony zbiór procesów serwera, jego część będzie stronicowana na dysku, co może negatywnie odbić się na wydajności.

Po włączeniu opcji **Optymalizuj ze względu na pamięć** wartość procentowa fizycznej pamięci RAM jest alokowana dla algorytmu zgodnie z wartością opcji konfiguracji *Modelowanie wartości procentowej limitu pamięci* programu IBM SPSS Modeler Server. W celu użycia większej ilości miejsca na potrzeby uczenia sieci neuronowych należy albo udostępnić większą ilość pamięci RAM, albo zwiększyć wartość tej opcji. Należy jednak zwrócić uwagę, że ustawienie zbyt wysokiej wartości tej opcji spowoduje stronicowanie.

Czas wykonywania algorytmów sieci neuronowej zależy od wymaganego poziomu dokładności. Czas wykonywania można kontrolować, ustawiając warunek zatrzymania w oknie dialogowym węzła.

K-średnie. Algorytm grupowania K-średnie oferuje te same opcje wykorzystania pamięci, co algorytmy sieci neuronowej. Wydajność pracy z danymi przechowywanymi na dysku jest jednak lepsza, ponieważ dostęp do danych jest sekwencyjny.

Wydajność: wyrażenia CLEM

Funkcje kolejności CLEM („funkcje @”) wyszukujące wstecz w strumieniu danych muszą mieć możliwość zapisania wystarczającej ilości danych dla najdłuższego z wyszukiwań wstecz. W przypadku operacji, których stopień wyszukiwań wstecz jest nieograniczony, muszą zostać zapisane wszystkie wartości zmiennej. Operacja nieograniczona to taka, gdzie wartość przesunięcia nie jest dosłowną liczbą całkowitą; na przykład `@OFFSET(Sales, Month)`. Wartość przesunięcia jest zapisywana w zmiennej o nazwie *Month*, której wartość jest nieznana aż do chwili wykonania. Serwer musi zapisać wszystkie wartości zmiennej *Sales*, aby zagwarantować dokładność wyników. Jeśli górna granica jest znana, należy podać ją

jako dodatkowy argument; na przykład `@OFFSET(Sales, Month, 12)`. Operacja ta instruuje serwer o konieczności przechowania nie więcej niż 12 najnowszych wartości *Sales*. Funkcje sekwencji, ograniczone lub nie, niemal zawsze uniemożliwiają wygenerowanie kodu SQL.

Rozdział 17. Ułatwienia dostępu w programie IBM SPSS Modeler

Przegląd ułatwień dostępu w programie IBM SPSS Modeler

Program IBM SPSS Modeler zapewnia ułatwienia dostępu dla wszystkich użytkowników, a także specjalistyczną obsługę dla osób z upośledzeniem słuchu i wzroku. W tej sekcji opisano funkcje i sposoby pracy z zastosowaniem ułatwień dostępu, takich jak lektory ekranowe i skróty klawiaturowe.

Rodzaje funkcji ułatwiających dostęp

Osoby niedowidzące lub wykorzystujące do obsługi komputera tylko klawiaturę mogą używać szerokiego zestawu narzędzi usprawniających eksplorację danych na wiele różnych sposobów. Można na przykład tworzyć strumienie, określać opcje i odczytywać raporty wynikowe bez używania myszy. Dostępne skróty klawiaturowe przedstawiono w poniższym temacie. Ponadto program IBM SPSS Modeler współpracuje z lektorami ekranowymi, takimi jak JAWS for Windows. Użytkownicy mogą także zmienić schemat kolorów i w ten sposób zwiększać kontrast wyświetlanego obrazu. Obsługę wszystkich usprawnień przedstawiono w poniższych tematach.

Ułatwienia dostępu dla użytkowników słabo widzących

Istnieje wiele właściwości ułatwiających korzystanie z oprogramowania IBM SPSS Modeler.

Opcje wyświetlania

Użytkownik może zdefiniować kolor wyświetlanych wykresów. W oprogramowaniu można także używać określonych ustawień zdefiniowanych w systemie Windows. W ten sposób użytkownik może jeszcze dokładniej dostosować kontrast.

1. Aby ustawić opcje wyświetlania, w menu Narzędzia wybierz pozycję **Opcje użytkownika**.
2. Kliknij kartę **Wyświetl**. Opcje na tej karcie obejmują schemat kolorów oprogramowania, kolory wykresów i rozmiary czcionek węzłów.

Uwaga: Lektor ekranowy nie jest w stanie odczytywać wykresów, dlatego nie są one dostępne dla użytkowników słabowidzących.

Powiadomienia dźwiękowe

Włączając lub wyłączając dźwięki, można sterować sposobem powiadamiania o wykonaniu określonych operacji w oprogramowaniu. Przykładowo: powiadomienia dźwiękowe można przypisać do takich operacji, jak tworzenie i usuwanie węzłów lub tworzenie nowych wyników lub modeli.

1. Aby ustawić opcje powiadomień, w menu Narzędzia wybierz pozycję **Opcje użytkownika**.
2. Kliknij kartę **Powiadomienia**.

Sterowanie automatycznym uruchamianiem nowych okien

Za pomocą karty Powiadomienia dostępnej w oknie dialogowym Opcje użytkownika można określić, że nowe raporty wynikowe (np. tabele lub wykresy) są uruchamiane w oddzielnych oknach. Może się okazać, że dla użytkownika wygodniejsze jest wyłączenie tej opcji i otwieranie okna wyników tylko wtedy, gdy jest to wymagane.

1. Aby ustawić te opcje, w menu Narzędzia wybierz pozycję **Opcje użytkownika**.
2. Kliknij kartę **Powiadomienia**.

3. W oknie dialogowym wybierz pozycję **Gdy pojawia się nowy wynik** dostępną na liście w grupie **Powiadomienia wizualne**.
4. W obszarze **Otwórz okno wynikowe** wybierz ustawienie **Nigdy**.

Rozmiar węzła

Węzły można wyświetlać w rozmiarze standardowym lub pomniejszonym. Rozmiary można dostosować do własnych preferencji.

1. Aby ustawić opcje rozmiaru węzłów, w menu Plik wybierz pozycję **Właściwości strumienia**.
2. Kliknij kartę **Układ**.
3. Na liście **Rozmiar ikony** wybierz pozycję **Standardowy**.

Ułatwienia dostępu dla użytkowników niewidomych

Wsparcie dla użytkowników niewidomych polega głównie na użyciu lektora ekranowego, takiego jak np. JAWS w systemie Windows. Współpracę lektora ekranowego z programem IBM SPSS Modeler można zoptymalizować za pomocą wielu ustawień.

Opcje wyświetlania

Lektory ekranowe działają zazwyczaj lepiej, gdy kontrast ekranu jest podwyższony. Jeśli w systemie Windows już wybrano opcje wyższego kontrastu, te ustawienia można zastosować w lektorze.

1. Aby ustawić opcje wyświetlania, w menu Narzędzia wybierz pozycję **Opcje użytkownika**.
2. Kliknij kartę **Wyświetl**.

Uwaga: Lektor ekranowy nie jest w stanie odczytywać wykresów, dlatego nie są one dostępne dla użytkowników niewidomych.

Powiadomienia dźwiękowe

Włączając lub wyłączając dźwięki, można decydować o formie powiadomień dotyczących określonych operacji wykonywanych w oprogramowaniu. Przykładowo: powiadomienia dźwiękowe można przypisać do takich operacji, jak tworzenie i usuwanie węzłów lub tworzenie nowych wyników lub modeli.

1. Aby ustawić opcje powiadomień, w menu Narzędzia wybierz pozycję **Opcje użytkownika**.
2. Kliknij kartę **Powiadomienia**.

Sterowanie automatycznym uruchamianiem nowych okien

Za pomocą karty Powiadomienia dostępnej w oknie dialogowym Opcje użytkownika można określić, że nowe raporty wynikowe są uruchamiane w oddzielnych oknach. Może się okazać, że dla użytkownika wygodniejsze jest wyłączenie tej opcji i otwieranie okna wyników tylko wtedy, gdy jest to wymagane.

1. Aby ustawić te opcje, w menu Narzędzia wybierz pozycję **Opcje użytkownika**.
2. Kliknij kartę **Powiadomienia**.
3. W oknie dialogowym wybierz pozycję **Gdy pojawia się nowy wynik** dostępną na liście w grupie **Powiadomienia wizualne**.
4. W obszarze **Otwórz okno wynikowe** wybierz ustawienie **Nigdy**.

Ułatwienia dostępu z użyciem klawiatury

Funkcjami produktu można sterować za pomocą klawiatury. Podobnie jak w przypadku dowolnej aplikacji systemu Windows, klawisz Alt w połączeniu z innymi klawiszami (np. kombinacja Alt+P rozwija menu Plik) służy do aktywacji menu poszczególnych okien, natomiast klawisz Tab do przechodzenia do kolejnych elementów sterujących w oknie dialogowym. Jednak w głównym oknie każdego produktu znajdują się również funkcje specjalne oraz wskazówki ułatwiające nawigację w oknach dialogowych.

W niniejszej sekcji przedstawiono podstawowe wskazówki dotyczące obsługi programu za pomocą klawiatury: od otwierania strumieni do używania okien dialogowych węzłów podczas pracy z wynikami. Dostępne są również szczegółowe listy klawiszy skrótów jeszcze bardziej ułatwiające nawigację w programie.

Skróty do nawigacji w oknie głównym

Większość operacji związanych z eksploracją danych wykonuje się w oknie głównym programu IBM SPSS Modeler. Główny obszar nazywany **obszarem roboczym strumienia** służy do tworzenia i uruchamiania strumieni danych. W dolnej części okna znajdują się **palety węzłów** zawierające wszystkie dostępne węzły. Palety rozmieszczono na kartach odpowiadających typom operacji eksplorowania danych wykonywanych w określonych grupach węzłów. Przykładowo: węzły służące do wprowadzania danych do programu IBM SPSS Modeler są zgrupowane na karcie Źródła, a węzły pozwalające na wyliczanie, filtrowanie bądź wprowadzanie zmiennych znajdują się na karcie Zmienne.

Po prawej stronie okna jest dostępnych kilka narzędzi pozwalających na zarządzanie strumieniami, wynikami i projektami. W górnej części okna po prawej stronie widoczny jest obszar **Zarządzanie**, który składa się z trzech kart umożliwiających zarządzanie strumieniami, wynikami i utworzonymi modelami. Te pozycje są dostępne po wybraniu odpowiedniej karty oraz obiektu na liście. W dolnej części okna po prawej stronie widoczny jest **panel projektu** umożliwiający organizowanie pracy w postaci projektów. Dwie karty widoczne w tym obszarze odpowiadają dwóm różnym widokom projektu. W **widoku Klasy** obiekty należące do danych projektów są sortowane wg typów, natomiast w **widoku CRISP-DM** obiekty są sortowane wg odpowiedniej fazy eksploracji danych, np. Przygotowanie danych lub Modelowanie. Rożne tryby pracy okna programu IBM SPSS Modeler omówiono w systemie pomocy i podręczniku użytkownika.

Poniżej przedstawiono listę skrótów używanych podczas pracy w oknie głównym programu IBM SPSS Modeler oraz podczas tworzenia strumieni. Skróty klawiaturowe do okien dialogowych i wyników przedstawiono w tematach zawartych poniżej. Należy pamiętać, że te skróty są dostępne tylko z poziomu okna głównego.

Tabela 39. Okno główne — skróty

Klawisz skrótu	**Funkcja**
Ctrl+F5	Przenosi fokus do palety węzłów.
Ctrl+F6	Przenosi fokus do obszaru roboczego strumienia.
Ctrl+F7	Przenosi fokus do panelu zarządzania.
Ctrl+F8	Przenosi fokus do panelu projektu.

Tabela 40. Węzeł i strumień — skróty

Klawisz skrótu	**Funkcja**
Ctrl+N	Tworzy nowy pusty obszar roboczy strumienia.
Ctrl+O	Wyświetla okno dialogowe Otwórz, za pomocą którego można wybrać i otworzyć istniejący strumień.
Ctrl+klawisze numeryczne	Przenosi fokus do odpowiedniej karty w oknie lub w panelu. Przykładowo: w oknie lub panelu zawierającym karty kombinacja klawiszy Ctrl+1 powoduje przejście do pierwszej karty od lewej, Ctrl+2 — do drugiej itd.
Ctrl+strzałka w dół	W palecie węzłów przenosi fokus z karty palety do pierwszego węzła poniżej karty.
Ctrl+strzałka w górę	W palecie węzłów przenosi fokus z węzła do karty palety.

Tabela 40. Węzeł i strumień — skróty (kontynuacja)	
Klawisz skrótu	**Funkcja**
Wprowadzanie	Po wybraniu węzła w palecie węzłów (dotyczy też udoskonalonych modeli na utworzonej palecie modeli) naciśnięcie tych klawiszy dodaje węzeł do obszaru roboczego strumienia. Naciśnięcie klawisza **Enter**, gdy węzeł jest już wybrany w obszarze roboczym, otwiera okno dialogowe tego węzła.
Ctrl+Enter	Po wybraniu węzła na tej palecie jest on dodawany do obszaru roboczego strumienia bez wybierania go. Fokus jest przenoszony do pierwszego węzła na palecie.
Alt+Enter	Po wybraniu węzła na tej palecie jest on dodawany do obszaru roboczego strumienia i wybierany. Fokus jest przenoszony do pierwszego węzła w palecie.
Shift+spacja	Gdy w palecie fokus jest ustawiony na węźle lub komentarzu, zaznacza lub usuwa zaznaczenie tego węzła lub komentarza. Jeśli zaznaczone są inne węzły lub komentarze, użycie tej operacji powoduje usunięcie ich zaznaczenia.
Ctrl+Shift+spacja	Gdy w strumieniu fokus jest ustawiony na węźle lub komentarzu lub na palecie fokus jest ustawiony na komentarzu bądź węźle, zaznacza lub usuwa zaznaczenie tego węzła lub komentarza. Nie wpływa to na żadne inne wybrane węzły ani komentarze.
Strzałka w lewo/prawo	Jeśli fokus jest ustawiony na obszarze roboczym strumienia, przenosi cały strumień poziomo na ekranie. Jeśli fokus jest ustawiony na karcie palety, przełącza karty. Jeśli fokus jest ustawiony na węźle palety, przełącza między węzłami w palecie.
Strzałka w górę/w dół	Jeśli fokus jest ustawiony na obszarze roboczym strumienia, przenosi cały strumień pionowo na ekranie. Jeśli fokus jest ustawiony na węźle palety, przełącza między węzłami w palecie. Jeśli fokus jest ustawiony na podpalecie, przełącza między innymi podpaletami na tej karcie palety.
Alt+strzałka w lewo/prawo	Przenosi wybrane węzły i komentarze na obszarze roboczym strumienia poziomo w kierunku wybranym za pomocą klawiszy strzałek.
Alt+strzałka w górę/w dół	Przenosi wybrane węzły i komentarze na obszarze roboczym strumienia pionowo w kierunku wybranym za pomocą klawiszy strzałek.
Ctrl+A	Zaznacza wszystkie węzły w strumieniu.
Ctrl+Q	Gdy fokus jest ustawiony na węźle, zaznacza ten węzeł i wszystkie poniższe węzły oraz usuwa zaznaczenie wszystkich powyższych węzłów.
Ctrl+W	Gdy fokus jest ustawiony na zaznaczonym węźle, usuwa zaznaczenie tego węzła i wszystkich węzłów zaznaczonych poniżej.
Ctrl+Alt+D	Powiela wybrany węzeł.
Ctrl+Alt+L	Po wybraniu modelu użytkowego w strumieniu otwierane jest okno dialogowe Wstaw umożliwiające załadowanie do strumienia modelu zapisanego w pliku `.nod`.

Tabela 40. Węzeł i strumień — skróty (kontynuacja)	
Klawisz skrótu	**Funkcja**
Ctrl+Alt+R	Wyświetla kartę Adnotacje odpowiadającą wybranemu węzłowi i umożliwia zmianę nazwy węzła.
Ctrl+Alt+U	Tworzy Dane niestandardowe (węzeł danych wejściowych użytkownika).
Ctrl+Alt+C	Włącza lub wyłącza pamięć podręczną węzła.
Ctrl+Alt+F	Opróżnia pamięć podręczną węzła.
Tabulator	W obszarze roboczym strumienia pozwala przechodzić między wszystkimi węzłami źródłowymi i komentarzami w bieżącym strumieniu. Na palecie węzłów przechodzi miedzy węzłami dostępnymi w palecie. Na wybranej podpalecie przechodzi do pierwszego węzła dostępnego w podpalecie.
Shift+Tab	Wykonuje operację analogiczną do naciśnięcia klawisza Tab, ale w odwrotnej kolejności.
Ctrl+Tab	Gdy fokus jest ustawiony na panelu zarządzania lub projektu, przenosi fokus do obszaru roboczego strumienia. Gdy fokus jest ustawiony na palecie węzłów, przenosi fokus między węzłem i jego kartą palet.
Dowolny klawisz alfabetyczny	Gdy fokus znajduje się na węźle w bieżącym strumieniu, przechodzi fokus i przechodzi do następnego węzła, którego nazwa zaczyna się od litery wybranej na klawiaturze.
F1	Otwiera system pomocy, wybierając temat odpowiadający elementowi, na którym ustawiono fokus.
F2	Rozpoczyna proces łączenia w przypadku węzła wybranego w obszarze roboczym. Za pomocą klawisza Tab można przejść do węzła wybranego w obszarze roboczym. Naciśnięcie klawiszy Shift+spacja powoduje zakończenie połączenia.
F3	Usuwa wszystkie połączenia w przypadku węzła wybranego w obszarze roboczym.
F6	Przenosi fokus między panelem zarządzania, panelem projektu i paletami węzłów.
F10	Otwiera menu Plik.
Shift+F10	Otwiera menu podręczne węzła lub strumienia.
Usuń	Usuwa wybrany węzeł z obszaru roboczego.
Esc	Zamyka menu podręczne lub okno dialogowe.
Ctrl+Alt+X	Rozwija Superwęzeł.
Ctrl+Alt+Z	Powiększa Superwęzeł.
Ctrl+Alt+Shift+Z	Pomniejsza Superwęzeł.
Ctrl+E	Gdy fokus znajduje się w obszarze roboczym strumienia, uruchamia bieżący strumień.

W programie IBM SPSS Modeler używanych jest wiele standardowych skrótów klawiaturowych, takich jak Ctrl+C (kopiowanie). Więcej informacji można znaleźć w temacie “Używanie skrótów klawiaturowych” na stronie 21.

Okna dialogowe i tabele — skróty

Wiele skrótów klawiaturowych i klawiszy służących do lektora ekranowego jest przydatnych podczas korzystania z okien dialogowych, tabel i tabel w oknach dialogowych. Pełna lista specjalnych skrótów klawiaturowych i skrótów do obsługi lektora ekranowego znajduje się poniżej.

Tabela 41. Okno dialogowe i konstruktor wyrażeń — skróty

Klawisz skrótu	**Funkcja**
Alt+4	Zamyka wszystkie otwarte okna dialogowe lub okna wyników. Wyniki można przywrócić z poziomu karty Wyniki dostępnej w panelu zarządzania.
Ctrl+End	Gdy fokus znajduje się na dowolnym elemencie sterującym w Konstruktorze wyrażeń, punkt wstawiania zostanie przeniesiony na koniec wyrażenia.
Ctrl+1	W Konstruktorze wyrażeń przenosi fokus do elementu sterującego edytowaniem wyrażeniem.
Ctrl+2	W Konstruktorze wyrażeń przenosi fokus do listy funkcji.
Ctrl+3	W Konstruktorze wyrażeń przenosi fokus do listy zmiennych.

Tabela — skróty

Skróty działające w tabelach umożliwiają wykonywanie operacji w tabelach wynikowych oraz korzystanie z elementów sterujących w oknach dialogowych dotyczących węzłów. W celu przejścia między komórkami w tabeli zazwyczaj jest używany klawisz Tab, a kombinacja Ctrl+Tab pozwala na zaprzestanie używania elementu sterującego tabelą. *Uwaga*: niekiedy program do odczytywania zawartości ekranu może natychmiast rozpocząć odczytywanie treści komórki. Jedno- lub dwukrotne naciśnięcie klawisza strzałki resetuje oprogramowanie i rozpoczyna się odczytywanie na głos.

Tabela 42. Tabela — skróty

Klawisz skrótu	**Funkcja**
Ctrl+W	W tabelach: odczytuje krótki opis wybranego **W**iersza. Przykładowo: „Wybrany wiersz 2, wartości to płeć, flaga, k/m itd.”
Ctrl+Alt+W	W tabelach: odczytuje długi opis wybranego **W**iersza. Przykładowo: „Wybrany wiersz 2, wartość to zmienna = płeć, typ = flaga, płeć = k/m itd.”
Ctrl+D	W tabelach: odczytuje krótki **O**pis wybranego obszaru. Przykładowo: „Wybrano obszar o rozmiarze: 1 wiersz na 6 kolumn”.
Ctrl+Alt+D	W tabelach: odczytuje długi **O**pis wybranego obszaru. Przykładowo: „Wybrano obszar o rozmiarze: 1 wiersz na 6 kolumn. Wybrane kolumny to Zmienna, Typ, Brak. Wybrano wiersz 1”.
Ctrl+T	W tabelach: odczytuje krótki opis wybranych kolumn. Przykładowo: „Zmienne, Typ, Brak”.
Ctrl+Alt+T	W tabelach: odczytuje długi opis wybranych kolumn. Przykładowo: „Wybrane kolumny to Zmienne, Typ, Brak”.
Ctrl+R	W tabelach: odczytuje liczbę **R**ekordów w tabeli.

Tabela 42. Tabela — skróty (kontynuacja)	
Klawisz skrótu	**Funkcja**
Ctrl+Alt+R	W tabelach: odczytuje liczbę **R**ekordów w tabeli oraz nazwy kolumn.
Ctrl+I	W tabelach: odczytuje **I**nformacje o komórce lub jej treść (dotyczy komórek z fokusem).
Ctrl+Alt+I	W tabelach: odczytuje długie opisy **I**nformacji o komórce (nazwa kolumny i treść komórki); dotyczy komórek z fokusem.
Ctrl+G	W tabelach: odczytuje krótkie informacje o**g**ólne o wybranych elementach.
Ctrl+Alt+G	W tabelach: odczytuje długie informacje o**g**ólne o wybranych elementach.
Ctrl+Q	W tabelach: pozwala na szybkie przełączanie komórek tabeli. Naciśnięcie klawiszy Ctrl+Q pozwala na odczytywanie długich opisów np. „Płeć=Kobieta” podczas poruszania się po tabeli za pomocą klawiszy strzałek. Ponowne naciśnięcie klawiszy Ctrl+Q spowoduje wybranie krótkich opisów (treść komórek).
F8	W tabelach: gdy fokus znajduje się w tabeli przenosi go na nagłówek kolumny.
Spację	W tabelach: gdy fokus znajduje się na nagłówku kolumny umożliwia sortowanie w kolumnie.

Komentarze — skróty

Podczas pracy z komentarzami ekranowymi można korzystać z poniższych skrótów.

Tabela 43. Komentarze — skróty	
Klawisz skrótu	**Funkcja**
Alt+C	Pokazuje/ukrywa komentarz.
Alt+M	Wstawia nowy komentarz, jeśli komentarze są aktualnie wyświetlane; pokazuje komentarze, jeśli są one aktualnie ukryte.
Tabulator	W obszarze roboczym strumienia pozwala przechodzić między wszystkimi węzłami źródłowymi i komentarzami w bieżącym strumieniu.
Wprowadzanie	Jeśli fokus znajduje się na komentarzu, określa rozpoczęcie edycji.
Alt+Enter lub Ctrl+Tab	Kończy edycję i zapisuje wprowadzone modyfikacje.
Esc	Anuluje edycję. Zmiany wprowadzone podczas edytowania zostaną utracone.
Alt+Shift+strzałka w górę	Zmniejsza wysokość obszaru tekstowego o jedną komórkę siatki (lub jeden piksel), jeśli funkcja wyrównania do siatki jest włączona (lub wyłączona).
Alt+Shift+strzałka w dół	Zwiększa wysokość obszaru tekstowego o jedną komórkę siatki (lub jeden piksel), jeśli funkcja wyrównania do siatki jest włączona (lub wyłączona).

Tabela 43. Komentarze — skróty (kontynuacja)	
Klawisz skrótu	**Funkcja**
Alt+Shift+strzałka w lewo	Zmniejsza szerokość obszaru tekstowego o jedną komórkę siatki (lub jeden piksel), jeśli funkcja wyrównania do siatki jest włączona (lub wyłączona).
Alt+Shift+strzałka w prawo	Zwiększa szerokość obszaru tekstowego o jedną komórkę siatki (lub jeden piksel), jeśli funkcja wyrównania do siatki jest włączona (lub wyłączona).

Przeglądarka skupień i Przeglądarka modelu — skróty

Podczas pracy w oknach Przeglądarki skupień i Przeglądarki modelu dostępne są skróty klawiaturowe.

Tabela 44. Skróty ogólne - Przeglądarka skupień i Przeglądarka modelu	
Klawisz skrótu	**Funkcja**
Tabulator	Przenosi fokus do następnego elementu sterującego ekranem.
Shift+Tab	Przenosi fokus do poprzedniego elementu sterującego ekranem.
Strzałka w dół	Jeśli fokus jest ustawiony na liście rozwijanej, otwiera listę lub przechodzi do kolejnej pozycji na liście. Jeśli fokus jest ustawiony na menu, przechodzi do kolejnej pozycji w menu. Jeśli fokus jest ustawiony na wykresie miniaturowym, przechodzi do kolejnej pozycji (lub do pierwszej pozycji, jeśli fokus się znajdował na ostatniej miniaturze).
Strzałka w górę	Jeśli lista rozwijana jest otwarta, przechodzi do poprzedniej pozycji na liście. Jeśli fokus jest ustawiony na menu, przechodzi do poprzedniej pozycji w menu. Jeśli fokus jest ustawiony na wykresie miniaturowym, przechodzi do poprzedniej pozycji (lub do ostatniej pozycji, jeśli fokus się znajdował na pierwszej miniaturze).
Wprowadzanie	Zamyka otwartą listę rozwijaną, lub określa wybór w otwartym menu.
F6	Przenosi fokus między lewym i prawym panelem okna.
Strzałki w lewo i w prawo	Jeśli fokus znajduje się na karcie, przechodzi do następnej lub poprzedniej karty. Jeśli fokus znajduje się na menu, przechodzi do następnego lub poprzedniego menu.
Alt+*litera*	Wybiera przycisk lub menu, w których nazwie dana litera jest podkreślona.
Esc	Zamyka otwarte menu lub menu rozwijane.

Tylko Przeglądarka skupień

W Przeglądarce skupień dostępny jest widok Skupienia zawierający tabelę skupień wg właściwości.

Aby wybrać widok Skupień zamiast Podsumowania modelu:

1. Naciskaj klawisz Tab do momentu zaznaczenia przycisku **Wyświetl**.
2. Dwukrotnie naciśnij klawisz strzałki w dół, aby wybrać pozycję **Skupienia**.

   Teraz można wybrać pojedynczą komórkę w siatce:
3. Naciskaj klawisz Tab do momentu przejścia do ostatniej ikony na pasku narzędzi wizualizacji.

   Rysunek 19. Ikona Pokaż drzewo wizualizacji
4. Naciśnij jeden raz klawisz Tab, następnie klawisz spacji i klawisz strzałki.

Teraz dostępne są następujące skróty klawiaturowe:

Tabela 45. Skróty Przeglądarki skupień

Klawisz skrótu	**Funkcja**
Klawisz strzałki	Przenosi fokus między pojedynczymi komórkami w siatce. Widok rozkładu komórek w prawym panelu ulega zmianie zgodnie ze zmianą położenia fokusa.
Ctrl+, (przecinek)	Zaznacza lub usuwa zaznaczanie całej kolumny w siatce, w której fokus jest ustawiony na komórce. Aby dodać kolumnę do wyboru, za pomocą klawiszy strzałek przejdź do komórki w danej kolumnie i ponownie naciśnij klawisze Ctrl +,.
Tabulator	Przenosi fokus poza siatkę do następnego elementu sterującego na ekranie.
Shift+Tab	Przenosi fokus poza siatkę z powrotem do poprzedniego elementu sterującego na ekranie.
F2	Uruchamia tryb edycji (tylko komórki opisu i etykiety).
Wprowadzanie	Zapisuje wprowadzone zmiany i wyłącza tryb edycji (tylko komórki opisu i etykiety).
Esc	Wyłącza tryb edycji bez zapisywania zmian (tylko komórki opisu i etykiety).

Klawisze skrótów — przykład: tworzenie strumieni

Aby przedstawić proces tworzenia strumienia w sposób bardziej przejrzysty dla użytkowników korzystających tylko z klawiatury lub lektora ekranowego, poniżej zawarto przykładową procedurę tworzenia strumienia bez użycia myszy. W tym przykładzie zostanie utworzony strumień zawierający węzeł pliku, wyliczenia oraz histogramu. W celu utworzenia tego strumienia:

1. **Uruchom program IBM SPSS Modeler.** Po uruchomieniu programu IBM SPSS Modeler fokus jest ustawiony na karcie Ulubione w palecie węzłów.
2. **Ctrl+strzałka w dół.** Przenosi fokus z karty do jej zawartości.
3. **Strzałka w prawo.** Przenosi fokus do węzła pliku.
4. **Spacja.** Wybiera węzeł pliku.

5. **Ctrl+Enter.** Dodaje węzeł pliku do obszaru roboczego strumienia. Użycie tej kombinacji klawiszy również nie powoduje zmiany wyboru w węźle pliku, więc kolejny dodany węzeł będzie połączony z tym węzłem.
6. **Tab.** Przenosi fokus z powrotem do palety węzłów.
7. **Strzałka w prawo x 4.** Przechodzi do węzła Wyliczenie.
8. **Spacja.** Wybiera węzeł Wyliczenie.
9. **Alt+Enter.** Dodaje węzeł Wyliczenie do obszaru roboczego i wybiera węzeł Wyliczenie. Teraz węzeł można połączyć z kolejnym dodanym węzłem.
10. **Tab.** Przenosi fokus z powrotem do palety węzłów.
11. **Strzałka w prawo x 5.** Przenosi fokus do węzła Histogram w palecie.
12. **Spacja.** Wybiera węzeł Histogram.
13. **Wprowadzanie.** Dodaje węzeł do strumienia i przenosi fokus do obszaru roboczego strumienia.

Przejdź do kolejnego przykładu lub zapisz strumień, aby następny przykład wykonać później.

Klawisze skrótów — przykład: edytowanie węzłów

W tym przykładzie zostanie użyty strumień utworzony w poprzednim przykładzie. Strumień składa się z węzła pliku, węzła wyliczenia oraz węzła histogramu. Początkowo fokus znajduje się na trzecim węźle w strumieniu — na węźle Histogram.

1. **Ctrl+strzałka w lewo x 2.** Przenosi fokus z powrotem do węzła pliku.
2. **Wprowadzanie.** Otwiera okno dialogowe Plik zmienny. Naciskając klawisz Tab, przejdź do pola Plik i wprowadź ścieżkę do pliku tekstowego oraz jego nazwę. Określony plik zostanie wybrany. Naciśnij klawisze Ctrl+Tab, aby przejść do dolnej części okna dialogowego. Naciskając klawisz Tab przejdź do przycisku OK i naciśnij klawisz Enter, aby zamknąć to okno dialogowe.
3. **Ctrl+strzałka w prawo.** Przenosi fokus na drugi węzeł — węzeł wyliczeń.
4. **Wprowadzanie.** Otwiera okno dialogowe węzła wyliczeń. Naciskając klawisz Tab, wybierz zmienne i określ warunki wyliczeń. Naciśnij klawisze Ctrl+Tab, aby przejść do przycisku OK, i naciśnij klawisz Enter, aby zamknąć to okno dialogowe.
5. **Ctrl+strzałka w prawo.** Przenosi fokus do trzeciego węzła — węzła Histogram.
6. **Wprowadzanie.** Otwiera okno dialogowe węzła Histogram. Naciskając klawisz Tab, wybierz zmienne i określ opcje wykresu. Na liście rozwijanej: za pomocą klawisza strzałki w dół otwórz listę i wyróżnij element listy, a następnie naciśnij klawisz Enter, aby wybrać daną pozycję. Naciskaj klawisz Tab, aby przejść do przycisku OK, i naciśnij klawisz Enter, aby zamknąć to okno dialogowe.

Teraz można dodać kolejne węzły lub uruchomić bieżący strumień. Podczas tworzenia strumieni należy pamiętać o następujących poradach:

- Podczas ręcznego łączenia węzłów za pomocą klawisza F2 można utworzyć punkt początkowy połączenia, użyć klawisza Tab, aby przejść do punktu końcowego, a następnie za pomocą klawiszy Shift+spacja zakończyć tworzenie połączenia.
- Używając klawisza F3, można usunąć wszystkie połączenia wybranego węzła w obszarze roboczym.
- Po utworzeniu strumienia za pomocą klawiszy Ctrl+E można uruchomić bieżący strumień.

Dostępna jest pełna lista skrótów klawiaturowych. Więcej informacji można znaleźć w temacie “Skróty do nawigacji w oknie głównym” na stronie 257.

Używanie lektora ekranowego

Na rynku dostępnych jest lektorów ekranowych. W programie IBM SPSS Modeler skonfigurowano opcje zapewniające obsługę oprogramowania JAWS dla systemu Windows z zastosowaniem rozwiązania Java Access Bridge instalowanego wraz z IBM SPSS Modeler. Jeśli na komputerze zainstalowano oprogramowanie JAWS, należy je uruchomić przed uruchomieniem IBM SPSS Modeler.

Uwaga: Aby używać oprogramowania JAWS wraz z SPSS Modeler potrzeba co najmniej 6 GB miejsca na dysku.

Ze względu na sposób przedstawiania procesów eksploracji danych w programie IBM SPSS Modeler z wykresami najlepiej pracuje się w sposób wizualny. Jednak możliwe jest również używanie wyników w trybie tekstowym — z zastosowaniem lektora ekranowego.

Uwaga: Niektóre funkcje wspomagające nie działają na systemach 64-bitowych. Jest to spowodowane niedostosowaniem rozwiązania Java Access Bridge do pracy w środowiskach 64-bitowych.

Używanie pliku słownika IBM SPSS Modeler

Plik słownika IBM SPSS Modeler (*Awt.JDF*) można uwzględnić w oprogramowaniu JAWS. Aby użyć tego pliku:

1. Przejdź do podkatalogu /*accessibility* w katalogu instalacyjnym IBM SPSS Modeler i skopiuj plik słownika (*Awt.JDF*).
2. Skopiuj plik do katalogu zawierającego skrypty JAWS.

Jeśli użytkownik korzysta z innych aplikacji JAVA, na komputerze już może się znajdować plik o nazwie *Awt.JDF*. W takiej sytuacji używanie tego pliku słownika będzie wymagało jego ręcznej edycji.

Używanie lektora ekranowego z wynikami HTML

Podczas wyświetlania wyników w formacie HTML w oprogramowaniu IBM SPSS Modeler (z użyciem lektora ekranowego) mogą wstępować pewne problemy. Dotyczy to wielu typów wyników, np.:

- Wyników wyświetlanych na karcie Zaawansowane w przypadku węzłów Regresja, Regresja logistyczna i Redukcja wymiarów.
- Sprawdzanie wyników działania węzła.

W każdym z tych okien lub okien dialogowych na pasku narzędzi dostępne jest narzędzie, za pomocą którego można wyświetlić wynik w domyślnej przeglądarce, co pozwala na obsługę standardowego programu do odczytywania zawartości ekranu. Następnie dzięki programowi do odczytywania zawartości ekranu użytkownik może uzyskać informacje wynikowe.

Funkcje ułatwień dostępu w oknie drzewa interaktywnego

Standardowy tryb wyświetlania modelu drzewa decyzyjnego w oknie drzewa interaktywnego może powodować problemy w przypadku korzystania z programów do odczytywania zawartości ekranu. Aby uzyskać dostęp do wersji z ułatwieniami dostępu, w menu drzewa interaktywnego kliknij pozycje:

Widok > **Dostępne okno**

Zostanie wyświetlony ekran podobny do standardowej mapy drzewa, ale oprogramowanie JAWS będzie go mogło odczytać poprawnie. Za pomocą klawiszy strzałek użytkownik może się poruszać we wszystkich kierunkach. Poruszanie się po oknie z ułatwieniami dostępu powoduje również przemieszczanie fokusu w oknie drzewa interaktywnego. Naciśnięcie klawisza spacja pozwala zmienić wybrany element, a kombinacja Ctrl+spacja powoduje rozszerzenie aktualnie wybranego zakresu.

Porady

Stosując kilka z opisanych w niniejszym dokumencie porad, można zwiększyć dostępność oprogramowania IBM SPSS Modeler dla użytkowników. Poniżej przedstawiono ogólne wskazówki dotyczące pracy z programem IBM SPSS Modeler.

- **Zamykanie rozszerzonych pól tekstowych.** Rozszerzone pola tekstowe można zamknąć za pomocą klawiszy Ctrl+Tab. Należy zauważyć, że kombinacja Ctrl+Tab jest również używana do zamykania elementów sterujących tabeli.
- **Korzystanie z klawisza Tab, a nie z klawiszy strzałek.** Podczas wybierania opcji w oknie dialogowym zaleca się korzystanie z klawisza Tab w celu przechodzenia między kolejnymi przyciskami opcji. W tym kontekście klawisze strzałek nie będą działać.

- **Lista rozwijana.** Na rozwijanej liście w oknach dialogowych pozycje można wybierać za pomocą klawisza spacja lub Esc. Za pomocą klawisza Esc można również zamykać listy rozwijane, które nie zamykają się po przejściu (za pomocą klawisza Tab) do kolejnego elementu sterującego.
- **Status wykonania**. Podczas używania strumienia w ramach dużej bazy danych w oprogramowaniu JAWS mogą wystąpić opóźnienia odczytu statusu strumienia. Aby aktualizować status, należy co pewien czas naciskać klawisz Ctrl.
- **Używanie palet węzłów.** Niekiedy po pierwszym przejściu do karty na paletach węzłów oprogramowanie JAWS odczyta ciąg „pole grupy”, zamiast nazwy węzła. W takiej sytuacji za pomocą klawiszy Ctrl+strzałka w prawo i Ctrl+strzałka w lewo można zresetować lektor ekranowy i odczytana zostanie nazwa węzła.
- **Odczytywanie menu.** Niekiedy po pierwszym otwarciu menu program JAWS może nie odczytać pierwszej pozycji w menu. Jeśli podejrzewasz, że pierwsza pozycja nie została odczytana, naciśnij kolejno klawisze strzałki w dół i strzałki w górę, aby usłyszeć nazwę pierwszej pozycji w menu.
- **Menu kaskadowe.** Program JAWS nie odczytuje pierwszego poziomu menu kaskadowego. Jeśli usłyszysz przerwę w odczycie podczas przechodzenia między pozycjami menu, naciśnij klawisz strzałki w prawo, aby odsłuchać pozycje menu podrzędnego.

Ponadto jeśli zainstalowano IBM SPSS Modeler Text Analytics, poniższe porady mogą ułatwić korzystanie z interaktywnego interfejsu pulpitu roboczego.

- **Przechodzenie do okien dialogowych.** Po przejściu do okna dialogowego wymagane może być naciśnięcie klawisza Tab, aby ustawić fokus na pierwszym elemencie sterującym.
- **Zamykanie rozszerzonych pól tekstowych.** Rozszerzone pola tekstowe można zamknąć za pomocą klawiszy Ctrl+Tab i przejść do kolejnego elementu sterującego. Należy zauważyć, że kombinacja Ctrl+Tab jest również używana do zamykania elementów sterujących tabeli.
- **Wprowadzanie pierwszej litery w celu wyszukania elementu w liście drzewa.** Wyszukując elementy w panelu kategorii, wyodrębnionych wyników lub w drzewie biblioteki, można wpisać pierwszą literę nazwy elementu, gdy fokus znajduje się na panelu. W ten sposób zostanie wybrane następne wystąpienie elementu, którego nazwa rozpoczyna się od wybranej litery.
- **Lista rozwijana.** Na liście rozwijanej w oknach dialogowych pozycje można wybierać za pomocą klawisza spacji.

Problemy podczas współpracy z innym oprogramowaniem

Podczas testowania kompatybilności oprogramowania IBM SPSS Modeler z lektorami ekranowymi, takimi jak JAWS, nasz zespół programistyczny zauważył, że używanie serwera Systems Management Server (SMS) w organizacji może negatywnie wpływać na odczytywanie przez program JAWS zawartości okien aplikacji napisanych w języku Java, np. IBM SPSS Modeler. Zablokowanie działania serwera SMS rozwiązuje ten problem. Dodatkowe informacje na temat serwera SMS zawarto w serwisie WWW firmy Microsoft.

JAWS i Java

Różne wersje programu JAWS w różny sposób współpracują z aplikacjami napisanymi w języku Java. Oprogramowanie IBM SPSS Modeler jest kompatybilne ze wszystkimi najnowszymi wersjami JAWS, jednak niektóre wersje mogą powodować niewielkie problemy w przypadku systemów opartych na języku Java. Należy się zapoznać z informacjami dotyczącymi programu JAWS dla systemu Windows dostępnymi pod adresem http://www.FreedomScientific.com.

Używanie wykresów w programie IBM SPSS Modeler

Wizualne reprezentacje informacji w postaci histogramów, wykresów ewaluacyjnych, wykresów wielokrotnych i wykresów rozrzutu stanowią problem dla programów do odczytywania zawartości ekranu. Należy jednak pamiętać, że wykresy sieciowe i rozkłady można wyświetlać z zastosowaniem podsumowania tekstowego. Ta funkcja jest dostępna z poziomu okna raportu wynikowego.

Rozdział 18. obsługa Unicode

Obsługa kodowania Unicode w IBM SPSS Modeler

Oprogramowanie IBM SPSS Modeler w pełni obsługuje kodowanie Unicode w rozwiązaniach IBM SPSS Modeler oraz IBM SPSS Modeler Server. Dzięki temu możliwa jest wymiana danych (takich jak wielojęzyczne bazy danych) z innymi aplikacjami obsługującymi kodowanie bez utraty informacji, do czego dochodzi najczęściej podczas przekształcania między regionalnymi systemami kodowania.

- Dane Unicode w oprogramowaniu IBM SPSS Modeler są przechowywane wewnętrznie i możliwe jest ich odczytywanie oraz zapisywanie w ramach baz danych bez utraty informacji.
- W języku IBM SPSS Modeler można odczytywać i zapisywać pliki UTF-8. Podczas importowania i eksportowania domyślnie wybierane jest kodowanie regionalne, ale można wybrać też kodowanie UTF-8. To ustawienie można określić w węzłach importowania i eksportowania plików. Kodowanie domyślne można też zmienić w oknie dialogowym właściwości strumienia. Więcej informacji można znaleźć w temacie “Określanie opcji ogólnych strumienia” na stronie 42.
- Pliki Statistics, SAS i pliki zawierające dane tekstowe przechowywane z zastosowaniem kodowania regionalnego zostaną przekształcone do formatu UTF-8 podczas importowania i przekodowane z powrotem w czasie eksportowania. Jeśli użytkownik zapisuje dane w dowolnym pliku, a zawierają one znaki Unicode, które nie istnieją w regionalnym zestawie znaków, to te znaki zostaną zastąpione, a następnie zostanie wyświetlone ostrzeżenie. Taka sytuacja zachodzi wyłącznie, gdy dane zaimportowano ze źródła danych obsługującego kodowanie Unicode (baza danych lub plik UTF-8) i zawierają one odmienne znaki regionalne lub pochodzące z wielu rodzajów kodowań bądź zestawów znaków.
- Obrazy w oprogramowaniu IBM SPSS Modeler Solution Publisher są kodowane z zastosowaniem kodowania UTF-8 i można je bezproblemowo przenosić między platformami i systemami.

Informacje na temat kodowania Unicode

Celem przyświecającym twórcom standardu Unicode było opracowanie spójnego sposobu kodowania tekstów w różnych językach zapewniającego proste udostępnianie plików miedzy różnymi krajami, regionami i w ramach różnych aplikacji. Standard Unicode (aktualnie w wersji 4.0.1) definiuje zestaw znaków stanowiący nadrzędny zestaw względem wszystkich powszechnie używanych zestawów na całym świecie; a do każdego znaku jest przypisana niepowtarzalna nazwa i wartość. Znaki wraz z wartościami są zgodne z zestawem Universal Character Set (UCS) zdefiniowanym w normie ISO-10646. Dodatkowe informacje: Serwis kodowania Unicode.

Uwagi

Niniejsza publikacja została przygotowana z myślą o produktach i usługach oferowanych w Stanach Zjednoczonych. Materiał ten jest również dostępny w IBM w innych językach. Jednakże w celu uzyskania dostępu do takiego materiału istnieje konieczność posiadania egzemplarza produktu w takim języku.

Produktów, usług lub opcji opisywanych w tym dokumencie IBM nie musi oferować we wszystkich krajach. Informacje o produktach i usługach dostępnych w danym kraju można uzyskać od lokalnego przedstawiciela IBM. Odwołanie do produktu, programu lub usługi IBM nie oznacza, że można użyć wyłącznie tego produktu, programu lub usługi IBM. Zamiast nich można zastosować ich odpowiednik funkcjonalny pod warunkiem, że nie narusza to praw własności intelektualnej IBM. Jednakże cała odpowiedzialność za ocenę przydatności i sprawdzenie działania produktu, programu lub usługi pochodzących od producenta innego niż IBM spoczywa na użytkowniku.

IBM może posiadać patenty lub złożone wnioski patentowe na towary i usługi, o których mowa w niniejszej publikacji. Przedstawienie niniejszej publikacji nie daje żadnych uprawnień licencyjnych do tychże patentów. Pisemne zapytania w sprawie licencji można przesyłać na adres:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
U.S.A.

Zapytania dotyczące zestawów znaków dwubajtowych (DBCS) należy kierować do lokalnych działów własności intelektualnej IBM (IBM Intellectual Property Department) lub wysłać je na piśmie na adres:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan, Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokio 103-8510, Japonia

INTERNATIONAL BUSINESS MACHINES CORPORATION DOSTARCZA TĘ PUBLIKACJĘ W STANIE, W JAKIM SIĘ ZNAJDUJE ("AS IS") BEZ UDZIELANIA JAKICHKOLWIEK GWARANCJI (RĘKOJMIĘ RÓWNIEŻ WYŁĄCZA SIĘ), WYRAŹNYCH LUB DOMNIEMANYCH, A W SZCZEGÓLNOŚCI DOMNIEMANYCH GWARANCJI PRZYDATNOŚCI HANDLOWEJ, PRZYDATNOŚCI DO OKREŚLONEGO CELU ORAZ GWARANCJI, ŻE PUBLIKACJA TA NIE NARUSZA PRAW OSÓB TRZECICH. Ustawodawstwa niektórych krajów nie dopuszczają zastrzeżeń dotyczących gwarancji wyraźnych lub domniemanych w odniesieniu do pewnych transakcji; w takiej sytuacji powyższe zdanie nie ma zastosowania.

Informacje zawarte w niniejszej publikacji mogą zawierać nieścisłości techniczne lub błędy drukarskie. Informacje te są okresowo aktualizowane, a zmiany te zostaną uwzględnione w kolejnych wydaniach tej publikacji. IBM zastrzega sobie prawo do wprowadzania ulepszeń i/lub zmian w produktach i/lub programach opisanych w tej publikacji w dowolnym czasie, bez wcześniejszego powiadomienia.

Wszelkie wzmianki w tej publikacji na temat stron internetowych firm innych niż IBM zostały wprowadzone wyłącznie dla wygody użytkowników i w żadnym razie nie stanowią zachęty do ich odwiedzania. Materiały dostępne na tych stronach nie są częścią materiałów opracowanych dla tego produktu IBM, a użytkownik korzysta z nich na własną odpowiedzialność.

IBM ma prawo do używania i rozpowszechniania informacji przysłanych przez użytkownika w dowolny sposób, jaki uzna za właściwy, bez żadnych zobowiązań wobec ich autora.

Licencjobiorcy tego programu, którzy chcieliby uzyskać informacje na temat programu w celu: (i) wdrożenia wymiany informacji między niezależnie utworzonymi programami i innymi programami (łącznie z tym opisywanym) oraz (ii) wspólnego wykorzystywania wymienianych informacji, powinni skontaktować się z:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
U.S.A.

Informacje takie mogą być udostępnione, o ile spełnione zostaną odpowiednie warunki, w tym, w niektórych przypadkach, zostanie uiszczona stosowna opłata.

Licencjonowany program opisany w niniejszej publikacji oraz wszystkie inne licencjonowane materiały dostępne dla tego programu są dostarczane przez IBM na warunkach określonych w Umowie IBM z Klientem, Międzynarodowej Umowie Licencyjnej IBM na Program lub w innych podobnych umowach zawartych między IBM i użytkownikami.

Dane dotyczące wydajności i cytowane przykłady zostały przedstawione jedynie w celu zobrazowania sytuacji. Faktyczne wyniki dotyczące wydajności mogą się różnić w zależności do konkretnych warunków konfiguracyjnych i operacyjnych.

Informacje dotyczące produktów innych podmiotów niż IBM zostały uzyskane od dostawców tych produktów, z ich publicznych ogłoszeń lub innych dostępnych publicznie źródeł. IBM nie testował tych produktów i nie może potwierdzić dokładności pomiarów wydajności, kompatybilności ani żadnych innych danych związanych z produktami firm innych niż IBM. Pytania dotyczące możliwości produktów firm innych niż IBM należy kierować do dostawców tych produktów.

Wszelkie stwierdzenia dotyczące przyszłych kierunków rozwoju i zamierzeń IBM mogą zostać zmienione lub wycofane bez powiadomienia.

Publikacja ta zawiera przykładowe dane i raporty używane w codziennej pracy. W celu kompleksowego ich zilustrowania, podane przykłady zawierają nazwiska osób prywatnych, nazwy przedsiębiorstw oraz nazwy produktów. Wszystkie te nazwy/nazwiska są fikcyjne i jakiekolwiek podobieństwo do istniejących nazw/nazwisk jest całkowicie przypadkowe.

Znaki towarowe

IBM, logo IBM i ibm.com są znakami towarowymi lub zastrzeżonymi znakami towarowymi International Business Machines Corp. zarejestrowanymi w wielu systemach prawnych na całym świecie. Pozostałe nazwy produktów i usług mogą być znakami towarowymi IBM lub innych przedsiębiorstw. Aktualna lista znaków towarowych IBM dostępna jest w serwisie WWW IBM, w sekcji "Copyright and trademark information" Informacje o prawach autorskich i znakach towarowych), pod adresem www.ibm.com/legal/copytrade.shtml.

Adobe, logo Adobe, PostScript oraz logo PostScript są znakami towarowymi lub zastrzeżonymi znakami towarowymi Adobe Systems Incorporated w Stanach Zjednoczonych i/lub w innych krajach.

Intel, logo Intel, Intel Inside, logo Intel Inside, Intel Centrino, logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium i Pentium są znakami towarowymi lub zastrzeżonymi znakami towarowymi Intel Corporation lub przedsiębiorstw podporządkowanych Intel Corporation w Stanach Zjednoczonych i w innych krajach.

Linux jest zastrzeżonym znakiem towarowym Linusa Torvaldsa w Stanach Zjednoczonych i/lub w innych krajach.

Microsoft, Windows, Windows NT oraz logo Windows są znakami towarowymi Microsoft Corporation w Stanach Zjednoczonych i/lub w innych krajach.

UNIX jest zastrzeżonym znakiem towarowym The Open Group w Stanach Zjednoczonych i/lub w innych krajach.

Java oraz wszystkie znaki towarowe i logo dotyczące języka Java są znakami towarowymi lub zastrzeżonymi znakami towarowymi Oracle i/lub przedsiębiorstw afiliowanych.

Warunki dotyczące dokumentacji produktu

Zezwolenie na korzystanie z tych publikacji jest przyznawane na poniższych warunkach.

Zakres stosowania

Niniejsze warunki stanowią uzupełnienie warunków używania serwisu WWW IBM.

Użytek osobisty

Użytkownik ma prawo kopiować te publikacje do własnego, niekomercyjnego użytku pod warunkiem zachowania wszelkich uwag dotyczących praw własności. Użytkownik nie ma prawa dystrybuować ani wyświetlać tych publikacji czy ich części, ani też wykonywać na ich podstawie prac pochodnych bez wyraźnej zgody IBM.

Użytek służbowy

Użytkownik ma prawo kopiować te publikacje, dystrybuować je i wyświetlać wyłącznie w ramach przedsiębiorstwa Użytkownika pod warunkiem zachowania wszelkich uwag dotyczących praw własności. Użytkownik nie ma prawa wykonywać na podstawie tych publikacji ani ich fragmentów prac pochodnych, kopiować ich, dystrybuować ani wyświetlać poza przedsiębiorstwem Użytkownika bez wyraźnej zgody IBM.

Prawa

Z wyjątkiem zezwoleń wyraźnie udzielonych w niniejszym dokumencie, nie udziela się jakichkolwiek innych zezwoleń, licencji ani praw, wyraźnych czy domniemanych, odnoszących się do tych publikacji czy jakichkolwiek informacji, danych, oprogramowania lub innej własności intelektualnej, o których mowa w niniejszym dokumencie.

IBM zastrzega sobie prawo do anulowania zezwolenia przyznanego w niniejszym dokumencie w każdej sytuacji, gdy, według uznania IBM, korzystanie z tych publikacji jest szkodliwe dla IBM lub jeśli IBM uzna, że warunki niniejszego dokumentu nie są przestrzegane.

Użytkownik ma prawo pobierać, eksportować lub reeksportować niniejsze informacje pod warunkiem zachowania bezwzględnej i pełnej zgodności z obowiązującym prawem i przepisami, w tym ze wszelkimi prawami i przepisami eksportowymi Stanów Zjednoczonych.

IBM NIE UDZIELA JAKICHKOLWIEK GWARANCJI, W TYM TAKŻE RĘKOJMI, DOTYCZĄCYCH TREŚCI TYCH PUBLIKACJI. PUBLIKACJE TE SĄ DOSTARCZANE W STANIE, W JAKIM SIĘ ZNAJDUJĄ ("AS-IS") BEZ UDZIELANIA JAKICHKOLWIEK GWARANCJI (RĘKOJMIĘ RÓWNIEŻ WYŁĄCZA SIĘ), WYRAŹNYCH CZY DOMNIEMANYCH, A W SZCZEGÓLNOŚCI DOMNIEMANYCH GWARANCJI PRZYDATNOŚCI HANDLOWEJ, PRZYDATNOŚCI DO OKREŚLONEGO CELU CZY NIENARUSZANIA PRAW OSÓB TRZECICH.

Indeks

A

adnotacje
  folder 230
  projekt 230
  przekształcanie na komentarze 60
  strumienie 56, 60
  węzły 56, 60
analiza Champion Challenger 204, 216
analiza w oparciu o drzewo
  typowe zastosowania 25
aplikacje do eksploracji danych 26
atrybut 25
automatyzacja 145

B

baza danych
  funkcje 153, 154
blokowanie obiektów repozytorium IBM SPSS Collaboration and Deployment Services Repository 212
blokowanie węzłów 41
błąd przepełnienia stosu 235
błąd w widoku renderowania
  za mało pamięci 235
braki danych
  w rekordach 140
  współrzędnych 141
  wyrażenia CLEM 143

C

CLEM
  funkcje 153, 154
  język 161
  przykłady 145
  sprawdzanie wyrażeń 157
  tworzenie wyrażeń 153
  typy danych 161–163
  wprowadzenie 23, 145
  wyrażenia 148, 161
close_to
  funkcje przestrzenne 176
cofanie dni 43
cofanie mapowania zmiennych 63
cofnij 18
Coordinator of Processes 11
COP 11
CRISP-DM
  widok projektów 227
CRISP-DM, model procesu 27, 28
crosses
  funkcje przestrzenne 176
czas wykonywania, wyświetlanie 49
czcionki
  w oknie struktury 109
czynnik 265

D

dane
  podgląd 40
dane uwierzytelniające
  dla repozytorium IBM SPSS Collaboration and Deployment Services Repository 205
dane z szumami 26
daty
  manipulowanie 191
  przekształcanie 191
distance
  funkcje przestrzenne 176
Dobór predyktorów, węzeł
  braki danych 140
dodawanie
  do projektu 228
dodawanie nagłówków grup 124
dodawanie połączeń z serwerem IBM SPSS Modeler Server 11
dokumentacja 3
domyślna
  faza projektu 227
domyślne kodowanie strumienia 42
dostępność
  funkcje w programie IBM SPSS Modeler 255
  porady w programie IBM SPSS Modeler 265
  przykład 263, 264
dostosowywanie karty palety 247
drukowanie
  dostosowywanie wielkości tabel 128, 130
  kontrola podziałów tabeli 134
  nagłówki i stopki 120
  numery stron 121
  odstęp między elementami raportu 121
  podgląd wydruku 120
  rozmiar wykresu 121
  strumienie 20, 38
  tabele przestawne 120
  warstwy 120, 128, 130
  wykresy 120
  wyniki tekstowe 120
drzewa decyzyjne
  dostępność 265
DTD 224
dzielenie tabel
  kontrola podziałów tabeli 134

E

Edytor danych
  opcje statystyk opisowych 137
  wiele otwartych plików z danymi 136
eksploracja danych
  przykłady aplikacji 34
  strategia 27
eksportowanie

eksportowanie *(kontynuacja)*
opisy strumieni 55
PMML 224
eksportowanie wykresów 113, 119
eksportowanie wyników
format Excel 113, 116
format HTML 113
format PDF 113, 117
format PowerPoint 113
format Word 113, 115
HTML 114
raport WWW 114
etykiety
usuwanie 124
wartość 224
wstawianie nagłówków grup 124
wyświetlanie 42
zmienna 224
etykiety wartości
w oknie struktury 137
w tabelach przestawnych 137
etykiety wersji, obiekt repozytorium IBM SPSS Collaboration and Deployment Services Repository 215
etykiety zmiennych
w oknach dialogowych 136
w oknie struktury 137
w tabelach przestawnych 137
etykiety, obiekt repozytorium IBM SPSS Collaboration and Deployment Services Repository 215

F

foldery, repozytorium IBM SPSS Collaboration and Deployment Services Repository 212, 213
format Excel
eksportowanie wyników 113, 116
format PowerPoint
eksportowanie wyników 113
format Word
eksportowanie wyników 113, 115
szerokie tabele 113
format wyświetlania waluty 44
formaty czasu 43, 163, 164
formaty daty 43, 163, 164
formaty współrzędnych geoprzestrzennych 47
formaty wyświetlania
liczby 44
miejsca dziesiętne 44
naukowy 44
symbol grupowania 44
użytkownika 44
współrzędne geoprzestrzenne 47, 48
formaty wyświetlania liczb 44
funkcja @BLANK 143, 169, 198
funkcja @DIFF 191
funkcja @FIELD 143, 199
funkcja @FIELDS_BETWEEN 143, 151, 199
funkcja @FIELDS_MATCHING 143, 151, 199
funkcja @INDEX 191
funkcja @LAST_NON_BLANK 191, 198
funkcja @MAX 191
funkcja @MEAN 191
funkcja @MIN 191
funkcja @MULTI_RESPONSE_SET 152, 199
funkcja @NULL 143, 169, 198
funkcja @OFFSET
czynniki wpływające na wydajność 252
funkcja @PARTITION_FIELD 199
funkcja @PREDICTED 199
funkcja @SDEV 191
funkcja @SINCE 191
funkcja @SUM 191
funkcja @TARGET 199
funkcja @TESTING_PARTITION 199
funkcja @THIS 191
funkcja @TODAY 186
funkcja @TRAINING_PARTITION 199
funkcja @VALIDATION_PARTITION 199
funkcja abs 174
funkcja allbutfirst 179
funkcja allbutlast 179
funkcja alphabefore 179
funkcja arccos 175
funkcja arccosh 175
funkcja arcsin 175
funkcja arcsinh 175
funkcja arctan 175
funkcja arctan2 175
funkcja arctanh 175
funkcja area 176
funkcja cdf_chisq 176
funkcja cdf_f 176
funkcja cdf_normal 176
funkcja cdf_t 176
funkcja close_to 176
funkcja cos 175
funkcja cosh 175
funkcja count_equal 151, 171
funkcja count_greater_than 151, 171
funkcja count_less_than 151, 171
funkcja count_non_nulls 171
funkcja count_not_equal 151, 171
funkcja count_nulls 143, 151, 171
funkcja count_substring 179
funkcja crosses 176
funkcja date_before 171
funkcja datetime_date 169
funkcja DIFF 191
funkcja distance 176
funkcja div 174
funkcja endstring 179
funkcja first_index 152, 171
funkcja first_non_null 152, 171
funkcja first_non_null_index 152, 171
funkcja fracof 174
funkcja hasendstring 179
funkcja hasmidstring 179
funkcja hasstartstring 179
funkcja hassubstring 179
funkcja INDEX 191
funkcja integer_bitcount 177
funkcja integer_leastbit 177
funkcja integer_length 177
funkcja intof 174
funkcja is_date 169
funkcja is_datetime 169
funkcja is_integer 169
funkcja is_number 169

funkcja is_real 169
funkcja is_string 169
funkcja is_time 169
funkcja is_timestamp 169
funkcja isalphacode 179
funkcja isendstring 179
funkcja islowercode 179
funkcja ismidstring 179
funkcja isnumbercode 179
funkcja isstartstring 179
funkcja issubstring 179
funkcja issubstring_count 179
funkcja issubstring_lim 179
funkcja isuppercode 179
funkcja last_index 152, 171
funkcja LAST_NON_BLANK 191
funkcja last_non_null 152, 171
funkcja last_non_null_index 152, 171
funkcja length 179
funkcja locchar 179
funkcja locchar_back 179
funkcja log 174
funkcja log10 174
funkcja lowertoupper 179
funkcja matches 179
funkcja max 171
funkcja MAX 191
funkcja max_index 152, 171
funkcja max_n 151, 171
funkcja MEAN 191
funkcja mean_n 151, 174
funkcja member 171
funkcja min 171
funkcja MIN 191
funkcja min_index 152, 171
funkcja min_n 151, 171
funkcja mod 174
funkcja negate 174
funkcja num_points 176
funkcja OFFSET 191
funkcja oneof 179
funkcja overlap 176
funkcja pi 175
funkcja power (wykładnicza) 174
funkcja random 179
funkcja random0 179
funkcja rem 174
funkcja replace 179
funkcja replicate 179
funkcja round 174
funkcja SDEV 191
funkcja sdev_n 151, 174
funkcja sign 174
funkcja sin 175
funkcja SINCE 191
funkcja sinh 175
funkcja skipchar 179
funkcja skipchar_back 179
funkcja soundex 185
funkcja soundex_difference 185
funkcja sqrt 174
funkcja startstring 179
funkcja stripchar 179
funkcja strmember 179
funkcja subscrs 179
funkcja substring 179
funkcja substring_between 179
funkcja SUM 191
funkcja sum_n 151, 174
funkcja tan 175
funkcja tanh 175
funkcja testbit 177
funkcja THIS 191
funkcja time_before 171
funkcja to_date 169, 186
funkcja to_dateline 186
funkcja to_datetime 169
funkcja to_integer 169
funkcja to_number 169
funkcja to_real 169
funkcja to_string 169
funkcja to_time 169, 186
funkcja to_timestamp 169, 186
funkcja trim 179
funkcja trim_start 179
funkcja trimend 179
funkcja undef 198
funkcja unicode_char 179
funkcja unicode_value 179
funkcja uppertolower 179
funkcja value_at 152, 171
funkcja within 176
funkcja wykładnicza 174
funkcje
  @BLANK 143
  @FIELD 153, 199
  @GLOBAL_MAX 197
  @GLOBAL_MEAN 197
  @GLOBAL_MIN 197
  @GLOBAL_SDEV 197
  @GLOBAL_SUM 197
  @PARTITION 199
  @PREDICTED 153, 199
  @TARGET 153, 199
  baza danych 153, 154
  funkcje zdefiniowane przez użytkownika (UDF) 153
  obsługa braków danych 143
  przykłady 145
  w wyrażeniach CLEM 153
funkcje bazodanowe
  funkcje zdefiniowane przez użytkownika (UDF) 154
  w wyrażeniach CLEM 154
funkcje bitowe 177
funkcje CLEM
  bitowe 177
  braki danych 143
  datetime 186
  funkcje specjalne 199
  globalne 197
  informacyjne 169
  lista dostępnych 167
  logiczne 173
  losowa 179
  łańcuch 179
  numerycznych 174
  porównywanie 171
  prawdopodobieństwo 176
  przekształcanie 169

funkcje CLEM *(kontynuacja)*
przestrzenne 176
sekwencja 191
trygonometryczne 175
wartości puste i null 198
funkcje czasu
time_before 171, 186
time_hours_difference 186
time_in_hours 186
time_in_mins 186
time_in_secs 186
time_mins_difference 186
time_secs_difference 186
funkcje datetime
datetime_date 186
datetime_day 186
datetime_day_name 186
datetime_day_short_name 186
datetime_hour 186
datetime_in_seconds 186
datetime_minute 186
datetime_month 186
datetime_month_name 186
datetime_month_short_name 186
datetime_now datetime_second 186
datetime_time 186
datetime_timestamp 186
datetime_weekday 186
datetime_year 186
funkcje daty
date_before 171, 186
date_days_difference 186
date_in_days 186
date_in_months 186
date_in_weeks 186
date_in_years 186
date_months_difference 186
date_weeks_difference 186
date_years_difference 186
funkcja @TODAY 186
funkcje daty i czasu 163, 164
funkcje globalne 197
funkcje if, then, else 173
funkcje informacyjne 169
funkcje logiczne 173
funkcje łańcuchowe 179
funkcje numeryczne 174
funkcje porównawcze 171
funkcje prawdopodobieństwa 176
funkcje przekształcania 169
funkcje przestrzenne 176
funkcje rozkładu 176
funkcje sekwencji 191
funkcje specjalne 199
funkcje trygonometryczne 175
funkcje zdefiniowane przez użytkownika (UDF) 153, 154

G

gałęzie, modelowanie i ocenianie 55, 217–219
generowanie kodu SQL
podgląd 46
rejestrowanie 46
grupowanie wierszy i kolumn 124

H

hasło
IBM SPSS Analytic Server 12
IBM SPSS Modeler Server 10
histogramy 82
HTML
eksportowanie wyników 113, 114

I

IBM SPSS Analytic Server
połączenie 12
wiele połączeń 12
IBM SPSS Collaboration and Deployment Services 204
IBM SPSS Collaboration and Deployment Services Repository
blokowanie i usuwanie blokady obiektów 212
dane uwierzytelniające 205
foldery 212, 213
łączenie się 204
pobieranie obiektów 210
pojedyncze uwierzytelnianie 204
przeglądanie 205
przenoszenie projektów 229
usuwanie obiektów i wersji 213
właściwości obiektu 214
wyszukiwanie w 211
zapisywanie obiektów 205
IBM SPSS Modeler
dokumentacja 3
funkcje ułatwień dostępu 255
opcje 235
pierwsze kroki 9
porady i skróty 65
przegląd 9, 235
uruchamianie z wiersza komend 9
IBM SPSS Modeler Advantage 204
IBM SPSS Modeler Server
hasło 10
identyfikator użytkownika 10
nazwa domeny (Windows) 10
nazwa hosta 10, 11
numer portu 10, 11
identyfikator użytkownika
IBM SPSS Modeler Server 10
ikony
określanie opcji 20, 46
importowanie
PMML 224

J

Jakość, węzeł
braki danych 140
Java 266
JAWS 255, 264, 266
język
opcje 235
zmiana języka wyników 126
justowanie
wyniki 108, 136

K

karta Wynik
  przenoszenie wyników 108
katalog
  domyślny 236
katalog tymczasowy 13
klasy 17, 227, 228
klawisze skrótów 256, 257, 260–262
klient
  katalog domyślny 236
kodowanie 42, 267
kodowanie tekstu 42
kodowanie UTF-8 42, 267
kokpit
  wykresy 104
kolor tła 131
kolory
  ustawienie 238
kolory skryptów
  ustawienie 239
kolory w tabelach przestawnych
  obramowania 130
kolumny
  zaznaczanie w tabelach przestawnych 134
  zmiana szerokości w tabelach przestawnych 133
komentarze
  klawisze skrótów 261, 262
  w węzłach i strumieniach 56
  wyświetlanie wszystkich w strumieniu 59
komórki w tabelach przestawnych
  formaty 129
  szerokość 133
  ukrywanie 127
  wybór 134
  wyświetlanie 127
komunikaty
  wyświetlanie wygenerowanego kodu SQL 46
komunikaty o błędach 48
Konstruktor wyrażeń
  przegląd 153
  uzyskiwanie dostępu 153
  używanie 153
konwencje 168
kopiowanie 18
kopiowanie i wklejanie wyniku do innych aplikacji 111
kopiuj specjalne 111
kreator wykresów
  galeria 68
  składniki 67
  układ 67
kropka 42

L

lektory ekranowe
  przykład 263, 264
liczby 150, 161, 162
liczby całkowite 161
liczby rzeczywiste 161, 162
linie siatki
  tabele przestawne 134
listy 161, 162
logowanie się do programu IBM SPSS Modeler Server 10

Ł

ładowanie
  stany 63
  węzły 63
łańcuchy
  dopasowanie 149
  manipulowanie w wyrażeniach CLEM 149
  zastępowanie 149
Łącze URL
  IBM SPSS Analytic Server 12

M

macierz wykresów rozrzutu 95, 97
macierze SPLOM 95
mapowanie danych 65
mapowanie zmiennych 63
mapy drzew 102
metody zaznaczania
  zaznaczanie wierszy i kolumn w tabelach przestawnych 134
miejsca dziesiętne
  formaty wyświetlania 44
minimalizowanie 19
modele
  dodawanie do projektów 228
  odświeżanie 218
  zapisywanie w repozytorium IBM SPSS Collaboration and Deployment Services Repository 209
modele odświeżające 218
modele PMML
  regresja liniowa 239
  regresja logistyczna 239
modele użytkowe
  zdefiniowane 16
modelowanie
  gałąź 55
models
  eksportowanie 239
  zastępowanie 237
monity, wykonywanie 50
mysz
  używanie w oprogramowaniu IBM SPSS Modeler 21, 36

N

nadawanie nazw węzłom i strumieniom 60
nagłówki 120, 131, 132
nagłówki grup 124
nakładane wykresy rozrzutu 95
narzędzie do mapowania danych 63, 64
nawigacja
  klawisze skrótów 256
nazwa domeny (Windows)
  IBM SPSS Modeler Server 10
nazwa hosta
  IBM SPSS Modeler Server 10, 11
nazwy strumieni 60
nazwy węzłów 60
nazwy zmiennych
  w oknach dialogowych 136
niezbędne zmienne 63, 65

notacja naukowa
blokowanie w wynikach 136
format wyświetlania 44
nowe funkcje 7
num_points
funkcje przestrzenne 176
numer portu
IBM SPSS Modeler Server 10, 11
numerowanie stron 121

O

obiekty
właściwości 231
obiekty wynikowe
zapisywanie w repozytorium IBM SPSS Collaboration and Deployment Services Repository 209
obracanie etykiet 124
obramowania
wyświetlanie ukrytych krawędzi 134
obserwacja 25
obsługa Unicode 267
obsługa wartości pustych
funkcje CLEM 198
obszar roboczy 14
obszar roboczy strumienia
settings 46
ocenianie
gałąź 55, 217–219
odstęp
usuwanie z łańcuchów 149, 179
odśwież
węzły źródłowe 42
odświeżenie modelu 216
okna dialogowe
porządek wyświetlania zmiennych 136
wyświetlanie etykiet zmiennych 136
wyświetlanie nazw zmiennych 136
okno dialogowe podczas uruchamiania 238
Okno drzewa interaktywnego
dostępność 265
okno główne 14
opcje
etykietowanie 137
IBM SPSS Modeler 235
ogólne 136
PMML 239
pokaż 238
Przeglądarka 136
składnia 239
statystyki opisowe w Edytorze danych 137
użytkownik 236
właściwości strumienia 41–44, 46–49
wygląd tabel przestawnych 137
opcje użytkownika 236
opcje wdrożenia 216
operator and 173
operator mniejszości 171
operator nierówności 171
operator not 173
operator or 173
operator równości 171
operator większości 171
operatory
operatory *(kontynuacja)*
łączenie łańcuchów 169
w wyrażeniach CLEM 153
opisy strumieni 53, 55
ostrzeżenia
określanie opcji 237
otwieranie
models 63
projekty 228
stany 63
strumienie 63
węzły 63
wyniki 63
overlap
funkcje przestrzenne 176

P

palety
dostosowywanie 245
palety modeli 209
pamięć
błąd przepełnienia stosu 235
zarządzanie 235, 236
pamięć podręczna
opróżnianie 39, 42
uaktywnianie 236
ustawianie pamięci podręcznej 38
zapisywanie 39
parametry
budowanie modelu 217
monity wykonywania 50
ocenianie 217
sesja 50, 51
strumień 50, 51
typ 51
w wyrażeniach CLEM 157
parametry sesji 50, 51
parametry strumienia 50, 51
pasek narzędzi 18
PDF
eksportowanie wyników 113, 117
pierwszeństwo 164
pierwszeństwo operatorów 164
pionowy tekst etykiet 124
piramidy populacyjne 82, 92
plik słownika 264
pliki
dodawanie pliku tekstowego do Edytora raportów 110
pliki BMP
eksportowanie wykresów 113, 119
pliki danych
wiele otwartych plików z danymi 136
pliki danych tekstowych
kodowanie 267
pliki dzienników
wyświetlanie wygenerowanego kodu SQL 46
pliki EPS
eksportowanie wykresów 113, 119
pliki JPEG
eksportowanie wykresów 113, 119
pliki kopii zapasowej strumienia
przywracanie 61
pliki metafile

pliki metafile *(kontynuacja)*
eksportowanie wykresów 113
pliki PNG
eksportowanie wykresów 113, 119
pliki PostScript (encapsulated)
eksportowanie wykresów 113, 119
Pliki statystyk
kodowanie 267
pliki TIFF
eksportowanie wykresów 113, 119
pliki w formacie SAS
kodowanie 267
pliki wynikowe
zapisywanie 62
PMML
eksportowanie modeli 224
importowanie modeli 224
opcje eksportu 239
pobieranie obiektów z repozytorium IBM SPSS Collaboration and Deployment Services Repository 210
podgląd
dane węzła 40
podmiot użytkujący
IBM SPSS Analytic Server 12
podpaleta
tworzenie 246
usuwanie z karty palety 246
wyświetlanie na karcie palety 246
podpowiedzi
dodawanie adnotacji do węzłów 60
podsumowujący wykres punktowy 95
podziały tabeli 134
pojedyncze uwierzytelnianie 10
pojedyncze uwierzytelnianie, repozytorium IBM SPSS Collaboration and Deployment Services Repository 203, 204
połączenia
klaster serwerów 11
z repozytorium IBM SPSS Collaboration and Deployment Services Repository 204
z serwerem IBM SPSS Analytic Server 12
z serwerem IBM SPSS Modeler Server 10, 11
połączenia łańcuchów 169
porady
dotyczące ułatwień dostępu 265
ogólne zastosowanie 65
porządek wyświetlania 124
PowerPoint
eksportowanie wyników w formacie PowerPoint 116
PowerPoint, pliki 228
powiadomienia
określanie opcji 237
powiększanie 18
powitalne okno dialogowe 238
Production Facility
używanie komend z pliku dziennika 136
programowanie wizualne 14
projekty
adnotacje 230
dodawanie obiektów 228
generowanie raportów 231
tworzenie 228
tworzenie nowego 228
ustawianie folderu domyślnego 227
ustawianie właściwości 230
projekty *(kontynuacja)*
w repozytorium IBM SPSS Collaboration and Deployment Services 229
widok CRISP-DM 227
widok Klasy 228
właściwości folderu 230
właściwości obiektu 231
zamykanie 231
zapisywanie w repozytorium IBM SPSS Collaboration and Deployment Services Repository 208
przecinek 42
Przeglądarka
odstęp między elementami raportu 121
okno struktury 107
okno wyników 107
opcje wyświetlania 136
rozwijanie widoku struktury 109
struktura 108
ukrywanie wyników 107
usuwanie wyników 108
wyszukiwanie i zastępowanie informacji 110
wyświetlanie etykiet wartości 137
wyświetlanie etykiet zmiennych 137
wyświetlanie nazw zmiennych 137
wyświetlanie wartości danych 137
zapisywanie dokumentu 122
zmiana poziomu struktury 109
zmiana rozmiarów struktury 109
zmienianie czcionki struktury 109
znajdowanie i zastępowanie informacji 110
zwijanie widoku struktury 109
przenoszenie wierszy i kolumn 124
przewijanie
określanie opcji 46
przykłady
Applications — podręcznik 3
przegląd 4
przykłady aplikacji 3
przypisy
przenumerowywanie 132
symbole 129

R

radiany
jednostki miar 44
Raport aktywny Cognos 114
raport WWW
eksportowanie wyników 114
raporty
dodawanie do projektów 228
generowanie 231
ustawianie właściwości 231
zapisywanie wyników 62
regresja 265
regresja liniowa
eksportowanie jako PMML 239
regresja logistyczna
eksportowanie jako PMML 239
rekordy
braki danych 140
systemowe braki danych 141
rozkład chi-kwadrat
funkcje prawdopodobieństwa 176

rozkład f
funkcje prawdopodobieństwa 176
rozkład normalny
funkcje prawdopodobieństwa 176
rozkład t
funkcje prawdopodobieństwa 176
rozmiary
w strukturze 109

S

separator dziesiętny
formaty wyświetlania liczb 42
serwer
dodawanie połączeń 11
katalog domyślny 236
logowanie 10
wyszukiwanie serwerów w usłudze COP 11
skalowanie
tabele przestawne 128, 130
skalowanie strumieni do widoku 20
skróty
klawiatura 21, 256, 257, 260–262
ogólne zastosowanie 65
skróty klawiaturowe 21
słowa kluczowe
dodawanie adnotacji do węzłów 60
sortowanie
wiersze tabeli przestawnej 125
spacje
usuwanie z łańcuchów 149, 179
sprawdzanie wyrażeń CLEM 157
stany
ładowanie 63
zapisywanie 62
starsze wersje tabel 135
stopki 120
stopnie
jednostki miar 44
struktura
rozwijanie 109
w oknie raportu 108
zmiana poziomu 109
zwijanie 109
strumienie
adnotacje 56, 60
dodawanie do projektów 228
dodawanie komentarzy 56
dodawanie węzłów 36, 38
geoprzestrzenny układ współrzędnych 48
ładowanie 63
łączenie węzłów 36
opcje 41–44, 46, 47
opcje wdrożenia 216
pliki kopii zapasowej 61
pomijanie węzłów 37
skalowanie do widoku 20
tworzenie 35
uruchamianie 55
wyłączanie węzłów 37
wyświetlanie czasu wykonywania 49
zapisywanie 61
zapisywanie na serwerze IBM Cloud Pak for Data 221
strumienie *(kontynuacja)*
zapisywanie w repozytorium IBM SPSS Collaboration and Deployment Services Repository 207
zmiana nazwy 53, 60
strumienie danych
tworzenie 35
strumień 14
symbol grupowania
formaty wyświetlania liczb 42
system
opcje 235
system pomiarowy 136
systemowe braki danych
w rekordach 141
szablony
wykresy 104
Szablony TableLook
stosowanie 128
tworzenie 128
szerokie tabele
wklejanie do programu Microsoft Word 111
szerokość kolumny
określanie domyślnej szerokości 137
określanie maksymalnej szerokości 128
określanie szerokości dla tekstu zawijanego 128
tabele przestawne 133
szybkie tabele przestawne 137

Ś

środkowy przycisk myszy
emulowanie 21, 36

T

tabele
czcionki 131
dodawanie do projektów 228
kolor tła 131
kontrola podziałów tabeli 134
marginesy 131
właściwości komórki 131
wyrównanie 131
zapisywanie wyników 62
tabele przestawne
cofanie zmian 126
czcionki 131
domyślny szablon dla nowych tabel 137
dostosowanie domyślnej szerokości kolumn 137
dostosowywanie do rozmiaru strony 128, 130
drukowanie dużych tabel 134
drukowanie warstw 120
edycja 123
eksportowanie w formacie HTML 113
etykiety wartości 125
etykiety zmiennych 125
formaty komórek 129
grupowanie wierszy i kolumn 124
język 126
kolor tła 131
kontrola podziałów tabeli 134
linie siatki 134
manipulowanie 123

tabele przestawne *(kontynuacja)*
marginesy 131
nagłówki 131, 132
obracanie etykiet 124
obramowania 130
przenoszenie wierszy i kolumn 124
przestawianie 123
przypisy 131–133
rozdzielanie grup wierszy lub kolumn 124
sortowanie wierszy 125
starsze wersje tabel 135
szerokość komórek 133
szybkie tabele przestawne 137
szybsze wyświetlanie tabel 137
tekst kontynuacji 130
transponowanie wierszy i kolumn 124
tworzenie wykresów na podstawie tabel 135
ukrywanie 107
usuwanie etykiet grup 124
używanie ikon 123
warstwy 126
wklejanie do innych aplikacji 111
wklejanie jako tabel 111
właściwości 128
właściwości komórki 131
właściwości ogólne 128
właściwości przypisów 129
wstawianie nagłówków grup 124
wstawianie wierszy i kolumn 125
wybór ilości wierszy do wyświetlenia 129
wyrównanie 131
wyświetlanie i ukrywanie komórek 127
wyświetlanie ukrytych krawędzi 134
zaznaczanie wierszy i kolumn 134
zmiana porządku wyświetlania 124
zmiana szablonu 127
zmiany kolorów wierszy 129
tekst
dodawanie pliku tekstowego do Edytora raportów 110
dodawanie w Edytorze raportów 110
eksportowanie wyników jako tekstu 113, 118
tekst kontynuacji
w tabelach przestawnych 130
transponowanie wierszy i kolumn 124
tryb audytu danych
wykorzystanie w eksploracji 25
tworzenie niestandardowych palet
tworzenie podpalety 246
tworzenie skryptów 23, 145
typ wdrożenia 216
typowe zastosowania 25
typy danych
w parametrach 51
tytuły
dodawanie w Edytorze raportów 110

U

uczenie maszynowe 25
ukośnik odwrotny w wyrażeniach CLEM 162
ukrywanie
etykiety wymiarów 127
nagłówki 132
przypisy 132
ukrywanie *(kontynuacja)*
tytuły 127
wierszy i kolumn 127
wyników działania procedury 108
uruchamianie strumieni 55
usługa map 85
ustawienia regionalne
opcje 235
ustawienia strony
nagłówki i stopki 120
rozmiar wykresu 121
usuwanie blokady obiektów repozytorium IBM SPSS Collaboration and Deployment Services Repository 212
usuwanie etykiet grup 124
usuwanie wyników 108

W

warstwowe
funkcje przestrzenne 176
warstwy
drukowanie 120, 128, 130
tworzenie 126
w tabelach przestawnych 126
wyświetlanie 126, 127
wartości
dodawanie do wyrażeń CLEM 157
wyświetlanie z poziomu audytu danych 157
wartości data/czas 150
wartości globalne
w wyrażeniach CLEM 157
wartości null 150
wartości puste 139, 150
warunki 148
wdrażanie 204
węzeł Agregacja
wydajność 251
węzeł Audyt danych
wykorzystanie do eksploracji danych 26
węzeł ewaluacji
wydajność 251
węzeł K-średnie
duże zestawy 42
wydajność 252
węzeł kategoryzacji
wydajność 251
węzeł Kohonena
duże zestawy 42
wydajność 252
węzeł Łączenie
wydajność 251
węzeł pliku pamięci podręcznej
ładowanie 63
węzeł Powtórzenia
wydajność 251
węzeł sieci neuronowej
duże zestawy 42
wydajność 252
Węzeł Sortowanie
wydajność 251
węzeł tworzenia reguły
ładowanie 63
węzeł typu
braki danych 143

węzeł typu *(kontynuacja)*
wydajność 251
węzeł wypełniania
braki danych 143
węzły
adnotacje 56, 60
blokowanie 41
czasy wykonania 49
dane, podgląd 40
dodawanie 36, 38
dodawanie do projektów 228
dodawanie komentarzy do 56
dostosowywanie karty palety 247
duplikowanie 38
edycja 38
kolejność 249
ładowanie 63
łączenie w strumieniu 36
określanie opcji 38
podgląd danych 40
pomijanie w strumieniu 37
tworzenie niestandardowych palet 245
tworzenie podpalety niestandardowej 246
uaktywnianie 37
usuwanie 36
usuwanie połączeń 38
usuwanie z palety 246
wprowadzenie 35
wydajność 251, 252
wyłączanie 37, 38
wyłączanie w strumieniu 37
wyszukiwanie 52
wyświetlanie na palecie 246
zapisywanie 61
zapisywanie w repozytorium IBM SPSS Collaboration and Deployment Services Repository 208
węzły końcowe 35
węzły modelowania
dostosowywanie karty palety modelowania 247
wydajność 252
węzły procesowe
wydajność 251
węzły wyników 35
węzły źródłowe
mapowanie danych 64
odświeżanie 42
wiele otwartych plików z danymi 136
wiele sesji programu IBM SPSS Modeler 13
wielokąty częstości 82
wiersz komend
uruchamianie programu IBM SPSS Modeler 9
wiersze
zaznaczanie w tabelach przestawnych 134
within
funkcje przestrzenne 176
wklejanie 18
wklejanie wyniku do innych aplikacji 111
właściwości
dla strumieni danych 41
fazy raportu 231
folder projektu 230
tabele 128
tabele przestawne 128
właściwości komórki 131
właściwości obiektu, repozytorium IBM SPSS Collaboration and Deployment Services Repository 214
właściwości strumienia
Analytic Server 47
włączanie węzłów 37
wprowadzenie 161
wskazówki
ogólne zastosowanie 65
współrzędne geoprzestrzenne
format wyświetlania 47
wybór układów 48
wstawianie nagłówków grup 124
wstęp
IBM SPSS Modeler 9, 235
wybór geoprzestrzennego układu współrzędnych 48
wybór ilości wierszy do wyświetlenia 128
wybór palety węzłów 246
wycinanie 18
wydajność
węzłów modelowania 252
węzłów procesowych 251
wyrażenia CLEM 252
wygenerowane palety modeli 16
wykres K-K 92
wykres tabeli 135
wykresy
3-W 69
bąbelkowy 73
chmura słów 103
dodawanie do projektów 228
Drzewo 101
eksportowanie 113
Ewaluacja 78
histogram 82
kokpit 104
kolumna 70
kołowy 90
krzywa matematyczna 87
linia 83
linia rzutowania 95
mapa 84, 85
Mapa drzewa 102
mapa natężeń 80
pierścieniowy 98
piramida populacyjna 82, 92
podsumowujący wykres punktowy 95
podwójne osie Y 76
radarowy 94
relacja 95
równoległy 90
rzeka tematyczna 100
słupek 70
słupek błędu 70, 77
szablony 104
świeca 74
t-SNE 103
tworzenie na podstawie tabel przestawnych 135
ukrywanie 107
upakowane koła 75
użytkownika 76
wiele osi 76
wiele szeregów 89
wielokąt częstości 82
wykres K-K 92

wykresy *(kontynuacja)*
wykres punktowy 95
wykres rozrzutu 95
wykres skrzynkowy 73
wykres złożony 88
wykresy sekwencyjne 99
zapisywanie wyników 62
wykresy 3-W 69
wykresy bąbelkowe 73
wykresy chmury słów 103
wykresy dla wielu szeregów 89
wykresy drzewa 101
Wykresy ewaluacyjne 78
wykresy kolumnowe 70
wykresy kołowe 90
wykresy krzywych matematycznych 87
wykresy linii rzutowania 95, 97
wykresy liniowe
linia rzutowania 95, 97
wykresy sekwencyjne 99
wykresy map 84
wykresy map natężeń 80
wykresy niestandardowe 76
wykresy pierścieniowe 98
wykresy radarowe 94
wykresy relacji 95
wykresy rozrzutu
1-W 95
3-W 95
macierz 95, 97
nakładane 95
prosty 95
wykresy punktowe 95
zgrupowane 95
wykresy rozrzutu 3-W 95
wykresy równoległe 90
wykresy Rzeka tematyczna 100
wykresy sekwencyjne 99
wykresy skrzynkowe 73
wykresy słupkowe 70
wykresy słupkowe trójwymiarowe 70
wykresy słupków błędu 70, 77
wykresy świecowe 74
Wykresy t-SNE 103
wykresy Upakowane koła 75
wykresy wieloliniowe 83
wykresy z dwiema osiami 76
wykresy z dwiema osiami Y 76
wykresy z wieloma osiami 76
wykresy złożone 88
wykrywanie wiedzy 25
wyłączanie węzłów 37, 38
wyniki
eksportowanie 113
interaktywne 112
kopiowanie 108
Przeglądarka 107
przenoszenie 108
szyfrowanie 122
ukrywanie 107
usuwanie 108
wklejanie do innych aplikacji 111
wyrównanie 108, 136
wyśrodkowanie 108, 136
wyniki *(kontynuacja)*
wyświetlanie 107
zapisywanie 122
zmiana języka wyników 126
wyniki HTML
lektor ekranowy 265
wyniki interaktywne 112
wyrażenia 161
wyrażenia CLEM
wydajność 252
wyrównanie
wyniki 108, 136
wyszukiwanie
węzłów w strumieniu 52
wyszukiwanie i zastępowanie
dokumenty programu Viewer 110
wyszukiwanie obiektów w repozytorium IBM SPSS Collaboration and Deployment Services Repository 211
wyszukiwanie połączeń w usłudze COP 11
wyśrodkowanie wyników 108, 136
wyświetlanie
etykiety wymiarów 127
nagłówki 132
przypisy 132
tytuły 127
wiersze lub kolumny 127
wyniki 107
wyświetlanie wszystkich komentarzy w strumieniu 59

Z

za mało pamięci 235
zapisywanie
obiekty wielokrotne 62
obiekty wynikowe 62
stany 62
strumienie 61
węzły 61
zapisywanie obiektów w repozytorium IBM SPSS Collaboration and Deployment Services Repository 205
zapisywanie wykresów
pliki BMP 113, 119
pliki EMF 113
pliki EPS 113, 119
pliki JPEG 113, 119
pliki metafile 113
pliki PICT 113
pliki PNG 119
pliki PostScript 119
pliki TIFF 119
zapisywanie wyników
format Excel 113, 116
format HTML 113
format PDF 113, 117
format PowerPoint 113, 116
format tekstowy 113, 118
format Word 113, 115
HTML 113, 114
raport WWW 114
zarządzanie 16
zastępowanie modeli 237
zastosowania 25
zatrzymaj wykonywanie 18
zawijanie

zawijanie *(kontynuacja)*
określanie szerokości kolumny dla tekstu zawijanego 128
zestawione wykresy słupkowe 70
zestawy 42
zestawy reguł
ocenianie 42
zestawy wielokrotnych dychotomii
w wyrażeniach CLEM 152
zestawy wielokrotnych odpowiedzi
w wyrażeniach CLEM 152, 157
zestawy wielu kategorii
w wyrażeniach CLEM 152
zgodność z sekcją 508 255
zgrupowane wykresy liniowe 83
zgrupowane wykresy rozrzutu 95
zgrupowane wykresy słupkowe 70
zmiana nazwy
strumienie 53
węzły 60
zmiana porządku wierszy i kolumn 124
zmiany kolorów wierszy
tabele przestawne 129
zmienianie rozmiaru 19
zmienne
porządek wyświetlania w oknach dialogowych 136
w wyrażeniach CLEM 157
wyświetlanie wartości 157
zmienne czasu
przekształcanie 191
zmienne obowiązkowe 65
zmienne szablonów 65
znajdowanie i zastępowanie
dokumenty programu Viewer 110
znaki 161, 162
znaki specjalne
usuwanie z łańcuchów 149