IBM SPSS Custom Tables 27

Uwaga

Przed skorzystaniem z niniejszych informacji oraz produktu, którego one dotyczą, należy zapoznać się z informacjami zamieszczonymi w sekcji “Informacje” na stronie 87.

Informacje o produkcie

Niniejsze wydanie publikacji dotyczy wersji V27, wydania 0, modyfikacji 0 produktu IBM® SPSS Statistics oraz wszystkich następnych wersji i modyfikacji do czasu, aż w kolejnym wydaniu publikacji zostanie zawarta informacja o stosownej zmianie.

Spis treści

Rozdział 1. Tabele użytkownika..... 1
Interfejs konstruktora tabel.....1
Interfejs konstruktora tabel.....1
Tworzenie tabel.....1
Tabele użytkownika: karta Opcje.....15
Tabele użytkownika: karta Tytuły.....17
Tabele użytkownika: karta Testy.....17
Proste tabele dla zmiennych kategorialnych.....19
Proste tabele dla zmiennych kategorialnych.....19
Pojedyncza zmienna kategorialna.....19
Tabela krzyżowa.....21
Sortowanie i wykluczanie kategorii.....23
Zestawianie, zagnieżdżanie oraz warstwy w przypadku zmiennych kategorialnych.....24
Zestawianie zmiennych kategorialnych.....25
Zagnieżdżanie zmiennych kategorialnych.....26
Warstwy.....28
Podsumowania i sumy pośrednie dla zmiennych kategorialnych.....30
Proste podsumowanie dla jednej zmiennej.....30
Sumowane są tylko wyświetlane kategorie.....31
Miejsce wyświetlania podsumowań.....32
Podsumowania dla tabel zagnieżdżonych.....32
Podsumowania dla zmiennych w warstwach.....33
Sumy pośrednie.....33
Przeliczone kategorie dla zmiennych kategorialnych.....35
Prosta przeliczona kategoria.....35
Ukrywanie kategorii w przeliczonej kategorii.....36
Sumy pośrednie odniesień w przeliczonej kategorii.....37
Używanie przeliczonych kategorii do wyświetlania niewyczerpujących się sum pośrednich.....38
Tabele zmiennych zawierających wspólne kategorie.....39
Tabela liczebności.....40
Tabela procentów.....41
Element sterujący Podsumowania i kategorie.....41
Zagnieżdżanie zmiennych w tabelach zawierających wspólne kategorie.....42
Statystyki podsumowujące.....43
Zmienna podsumowująca statystyki źródłowej.....43
Zmienne zestawione.....45
Statystyki podsumowania wybierane przez użytkownika dla zmiennych kategorialnych.....46
Podsumowywanie zmiennych ilościowych.....48
Podsumowywanie zmiennych ilościowych.....48
Zestawione zmienne ilościowe.....48
Wielokrotne statystyki podsumowujące.....48
Liczebność, N ważnych i Braki danych.....49
Różne podsumowania dla różnych zmiennych.....50
Podsumowania grupowe w kategoriach.....51
Przedziały ufności.....53
Testy.....53
Testy.....53
Testy niezależności (chi-kwadrat).....54
Porównywanie średnich kolumnowych.....56
Porównywanie proporcji kolumnowych.....59
Uwaga na temat wag i zestawów wielokrotnych odpowiedzi.....64

Zestawy wielokrotnych odpowiedzi..... 64
Liczebności, odpowiedzi, procenty i podsumowania..... 65
Stosowanie zestawów wielokrotnych odpowiedzi z innymi zmiennymi.....67
Zestawy wielokrotnych odpowiedzi a odpowiedzi powtórzone..... 68
Testowanie istotności z zestawami wielokrotnych odpowiedzi..... 69
Braki danych..... 71
Tabele bez braków danych..... 71
Uwzględnianie braków danych w tabelach..... 72
Formatowanie i dostosowywanie tabel.....73
Formatowanie i dostosowywanie tabel..... 73
Format wyświetlania statystyk podsumowujących.....74
Etykiety statystyk podsumowujących..... 75
Szerokość kolumny..... 75
Wartości wyświetlane w pustych komórkach..... 76
Pliki przykładowe..... 77

Informacje..... 87
Znaki towarowe.....88

Indeks..... 91

Rozdział 1. Tabele użytkownika

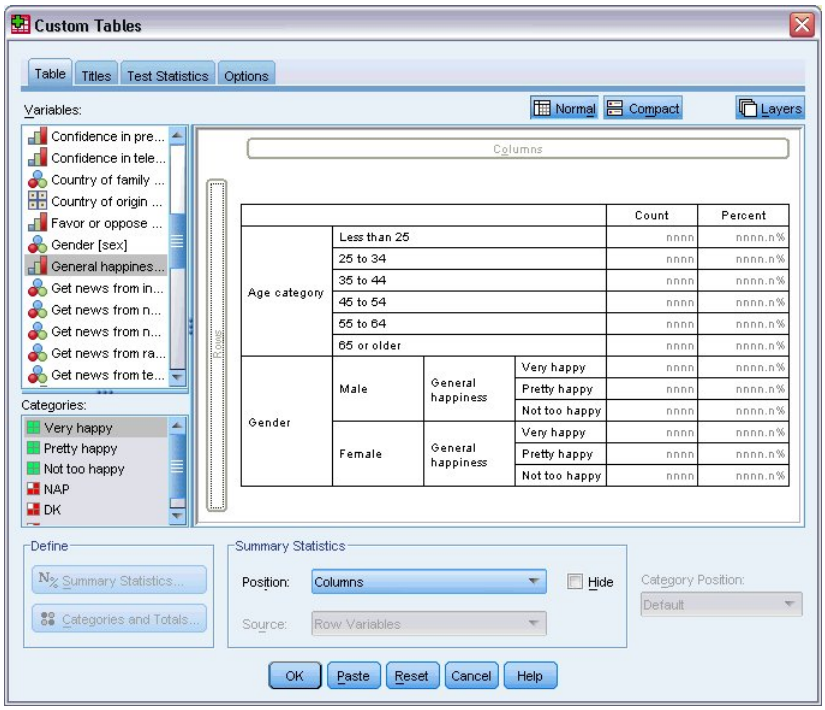
Następujące funkcje dotyczące tabel użytkownika są zwarte w SPSS Statistics Standard Edition lub komponencie opcjonalnym Custom Tables.

Interfejs konstruktora tabel

Interfejs konstruktora tabel

Moduł Tabele użytkownika wykorzystuje prosty interfejs konstruktora tabel, obsługujący funkcję „przeciągnij i upuść”, umożliwiający podgląd tabeli przy wybieraniu zmiennych i opcji. Udostępnia on również funkcje nie występujące w typowym „oknie dialogowym”, w tym możliwość zmiany rozmiaru okna i jego poszczególnych paneli.

Tworzenie tabel



Rysunek 1. Okno dialogowe Tabele użytkownika, karta Opcje

Zmienne i miary podsumowujące, które będą występować w tabelach, są wybierane na karcie Tabela w konstruktorze tabel.

Lista zmiennych. Zmienne zawarte w pliku danych są wyświetlane w lewym górnym panelu okna. Moduł Tabele użytkownika stosuje dwa różne poziomy pomiarów zmiennych i postępuje z nimi w sposób zależny od poziomu pomiaru danej zmiennej:

Jakościowe. Dane posiadające ograniczoną liczbę odrębnych wartości lub kategorii (np. płeć czy religia). Zmienne kategorialne mogą być zmiennymi łańcuchowymi (alfanumerycznymi) lub numerycznymi, wykorzystującymi kody liczbowe reprezentujące kategorie (np. 0 = *mężczyzna* i 1 = *kobieta*). Zmienne kategorialne nazywane są również danymi jakościowymi. Zmienne kategorialne mogą być **nominalne** lub **porządkowe**.

- *Nominalny*. Zmienna może być traktowana jako nominalna, gdy jej wartości reprezentują kategorie bez wewnętrznego rangowania; na przykład wydział, na którym są zatrudnieni pracownicy. Przykładami zmiennych nominalnych są: region, kod pocztowy lub wyznanie.
- *Porządkowy*. Zmienna może być traktowana jako porządkowa, gdy jej wartości reprezentują kategorię z wewnętrznym rangowaniem, na przykład poziomy zadowolenia z usługi od bardzo niezadowolonego do bardzo zadowolonego). Przykładami zmiennych porządkowych mogą być oceny opinii reprezentujące stopień satysfakcji lub przekonania oraz oceny preferencji.

Ilościowe. Dane mierzone na skali interwałowej lub ilorazowej, których wartości określają zarówno ich porządek, jak i odległość między nimi. Na przykład roczna pensja w wysokości 72 195 PLN jest wyżej niż pensja wynosząca 52 398 PLN, a odległość między tymi dwiema wartościami wynosi 19 797 PLN. Zmienne ilościowe są również zwane danymi ilościowymi lub ciągłymi.

Zmienne kategorialne definiują kategorie (w wierszach, kolumnach i warstwach) tabeli, a domyślną statystyką podsumowującą jest liczebność (liczba obserwacji w każdej kategorii). Na przykład domyślna tabela zmiennej jakościowej Płeć pokazuje po prostu liczbę kobiet i mężczyzn.

Zmienne ilościowe są zwykle podsumowywane w obrębie kategorii zmiennych kategorialnych, a domyślną statystyką podsumowującą jest średnia. Na przykład domyślna tabela zmiennej Dochód w obrębie kategorii płci pokazuje średni dochód kobiet i mężczyzn.

Podsumowania mogą również obejmować same zmienne ilościowe, bez ich grupowania według zmiennej kategorialnej. Jest to wykorzystywane głównie do **zestawiania** podsumowań wielu zmiennych ilościowych. Więcej informacji zawiera temat “Zestawianie zmiennych” na stronie 4.

Zestawy wielokrotnych odpowiedzi

Moduł Tabele użytkownika obsługuje również specjalny rodzaj „zmiennej” o nazwie **zestaw wielokrotnych odpowiedzi**. Zestawy wielokrotnych odpowiedzi nie są „zmiennymi” w normalnym tego słowa znaczeniu. Nie występują w Edytorze danych, a inne procedury nie rozpoznają ich. Zestawy wielokrotnych odpowiedzi wykorzystują wiele zmiennych do rejestracji odpowiedzi na pytania w przypadku, gdy respondent może udzielić więcej niż jednej odpowiedzi. Są one traktowane podobnie jak zmienne kategorialne i można je, w większości przypadków, poddawać podobnym operacjom. Więcej informacji zawiera temat “Zestawy wielokrotnych odpowiedzi” na stronie 64.

Ikona obok każdej zmiennej na liście zmiennych określa jej rodzaj.

Tabela 1. Ikony poziomu pomiaru

	Liczba	**Łańcuch**	**Data**	**Czas**
Zmienna (ilościowa)		n/a		
Porządkowa				
Nominalna				

Tabela 2. Ikony zestawów wielokrotnych odpowiedzi

Typ zestawu wielokrotnych odpowiedzi	**Ikona**
Zestaw wielokrotnych odpowiedzi, wielokrotne kategorie	
Zestaw wielokrotnych odpowiedzi, wielokrotne dychotomie	

Poziom pomiaru zmiennej w konstruktorze tabel można zmienić klikając ją prawym przyciskiem myszy na liście zmiennych i wybierając z menu podręcznego opcję **Jakościowa** lub **Ilościowa**. Poziom ten można zmienić na stałe w widoku Zmienne w Edytorze danych. Zmienne zdefiniowane jako **nominalne** lub **porządkowe** są traktowane przez moduł Tabele użytkownika jako zmienne jakościowe.

Kategorie. Po wybraniu z listy zmiennych zmiennej kategorialnej kategorie zdefiniowane dla niej są wyświetlane na liście Kategorie. Po zastosowaniu zmiennej w tabeli są one również wyświetlane w panelu obszaru roboczego. Jeśli dla zmiennej nie zdefiniowano żadnych kategorii, na liście Kategorie i w panelu obszaru roboczego wyświetlane są dwie kategorie zastępcze: *Kategoria 1* i *Kategoria 2*.

Zdefiniowane kategorie wyświetlane w konstruktorze tabel odpowiadają **etykietom wartości**, czyli opisowym etykietom przypisanym różnym wartościom danych (np. wartości liczbowe 0 i 1, z etykietami *mężczyzna* i *kobieta*). Etykiety wartości można zdefiniować w widoku Zmienne w Edytorze danych lub za pomocą opcji Definiuj zmienne w menu Dane w oknie Edytora danych.

Panel obszaru roboczego. Tabelę tworzy się przeciągając zmienne do wierszy i kolumn w panelu obszaru roboczego. Pokazywany jest tam podgląd tabeli, która zostanie utworzona. W panelu obszaru roboczego nie są pokazywane rzeczywiste wartości danych w poszczególnych komórkach, lecz dosyć wiernie przedstawia on układ finalnej tabeli. W przypadku zmiennych kategorialnych, jeśli plik danych zawiera unikalne wartości, dla których nie zdefiniowano etykiet, rzeczywista tabela może zawierać większą liczbę kategorii niż pokazywanych jest na podglądzie.

- W widoku **Normalnym** pokazywane są wszystkie wiersze i kolumny, które występować będą w tabeli, w tym wiersze i/lub kolumny przeznaczone dla statystyk podsumowujących i kategorii zmiennych kategorialnych.
- W widoku **Kompaktowym** pokazywane są tylko zmienne, które będą się znajdować w tabeli, bez podglądu wierszy i kolumn.

Podstawowe zasady i ograniczenia przy tworzeniu tabel

- W przypadku zmiennych kategorialnych statystyki podsumowujące są obliczane na podstawie najbardziej wewnętrznej zmiennej w wymiarze statystyki źródłowej.
- Domyślny wymiar statystyki źródłowej (wiersz lub kolumna) w przypadku zmiennych kategorialnych zależy od kolejności przeciągania zmiennych do panelu obszaru roboczego. Jeśli najpierw, na przykład zmienna zostanie przeciągnięta do pola wierszy, wymiar wierszowy będzie domyślnym wymiarem statystyki źródłowej.
- Zmienne ilościowe mogą być podsumowywane tylko w obrębie kategorii najbardziej wewnętrznej zmiennej w wymiarze wierszy lub kolumn. (Zmienną ilościową można umieścić na dowolnym poziomie tabeli, lecz jest ona podsumowywana na najbardziej wewnętrznym poziomie).
- Zmiennych ilościowych nie można podsumowywać w obrębie innych zmiennych ilościowych. Można jednak zestawiać podsumowania wielu zmiennych ilościowych lub podsumowywać je w obrębie kategorii zmiennych kategorialnych. Nie można zagnieżdżać jednej zmiennej ilościowej w innej ani umieszczać jednej zmiennej ilościowej w wymiarze wierszowym, a innej w wymiarze kolumnowym.
- Jeśli dowolna zmienna w aktywnym zbiorze danych zawiera więcej niż 12 000 zdefiniowanych etykiet wartości, nie można użyć kreatora tabel do tworzenia tabel. Jeśli zmienne przekraczające to ograniczenie nie muszą być uwzględniane w tabelach, można zdefiniować i zastosować zbiory danych, które wyłączają te zmienne. Jeśli konieczne jest uwzględnienie zmiennych zawierających ponad 12 000 zdefiniowanych etykiet wartości, do utworzenia tabel można użyć składni komend CTABLES.

Tworzenie tabeli

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. Przeciągnij jedną lub kilka zmiennych do pola wierszy i/lub kolumn w panelu obszaru roboczego.
3. Kliknij przycisk **OK**, aby utworzyć tabelę.
4. Zaznacz (kliknij) zmienną w panelu obszaru roboczego.
5. Przeciągnij zmienną w dowolne miejsce poza panelem obszaru roboczego lub naciśnij klawisz Delete.

Aby zmienić poziom pomiaru zmiennej:

6. Kliknij zmienną prawym przyciskiem myszy na liście zmiennych (można tego dokonać tylko na liście zmiennych, a nie w obszarze roboczym).
7. Z menu kontekstowego wybierz opcję **Jakościowa** lub **Ilościowa**.

Zestawianie zmiennych

Zestawianie polega na łączeniu oddzielnych tabel w celu jednoczesnego wyświetlenia informacji w nich zawartych. Można na przykład wyświetlić informacje na temat *płci* i *kategorii wiekowych* w oddzielnych sekcjach jednej tabeli.

Zestawianie zmiennych

1. Na liście zmiennych zaznacz wszystkie zmienne do zestawienia, a następnie przeciągnij je razem do wierszy lub kolumn w panelu obszaru roboczego.

   lub

2. Przeciągnij zmienne pojedynczo, upuszczając każdą z nich nad lub pod zmiennymi występującymi w wierszach z prawej lub lewej strony zmiennych występujących w kolumnach.

Tabela 3. Zmienne kategorialne zestawione

Lista zmiennych	Kategorie	Statystyka podsumowująca
Zmienna 1	Kategoria 1	123
	Kategoria 2	456
Zmienna 2	Kategoria 1	123
	Kategoria 2	456
	Kategoria 3	789

Więcej informacji zawiera temat “Zestawianie zmiennych kategorialnych” na stronie 25.

Zagnieżdżanie zmiennych

Zagnieżdżanie, podobnie jak umieszczanie w tabelach krzyżowych, umożliwia pokazanie zależności pomiędzy dwiema zmiennymi kategorialnymi, z tą różnicą, że jedna zmienna jest zagnieżdżona w drugiej w tym samym wymiarze. Można na przykład zagnieździć zmienną *Płeć* w zmiennej *Kategoria wieku* w wymiarze wierszowym, pokazując liczbę kobiet i mężczyzn w każdej kategorii wiekowej.

Można również zagnieździć zmienną ilościową w zmiennej kategorialnej. Na przykład zagnieżdżenie zmiennej *Dochód* w zmiennej *Płeć* spowoduje wyświetlenie odrębnej średniej (lub mediany czy innej miary podsumowującej) wartości dochodów kobiet i mężczyzn.

Zagnieżdżanie zmiennych

1. Przeciągnij zmienną kategorialną do pola wierszy lub kolumn w panelu obszaru roboczego.
2. Przeciągnij zmienną kategorialną lub ilościową do lewej lub prawej strony zmiennej kategorialnej w wierszu albo nad lub pod zmienną kategorialną w kolumnie.

Tabela 4. Zagnieżdżone zmienne kategorialne

Zmienna 1	Zmienna 2	Statystyka podsumowująca
Kategoria 1	Kategoria 1	12
	Kategoria 2	34
	Kategoria 3	56
Kategoria 2	Kategoria 1	12
	Kategoria 2	34

Tabela 4. Zagnieżdżone zmienne kategorialne (kontynuacja)		
Zmienna 1	**Zmienna 2**	**Statystyka podsumowująca**
	Kategoria 3	56

Więcej informacji zawiera temat “Zagnieżdżanie zmiennych kategorialnych” na stronie 26.

Uwaga: Tabele użytkownika nie uwzględniają przetwarzania plików powstałych w wyniku podzielenia danych wg warstw. Aby uzyskać taki sam wynik, jak dla plików powstałych w wyniku podzielenia danych wg warstw, umieść zmienne podziału danych w warstwach na skrajnym zewnętrznym poziomie zagnieżdżenia w tabeli.

Warstwy

Warstwy umożliwiają dodanie do tabel wymiaru głębokości, przekształcając je w trójwymiarowe „kostki” danych. Warstwy przypominają zagnieżdżanie lub zestawianie; podstawową różnicą jest fakt, że jednocześnie widoczna jest kategoria tylko jednej warstwy. Na przykład wykorzystując *Kategoria wieku* jako zmienną w wierszu i *Płeć* jako zmienną w warstwie można utworzyć tabelę, w której informacje dotyczące kobiet i mężczyzn wyświetlane są w różnych warstwach.

Tworzenie warstw

1. Kliknij przycisk **Warstwy** na karcie Tabela w konstruktorze tabel, aby wyświetlić listę warstw.
2. Przeciągnij zmienne ilościowe lub kategorialne, które definiować będą warstwy, na listę warstw.

Na liście warstw nie mogą jednocześnie występować zmienne ilościowe i kategorialne. Wszystkie zmienne muszą być tego samego rodzaju. Zestawy wielokrotnych odpowiedzi są traktowane na liście warstw jako zmienne jakościowe. Zmienne ilościowe w warstwach są zawsze zestawione.

Jeśli w warstwach występuje wiele zmiennych jakościowych, warstwy takie można zestawić lub zagnieździć.

- Opcja **Oddzielne kategorie (zestawienie)** jest równoważna zestawianiu. Dla każdej kategorii każdej zmiennej w warstwie, zostanie wyświetlona oddzielna warstwa. Całkowita liczba warstw jest po prostu *sumą* liczby kategorii każdej zmiennej w warstwie. Na przykład jeśli istnieją trzy zmienne w warstwach, każda z trzema kategoriami, tabela zawierać będzie dziewięć warstw.
- Opcja **Kombinacje kategorii (zagnieżdżanie)** jest równoważna zagnieżdżaniu lub umieszczaniu warstw w tabeli krzyżowej. Całkowita liczba warstw jest *iloczynem* liczby kategorii każdej zmiennej w warstwie. Na przykład jeśli istnieją trzy zmienne w warstwach, każda z trzema kategoriami, tabela zawierać będzie 27 warstw.

Wyświetlanie lub ukrywanie nazw i/lub etykiet zmiennych

Dostępne są następujące opcje wyświetlania nazw i etykiet zmiennych:

- Pokaż tylko etykiety zmiennych. W przypadku zmiennych bez zdefiniowanych etykiet wyświetlana jest nazwa zmiennej. Jest to ustawienie domyślne.
- Pokaż tylko nazwy zmiennych.
- Pokaż nazwy i etykiety zmiennych.
- Nie pokazuj nazw lub etykiet zmiennych. Chociaż kolumna/wiersz zawierające etykietę lub nazwę zmiennej będą wyświetlane na podglądzie tabeli w panelu obszaru roboczego, w samej tabeli będą niewidoczne.

Wyświetlanie lub ukrywanie etykiet lub nazw zmiennych:

1. Kliknij prawym przyciskiem myszy zmienną na podglądzie tabeli w panelu obszaru roboczego.
2. Z menu podręcznego wybierz opcję **Pokaż etykietę zmiennej** lub **Pokaż nazwę zmiennej**, aby włączyć lub wyłączyć wyświetlanie etykiet lub nazw zmiennych. Znacznik obok wybranej opcji oznacza, że etykiety lub nazwy będą wyświetlane.

Statystyki podsumowujące

Okno dialogowe Statystyki podsumowujące umożliwia:

- Dodawanie i usuwanie statystyk podsumowujących do/z tabeli.
- Zmianę etykiet statystyk.
- Zmianę kolejności statystyk.
- Zmianę formatu statystyk, w tym liczby miejsc dziesiętnych.

Dostępne tutaj statystyki podsumowujące (i inne opcje) zależą od poziomu pomiaru zmiennej statystyki źródłowej, wyświetlanej w górnej części okna dialogowego. Źródło statystyk podsumowujących (zmienna, na podstawie której obliczane są statystyki podsumowujące) jest określone przez:

- **Poziom pomiaru**. Jeśli tabela (lub sekcja tabeli w przypadku tabeli zestawionej) zawiera zmienną ilościową, statystyki podsumowujące obliczane są na podstawie tej zmiennej.
- **Kolejność wybierania zmiennych**. Domyślny wymiar statystyki źródłowej (wiersz lub kolumna) w przypadku zmiennych kategorialnych zależy od kolejności przeciągania zmiennych do panelu obszaru roboczego. Jeśli najpierw, na przykład zmienna zostanie przeciągnięta do pola wierszy, wymiar wierszowy będzie domyślnym wymiarem statystyki źródłowej.
- **Zagnieżdżanie**. W przypadku zmiennych kategorialnych statystyki podsumowujące są obliczane na podstawie najbardziej wewnętrznej zmiennej w wymiarze statystyki źródłowej.

Tabela zestawiona może zawierać wiele zmiennych podsumowujących statystyki źródłowej (zarówno ilościowych, jak i kategorialnych), jednak każda sekcja tabeli zawiera tylko jedną taką zmienną.

Zmiana wymiaru podsumowującej statystyki źródłowej

1. Wybierz wymiar (wiersze, kolumny lub warstwy) z rozwijanej listy **Źródło** w grupie Statystyki podsumowujące na karcie Tabela.

Określanie statystyk podsumowujących wyświetlanych w tabeli

1. Zaznacz (kliknij) zmienną podsumowującą statystyki źródłowej w panelu obszaru roboczego na karcie Tabela.
2. W grupie Definiuj na karcie Tabela kliknij **Statystyki podsumowujące**.

   lub
3. Kliknij prawym przyciskiem myszy zmienną podsumowującą statystyki źródłowej w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
4. Zaznacz statystyki podsumowujące, które mają zostać dołączone do tabeli. (Zaznaczone statystyki można przenieść z listy Statystyki do listy Pokaż w tabeli klikając przycisk ze strzałką lub przeciągając je).
5. Kliknij strzałkę skierowaną w górę lub w dół, aby zmienić położenie zaznaczonej statystyki podsumowującej.
6. Z rozwijanej listy Format wybierz format wyświetlania zaznaczonej statystyki podsumowującej.
7. W komórce Dziesiętne dla zaznaczonej statystyki podsumowującej wprowadź liczbę wyświetlanych miejsc dziesiętnych.
8. Kliknij przycisk **Zastosuj do wybranej**, aby zastosować zaznaczoną statystykę podsumowującą do zmiennych zaznaczonych w panelu obszaru roboczego.
9. Kliknij przycisk **Zastosuj do wszystkich**, aby zastosować zaznaczoną statystykę podsumowującą do wszystkich zmiennych zestawionych tego samego rodzaju, występujących w panelu obszaru roboczego.

Uwaga: Działanie przycisku **Zastosuj do wszystkich** różni się od działania przycisku **Zastosuj do wybranej** tylko w odniesieniu do zmiennych zestawionych tego samego rodzaju, już umieszczonych w panelu obszaru roboczego. W obu przypadkach zaznaczone statystyki podsumowujące zostają automatycznie zastosowane do wszystkich dodatkowych zmiennych zestawionych tego samego rodzaju, dodawanych do tabeli.

Statystyki podsumowujące dla zmiennych kategorialnych

Do podstawowych statystyk dostępnych dla zmiennych kategorialnych należą liczebności i procenty. Dla podsumowań ogółem i sum pośrednich można również określić statystyki użytkownika. Takie statystyki użytkownika obejmują miary tendencji centralnej (np. średnia i mediana) oraz rozproszenia (np. odchylenie standardowe), które mogą być odpowiednie dla niektórych porządkowych zmiennych kategorialnych. Więcej informacji zawiera temat “Statystyki podsumowania wybierane przez użytkownika dla zmiennych kategorialnych” na stronie 10.

Liczebności. Liczba obserwacji we wszystkich komórkach tabeli lub liczba odpowiedzi w przypadku zestawów wielokrotnych odpowiedzi. Jeśli stosowane jest ważenie, ta wartość jest ważoną liczebnością.

- Jeśli stosowane jest ważenie, wartość jest ważoną liczebnością.
- Liczebność ważona jest taka sama zarówno przy globalnym ważeniu zbioru danych (**Dane** > **Ważenie obserwacji**), jak i przy ważeniu według zmiennej będącej podstawą rzeczywistą (określa się ją na karcie **Opcje** okna dialogowego **Tabele użytkownika**.

Nieważona liczebność. Nieważona liczba obserwacji we wszystkich komórkach tabeli. Różni się od liczebności tylko, jeśli obowiązuje ważenie.

Skorygowana liczebność. Skorygowana liczebność jest używana przy obliczeniach z ważeniem według podstawy rzeczywistej. Jeśli nie jest używana zmienna będąca rzeczywistą podstawą ważenia (określona na karcie Opcje), to skorygowana liczebność jest taka sama, jak nieskorygowana.

Procent w kolumnie. Procent w każdej kolumnie. Suma procentów w każdej kolumnie podtabeli (w przypadku procentów prostych) wynosi 100%. Procenty w kolumnie są zwykle przydatne tylko wówczas, kiedy w *wierszu* występuje zmienna jakościowa.

Procent w wierszu. Procent w każdym wierszu. Suma procentów w każdym wierszu podtabeli (w przypadku procentów prostych) wynosi 100%. Procenty w wierszu są zwykle przydatne tylko wówczas, kiedy w *kolumnie* występuje zmienna jakościowa.

Wierszowy procent w warstwie i kolumnowy procent w warstwie. Suma procentów w wierszu lub kolumnie (w przypadku procentów prostych) wynosi 100% dla wszystkich podtabel w tabeli zagnieżdżonej. Jeśli tabela zawiera warstwy, suma procentów w wierszu lub w kolumnie wynosi 100% dla wszystkich podtabel zagnieżdżonych w każdej warstwie.

Procent w warstwie. Procent w każdej warstwie. W przypadku procentów prostych suma odsetka komórek w aktualnie widocznej warstwie wynosi 100%. Jeśli w warstwach nie występują żadne zmienne, procent ten jest równoważny procentowi w tabeli.

Procent w tabeli. Procenty dla każdej komórki obliczane są na podstawie liczby obserwacji w całej tabeli. Wszystkie odsetki komórek obliczane są na podstawie tej samej całkowitej liczby obserwacji, a ich suma w całej tabeli wynosi 100% (w przypadku procentów prostych).

Procent w podtabeli. Procenty dla każdej komórki są obliczane na podstawie liczby obserwacji w podtabeli. Wszystkie odsetki komórek w podtabeli obliczane są na podstawie tej samej całkowitej liczby obserwacji, a ich suma w podtabeli wynosi 100%. W tabelach zagnieżdżonych zmienna poprzedzająca najwyższy poziom zagnieżdżenia definiuje podtabele. Na przykład w tabeli określającej *stan cywilny* w podziale na *płeć* w podziale na *kategorie wieku*, zmienna *Płeć* definiuje podtabele.

Przedziały ufności

- Dostępna jest górna i dolna granica ufności liczebności, wartości procentowych, średniej, mediany, centyli i sumy.
- Łańcuch "&[Poziom ufności]" w etykiecie tabeli zawiera poziom ufności.
- Dostępny jest błąd standardowy liczebności, wartości procentowych, średniej i sumy.
- Przedziały ufności i błędy standardowe nie są dostępne w przypadku zestawów wielokrotnych odpowiedzi.

Poziom l

Poziom ufności przedziałów ufności wyrażony jako wartość procentowa. Wartość musi być liczbą większą od 0 i mniejszą niż 100.

Zestawy wielokrotnych odpowiedzi

Zestawy wielokrotnych odpowiedzi mogą zawierać procenty wyliczone na podstawie obserwacji, odpowiedzi lub liczebności. Więcej informacji zawiera temat “Statystyki podsumowujące dla zestawów wielokrotnych odpowiedzi” na stronie 8.

Zestawione tabele

Przy obliczaniu procentów każda sekcja tabeli zdefiniowana przez zmienną zestawiającą jest traktowana jako odrębna tabela. Sumy wierszowych procentów w warstwie, kolumnowych procentów w warstwie i procentów w tabeli wynoszą 100% (w przypadku procentów prostych) w każdej zestawionej sekcji tabeli. Podstawa obliczania wartości procentowej wykorzystywana przy obliczaniu różnych procentów opiera się na liczbie obserwacji w każdej zestawionej sekcji tabeli.

Podstawa obliczania wartości procentowej

Procenty mogą być obliczane na trzy różne sposoby, w zależności od sposobu traktowania braków danych w bazie obliczeniowej:

Procent prosty. Procenty obliczane są na podstawie liczby obserwacji wykorzystanych w tabeli, a ich suma zawsze wynosi 100%. Jeśli kategoria zostanie wykluczona z tabeli, obserwacje do niej należące nie są uwzględniane w bazie do obliczeń. Obserwacje zawierające systemowe braki danych nigdy nie są uwzględniane w bazie do obliczeń. Obserwacje zawierające braki danych zdefiniowane przez użytkownika nie są uwzględniane, jeśli z tabeli zostaną wykluczone kategorie braków danych zdefiniowanych przez użytkownika (ustawienie domyślne), natomiast są uwzględniane, jeśli takie kategorie zostaną włączone do tabeli. Każdy procent niezawierający w nazwie terminu *N ważnych* lub *N Ogółem* jest procentem prostym.

Procent z N ogółem. Obserwacje zawierające systemowe i zdefiniowane przez użytkownika braki danych są uwzględniane przy obliczaniu procentów prostych. Suma procentów może być mniejsza niż 100%.

Procent z N ważnych. Obserwacje zawierające braki danych zdefiniowane przez użytkownika nie są uwzględniane przy obliczaniu procentów prostych nawet wówczas, gdy tabela zawiera kategorie braków danych zdefiniowanych przez użytkownika.

Uwaga: Obserwacje w kategoriach wykluczonych ręcznie, poza kategoriami braków danych zdefiniowanych przez użytkownika, nigdy nie są uwzględniane przy obliczeniach.

Statystyki podsumowujące dla zestawów wielokrotnych odpowiedzi

Dla zestawów wielokrotnych odpowiedzi dostępne są następujące statystyki dodatkowe.

% z odpowiedzi w kolumnie/wierszu/warstwie. Procent na podstawie odpowiedzi.

% z odpowiedzi w kolumnie/wierszu/warstwie (Baza: liczebność). Licznikiem jest liczba odpowiedzi, a mianownikiem liczebność ogółem.

% z liczebności w kolumnie/wierszu/warstwie (Baza: odpowiedzi). Licznikiem jest liczebność, a mianownikiem całkowita liczba odpowiedzi.

Kolumnowy/Wierszowy % z odpowiedzi w warstwie %. Procent w podtabelach. Procent na podstawie odpowiedzi.

Kolumnowy/Wierszowy % z odpowiedzi w warstwie (Baza: liczebność). Procenty w podtabelach. Licznikiem jest liczba odpowiedzi, a mianownikiem liczebność ogółem.

Kolumnowy/Wierszowy % z odpowiedzi w warstwie (Baza: odpowiedzi). Procenty w podtabelach. Licznikiem jest liczebność, a mianownikiem całkowita liczba odpowiedzi.

Odpowiedzi. Liczebność odpowiedzi.

% z odpowiedzi w podtabeli/tabeli. Procent na podstawie odpowiedzi.

% z odpowiedzi w podtabeli/tabeli (Baza: liczebność). Licznikiem jest liczba odpowiedzi, a mianownikiem liczebność ogółem.

% z liczebności w podtabeli/tabeli (Baza: odpowiedzi). Licznikiem jest liczebność, a mianownikiem całkowita liczba odpowiedzi.

Statystyki podsumowujące dla zmiennych ilościowych i podsumowania użytkownika dla zmiennych jakościowych

Oprócz liczebności i procentów dostępnych dla zmiennych kategorialnych, dla zmiennych ilościowych dostępne są następujące statystyki podsumowujące oraz jako statystyki podsumowania wybierane przez użytkownika i sumy pośrednie dla zmiennych kategorialnych. Statystyki takie nie są dostępne dla zestawów wielokrotnych odpowiedzi ani zmiennych łańcuchowych (alfanumerycznych).

Średnia. Średnia arytmetyczna; suma podzielona przez liczbę obserwacji.

Mediana. Wartość, powyżej i poniżej której przypada połowa obserwacji; 50. percentyl.

Dominanta. Wartość najczęściej występująca. Jeżeli występuje wiązanie, pokazywana jest wartość najmniejsza.

Minimum. Najmniejsza (najniższa) wartość.

Maksimum. Największa (najwyższa) wartość.

Braki danych. Liczba braków danych (zdefiniowanych i systemowych).

Percentyl. Można uwzględnić 5., 25., 75., 95. i/lub 99. percentyl.

Przedział. Różnica pomiędzy wartością maksymalną a minimalną.

Odchylenie standardowe. Miara rozproszenia wokół średniej. W przypadku rozkładu normalnego, 68% obserwacji znajduje się w obszarze oddalonym o jedno odchylenie standardowe od średniej, zaś 95% - w przedziale oddalonym o dwa odchylenia standardowe. Na przykład jeśli średni wiek wynosi 45 lat, przy odchyleniu standardowym równym 10, w przypadku rozkładu normalnego 95% obserwacji mieściłoby się w przedziale od 25 do 65 lat (pierwiastek kwadratowy z wariancji).

Suma. Suma wartości.

Procent sumy. Procenty na podstawie sum. Statystyka dostępna dla wierszy i kolumn (w podtabelach), całych wierszy i kolumn (dla wszystkich podtabel), warstw, podtabel i całych tabel.

Ogółem N. Liczba wartości bez braków danych, ze zdefiniowanymi brakami danych i systemowymi brakami danych. Nie obejmuje obserwacji w kategoriach wykluczonych ręcznie innych niż kategorie braków danych zdefiniowanych przez użytkownika.

Skorygowane - N Ogółem. Skorygowana suma N używana przy obliczeniach z ważeniem według podstawy rzeczywistej. Jeśli nie jest używana zmienna będąca rzeczywistą podstawą ważenia (określona na karcie Opcje), to skorygowana suma N liczebność jest taka sama, jak nieskorygowana. Ta statystyka jest niedostępna w przypadku zestawów wielokrotnych odpowiedzi.

N ważnych. Liczba wartości bez braków danych. Nie obejmuje obserwacji w kategoriach wykluczonych ręcznie innych niż kategorie braków danych zdefiniowanych przez użytkownika.

Skorygowane - N ważnych. Skorygowana liczba N ważnych jest stosowana w obliczeniach z ważeniem według podstawy rzeczywistej. Jeśli nie jest używana zmienna będąca rzeczywistą podstawą ważenia (określona na karcie Opcje), to skorygowana liczba N ważnych jest taka sama, jak nieskorygowana. Ta statystyka jest niedostępna w przypadku zestawów wielokrotnych odpowiedzi.

Wariancja. Miara rozproszenia wokół średniej, równa sumie podniesionych do kwadratu odchyleń od średniej, podzielonej przez liczbę obserwacji minus jeden. Wariancja mierzona jest w jednostkach będących kwadratem jednostek samej zmiennej (kwadrat odchylenia standardowego).

Przedziały ufności

- Dostępna jest górna i dolna granica ufności liczebności, wartości procentowych, średniej, mediany, centyli i sumy.
- Łańcuch "&[Poziom ufności]" w etykiecie tabeli zawiera poziom ufności.
- Dostępny jest błąd standardowy liczebności, wartości procentowych, średniej i sumy.
- Przedziały ufności i błędy standardowe nie są dostępne w przypadku zestawów wielokrotnych odpowiedzi.

Poziom l

Poziom ufności przedziałów ufności wyrażony jako wartość procentowa. Wartość musi być liczbą większą od 0 i mniejszą niż 100.

Zestawione tabele

Każda sekcja tabeli zdefiniowana przez zmienną zestawiającą jest traktowana jako odrębna tabela i w ten sposób obliczane są statystyki podsumowujące.

Statystyki podsumowania wybierane przez użytkownika dla zmiennych kategorialnych

W przypadku tabel ze zmiennymi kategorialnymi, zawierających podsumowania ogółem lub sumy pośrednie, można stosować inne statystyki podsumowujące niż podsumowania wyświetlane dla każdej kategorii. Można na przykład wyświetlić liczebności i procenty w kolumnach dla porządkowej zmiennej kategorialnej w wierszu i wyświetlić medianę dla statystyki „suma”.

Aby utworzyć tabelę dla zmiennej kategorialnej ze statystyką podsumowania wybieraną przez użytkownika:

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**

   Otwarty zostanie kreator tabel.
2. Przeciągnij zmienną kategorialną do pola wierszy lub kolumn w obszarze roboczym.
3. Kliknij prawym przyciskiem myszy zmienną w obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
4. Kliknij (zaznacz) pole wyboru **Suma**, a następnie kliknij przycisk **Zastosuj**.
5. Ponownie kliknij zmienną prawym przyciskiem myszy w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
6. Kliknij (zaznacz) opcję **Statystyki użytkownika dla podsumowań ogółem i sum pośrednich**.

   Domyślnie wszystkie statystyki podsumowujące, w tym podsumowania wybierane przez użytkownika, są wyświetlane w wymiarze przeciwnym do wymiaru zawierającego zmienną kategorialną. Na przykład jeśli w wierszu występuje zmienna jakościowa, statystyki podsumowujące definiują kolumny tabeli, jak pokazano na poniższym rysunku:

Tabela 5. Kategorie zmiennych porządkowych w wierszach, statystyki podsumowujące liczebności i średnie w kolumnach

Lista zmiennych	Kategorie	Liczebność	Średnia
Zmienna 1	1 Zgadzam się	196	
	2 Opinia neutralna	936	
	3 Nie zgadzam się	744	
	Podsumowanie ogółem	1876	2,29

Aby wyświetlić statystyki podsumowujące w tym samym wymiarze, co zmienna kategorialna:

7. Na karcie Tabela w konstruktorze tabel, w grupie Statystyki podsumowujące z rozwijanej listy Pozycja wybierz żądany wymiar.

Jeśli, na przykład zmienna kategorialna jest wyświetlana w wierszach, z rozwijanej listy wybierz **Wiersze**.

Formaty wyświetlania statystyk podsumowujących

Dostępne są następujące opcje formatu wyświetlania:

nnnn. Prosta wartość liczbowa.

nnnn%. Symbol procentów dołączony na końcu wartości.

Auto. Zdefiniowany format wyświetlania zmiennych, wraz z określeniem liczby miejsc dziesiętnych.

N=nnnn. Przed wartością wyświetlane jest *N*=. Opcja ta może być przydatna w przypadku statystyk takich, jak liczebności, *N* ważnych i Ogółem *N* w tabelach, w których etykiety statystyk podsumowujących nie są wyświetlane.

(nnnn). Wszystkie wartości ujęte w nawiasy.

(nnnn)(wart. ujemna). Tylko ujemne wartości ujęte w nawiasy.

(nnnn%). Wszystkie wartości ujęte w nawiasy, z dołączonym na końcu symbolem procentów.

n,nnn.n. Format z przecinkiem. Przecinek jako symbol grupowania cyfr i kropka jako separator miejsc dziesiętnych, bez względu na ustawienia regionalne.

n.nnn,n. Format z kropką. Kropka jako symbol grupowania cyfr i przecinek jako separator miejsc dziesiętnych, bez względu na ustawienia regionalne.

$n,nnn.n. Format dolarowy. Symbol dolara wyświetlany przed wartością; przecinek jako symbol grupowania cyfr i kropka jako symbol dziesiętny, bez względu na ustawienia regionalne.

CCA, CCB, CCC, CCD, CCE. Formaty użytkownika. Na liście wyświetlany jest bieżący format zdefiniowany dla każdej waluty użytkownika. Formaty takie są definiowane na karcie Użytkownika, w oknie dialogowym Opcje (menu Edycja, Opcje).

Ogólne zasady i ograniczenia

- Z wyjątkiem formatu Auto liczba miejsc dziesiętnych jest określona ustawieniem Dziesiętne dla kolumn.
- Z wyjątkiem formatów z przecinkiem, symbolem dolara i kropką, wykorzystywany jest separator miejsc dziesiętnych określony w Ustawieniach regionalnych na Panelu sterowania systemu Windows.
- Chociaż w przypadku formatu z przecinkiem/symbolem dolara i kropką jako symbol grupowania cyfr wyświetlany jest odpowiednio przecinek lub kropka, w momencie tworzenia programu nie istnieje format wyświetlania symbolu grupowania cyfr zgodnego z bieżącymi ustawieniami regionalnymi, określonymi w Panelu sterowania systemu Windows.

Kategorie i podsumowania

Okno dialogowe Kategorie i podsumowania umożliwia:

- Zmianę kolejności i wykluczanie kategorii.
- Wstawianie podsumowań i sum pośrednich.
- Wstaw przeliczone kategorie.
- Uwzględnianie lub wykluczanie pustych kategorii.
- Uwzględnianie lub wykluczanie kategorii zdefiniowanych jako zawierające braki danych.
- Uwzględnianie lub wykluczanie kategorii nie posiadających zdefiniowanych etykiet wartości.
- To okno dialogowe jest dostępne tylko dla zmiennych kategorialnych i zestawów wielokrotnych odpowiedzi. Nie jest dostępne dla zmiennych ilościowych.
- W przypadku zaznaczenia wielu zmiennych posiadających różne kategorie nie można wstawić sum pośrednich, wstawić przeliczonych kategorii, wykluczyć kategorii ani ręcznie zmienić ich kolejności. Ma to miejsce tylko wówczas, gdy na podglądzie w obszarze roboczym zaznaczono wiele zmiennych i

otwarto to okno dialogowe jednocześnie dla wszystkich zaznaczonych zmiennych. Każdą z takich czynności można jednak wykonać oddzielnie dla każdej zmiennej.

- W przypadku zmiennej bez zdefiniowanych etykiet wartości można tylko sortować kategorie i wstawiać podsumowania.

Uzyskiwanie dostępu do okna dialogowego Kategorie i podsumowania

1. Przeciągnij zmienną kategorialną lub zestaw wielokrotnych odpowiedzi do panelu obszaru roboczego.
2. Kliknij prawym przyciskiem myszy zmienną w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.

   lub

3. Zaznacz (kliknij) zmienną w panelu obszaru roboczego, a następnie kliknij **Kategorie i podsumowania** w grupie Definiuj na karcie Tabela.

   Można również zaznaczyć kilka zmiennych kategorialnych występujących w tym samym wymiarze w panelu obszaru roboczego:

4. Naciśnij klawisz Ctrl i kliknij każdą zmienną w panelu obszaru roboczego.

   lub

5. Kliknij poza polem podglądu tabeli w panelu obszaru roboczego, a następnie kliknij i przeciągnij wskaźnik myszy, aby zaznaczyć obszar zawierający zmienne do zaznaczenia.

   lub

6. Kliknij prawym przyciskiem myszy dowolną zmienną w wymiarze i wybierz opcję **Zaznacz zmienne w [wymiar]**, aby zaznaczyć wszystkie zmienne w danym wymiarze.

Aby zmienić porządek kategorii

Aby ręcznie zmienić kolejność kategorii:

1. Zaznacz (kliknij) kategorię na liście.
2. Kliknij strzałkę skierowaną w górę lub w dół, aby przenieść kategorię w górę lub w dół listy.

   lub

3. Kliknij kolumnę Wartości dla danej kategorii i przeciągnij ją w inne miejsce.

Aby wykluczyć kategorie

1. Zaznacz (kliknij) kategorię na liście.
2. Kliknij strzałkę obok listy Wyklucz kategorie.

   lub

3. Kliknij kolumnę Wartości dla danej kategorii i przeciągnij ją poza listę.

W przypadku wykluczenia jakichkolwiek kategorii, wykluczone zostaną również kategorie bez zdefiniowanych etykiet wartości.

Aby posortować kategorie

Kategorie można sortować według wartości danych, etykiet wartości, liczby komórek lub statystyki podsumowującej, w porządku rosnącym lub malejącym.

1. W grupie Sortuj kategorie kliknij listę rozwijaną Według i wybierz kryterium sortowania: wartość, etykieta, liczebność lub statystyka podsumowująca (np. średnia, mediana lub dominanta). Dostępność statystyk podsumowujących sortowanie zależy od statystyk podsumowujących, które zostały wybrane z przeznaczeniem do wyświetlania w tabeli.
2. Kliknij rozwijaną listę Porządek, aby wybrać porządek sortowania (rosnąco lub malejąco).

Sortowanie kategorii jest niedostępne, jeśli jakiekolwiek kategorie zostały wykluczone.

Sumy pośrednie

1. Zaznacz (kliknij) na liście ostatnią kategorię z zakresu kategorii, które mają zostać uwzględnione w sumie pośredniej.
2. Kliknij przycisk **Dodaj sumę pośrednią...**.
3. W oknie dialogowym Definiuj sumy pośrednie w razie konieczności możesz zmodyfikować tekst etykiety sumy pośredniej.
4. Aby wyświetlać tylko sumę pośrednią i ukryć kategorie, definiujące sumę pośrednią, wybierz polecenie **Ukrywanie kategorii z sumami pośrednimi**.
5. Kliknij przycisk **Dalej** aby dodać sumę pośrednią.

Podsumowania

1. Kliknij pole wyboru **Suma**. Można zmodyfikować również tekst etykiety sumy.

Jeśli zaznaczona zmienna jest zagnieżdżona w innej zmiennej, podsumowania zostaną wstawione dla każdej podtabeli.

Położenie podsumowań i sum pośrednich

Podsumowania i sumy pośrednie mogą być wyświetlane nad lub pod kategoriami uwzględnionymi w każdym podsumowaniu.

- Jeśli w grupie Umiejscowienie podsumowań i sum pośrednich wybrana zostanie opcja **Poniżej kategorii**, podsumowania wyświetlane będą pod każdą podtabelą, a wszystkie kategorie znajdujące się powyżej zaznaczonej kategorii z nią włącznie (lecz poniżej poprzedniej sumy pośredniej) zostaną uwzględnione w każdej sumie pośredniej.
- Jeśli w grupie Umiejscowienie podsumowań i sum pośrednich wybrana zostanie opcja **Powyżej kategorii**, podsumowania wyświetlane będą nad każdą podtabelą, a wszystkie kategorie znajdujące się poniżej zaznaczonej kategorii z nią włącznie (lecz powyżej poprzedniej sumy pośredniej) zostaną uwzględnione w każdej sumie pośredniej.

Ważne: Położenie sum pośrednich należy wybrać przed zdefiniowaniem jakiejkolwiek z nich. Zmiana położenia ma wpływ na wszystkie sumy pośrednie (a nie tylko aktualnie zaznaczone) i powoduje również *zmianę kategorii uwzględnionych w sumach pośrednich.*

Przeliczone kategorie

Możesz wyświetlić przeliczone kategorie statystyk podsumowujących sum całkowitych i pośrednich i/lub stałych. Więcej informacji zawiera temat “Przeliczone kategorie” na stronie 14.

Statystyki podsumowania wybierane przez użytkownika i sumy pośrednie

Za pomocą okna dialogowego Statystyki podsumowujące w obszarach tabeli zawierających podsumowania i sumy pośrednie można wyświetlić statystyki inne niż „podsumowania”. Więcej informacji zawiera temat “Statystyki podsumowujące dla zmiennych kategorialnych” na stronie 7.

Uwaga: Jeśli wybranych jest wiele podsumowujących statystyk użytkownika, które są również zawarte w części głównej tabeli, i zostaną ukryte etykiety statystyk, podsumowania zostają przywrócone w tym samym porządku co w części głównej tabeli. Ponieważ etykiety nie są wyświetlane, więc można nie mieć pewności, co dane statystyki reprezentują. Ogólnie nie jest wskazane wybieranie wielu statystyk i ukrywanie etykiet.

Podsumowania, sumy pośrednie i wykluczone kategorie

W obliczeniach podsumowań nie są uwzględniane obserwacje należące do wykluczonych kategorii.

Braki danych, puste kategorie i wartości bez etykiet

Brakujące wartości. Określa sposób wyświetlania **braków danych zdefiniowanych przez użytkownika**, czyli wartości zdefiniowanych jako zawierające braki danych (np. kod 99 oznaczający „nie dotyczy” w odniesieniu do ciąży u mężczyzn). Domyślnie braki danych zdefiniowane przez użytkownika są wykluczane. Zaznaczenie tej opcji spowoduje włączenie do tabeli kategorii braków danych zdefiniowanych przez użytkownika. Chociaż zmienna może zawierać więcej niż jedną kategorię braków danych, na podglądzie tabeli w obszarze roboczym wyświetlana jest tylko jedna ogólna kategoria braków danych. Wszystkie kategorie braków danych zdefiniowanych przez użytkownika zostaną włączone do

tabeli. **Systemowe braki danych** (puste komórki dla zmiennych numerycznych w Edytorze danych) są zawsze wykluczane.

Puste kategorie. Puste kategorie to kategorie ze zdefiniowanymi etykietami wartości, w których w danej tabeli lub podtabeli nie występują żadne obserwacje. Domyślnie puste kategorie są włączane do tabel. Usunięcie zaznaczenia tej opcji spowoduje wykluczenie ich z tabeli.

Inne wartości znalezione podczas skanowania danych. Domyślnie wartości kategorii w pliku danych, nieposiadające zdefiniowanych etykiet, są automatycznie włączane do tabel. Usunięcie zaznaczenia tej opcji spowoduje wykluczenie z tabeli wartości bez zdefiniowanych etykiet. W przypadku wykluczenia jakichkolwiek kategorii ze zdefiniowanymi etykietami wartości, wykluczone zostaną również kategorie bez zdefiniowanych etykiet wartości.

Przeliczone kategorie

Oprócz wyświetlania łącznych wyników podsumowania statystyk, tabela może zawierać jedną lub więcej kategorii przeliczonych z tych łącznych wyników, z wartości stałych, sum pośrednich i łącznych oraz ich kombinacji. Wyniki określane są jako kategorie przeliczone i po przeliczeniu. Kategoria przeliczona zachowuje się jak kategoria w pojedynczej zmiennej z następującymi podobieństwami i różnicami:

- Przeliczona kategoria jest pozycjonowana tak jak inne kategorie.
- Kategoria przeliczona działa na tej samej statystyce, co inne kategorie.
- Kategorie przeliczone nie wypływają na sumy pośrednie, całkowite lub testy znaczenia.
- Domyślnie wartości przeliczonych kategorii stosują to samo formatowanie dla statystyk podsumowania, co inne kategorie. Można wymusić format podczas definiowania przeliczonej kategorii.

Ponieważ przeliczone kategorie mogą być używane z całkowitymi połączonymi wynikami, mogą być podobne do sum pośrednich. Jednak przeliczone kategorie mają następujące zalety w porównaniu z sumami pośrednimi.

- Przeliczone kategorie mogą być obliczane na podstawie wyników innych sum pośrednich.
- Przeliczone kategorie mogą na siebie zachodzić, operując w tych samych (lub częściowo tych samych) kategoriach.
- Przeliczone kategorie nie muszą obejmować wartości ze wszystkich innych kategorii powyżej lub poniżej przeliczonej kategorii. Oznacza to, że przeliczone kategorie nie wyczerpują się.
- Przeliczone kategorie mogą zawierać wartości z kategorii, które nie są sąsiednie.

W odróżnieniu od sum całkowitych i pośrednich, przeliczone kategorie są obliczane z łączonych danych, zamiast danych oryginalnych. Dlatego wartości przeliczonych kategorii mogą nie zgadzać się z wynikami sum całkowitych i pośrednich. Ponadto, ponieważ masz opcję ukrywania kategorii źródłowych podczas definiowania przeliczonej kategorii, interpretowanie sum pośrednich w tabeli wyniku może być trudne. Jeśli używa się przeliczonych kategorii zaleca się określenie niestandardowych etykiet dla sum pośrednich.

Aby zdefiniować przeliczoną kategorię

Przeliczone kategorie są dodawane z okna dialogowego Kategorie i podsumowania. Informacje o dostępie do tego okna dialogowego opisano w części “Kategorie i podsumowania” na stronie 11 .

1. W oknie dialogowym Kategorie i podsumowania kliknij **Dodaj kategorię...**
2. W opcji **Etykieta dla przeliczonej kategorii** określ etykietę dla przeliczonej kategorii. Możesz przeciągnąć kategorie z listy Kategorie, aby zawrzeć etykiety dla tych kategorii.
3. Zbuduj wyrażenie wybierając kategorie i/lub sumy oraz sumy pośrednie przy pomocy operatorów w celu zdefiniowania przeliczonych kategorii. Można również wpisać wartości stałe (np. 500) w celu ich dołączenia do wyrażenia.
4. Aby wyświetlać tylko przeliczoną kategorię i ukryć kategorie definiujące sumę pośrednią, wybierz polecenie **Ukrywanie kategorii używanych w wyrażeniach**.

5. Kliknij kartę **Wyświetl formaty** w celu zmiany formatu wyświetlania i liczby miejsc dziesiętnych dla przeliczonej kategorii. Więcej informacji zawiera temat “Formaty wyświetlania dla przeliczonych kategorii” na stronie 15.
6. Kliknij przycisk **Dalej**, aby dodać przeliczoną kategorię.

Formaty wyświetlania dla przeliczonych kategorii

Domyślnie przeliczona kategoria używa tego samego formatu wyświetlania i liczby miejsc po przecinku, podobnie jak inne kategorie w zmiennej. Można je nadpisać w tabeli Formaty wyświetlania w oknie dialogowym Przeliczona kategoria. Karta Format wyświetlania zestawia bieżącą statystykę podsumowania, w której funkcjonuje przeliczona kategoria jako uzupełnienie do wyświetlania formatów i liczby miejsc dziesiętnych dla tej statystyki.

Dla każdej statystyki podsumowania możliwe jest wykonanie następujących czynności:

1. Z rozwijanej listy Format wybierz format wyświetlania statystyki podsumowania. Pełna lista dostępnych formatów znajduje się w temacie “Formaty wyświetlania statystyk podsumowujących” na stronie 11 .
2. W komórce Dziesiętne dla zaznaczonej statystyki podsumowującej wprowadź liczbę wyświetlanych miejsc dziesiętnych.

Tabele zmiennych zawierających wspólne kategorie (tabele porównawcze)

Ankiety często zawierają wiele pytań ze wspólnym zestawem możliwych odpowiedzi. Wykorzystując zestawianie zmiennych można wyświetlić tak powiązane zmienne w jednej tabeli, a w jej kolumnach wyświetlić wspólne kategorie odpowiedzi.

Tworzenie tabel dla wielu zmiennych zawierających wspólne kategorie

Tabela 6. Zmienne zestawione zawierające wspólne kategorie odpowiedzi w kolumnach

Lista zmiennych	**Kategoria 1**	**Kategoria 2**	**Kategoria 3**
Zmienna 1	12	34	56
Zmienna 2	56	12	34
Zmienna 3	34	56	12

Więcej informacji zawiera temat “Tabele zmiennych zawierających wspólne kategorie” na stronie 39.

Dostosowywanie konstruktora tabel

W przeciwieństwie do standardowych okien dialogowych wielkość okna konstruktora tabel można zmienić w taki sam sposób, w jaki zmienia się wielkość standardowego okna aplikacji:

1. Kliknij i przeciągnij górną, dolną, boczną krawędź lub narożnik okna konstruktora tabel, aby je powiększyć lub zmniejszyć.

   Na karcie Tabela można również zmienić wielkość listy zmiennych, listy kategorii i panelu obszaru roboczego.
2. Kliknij i przeciągnij pasek poziomy rozdzielający listę zmiennych i listę kategorii w celu ich wydłużenia lub skrócenia. Przeciągnięcie w dół powoduje wydłużenie listy zmiennych i skrócenie listy kategorii. Przeciągnięcie w górę wywołuje skutek odwrotny.
3. Kliknij i przeciągnij pasek pionowy oddzielający listę zmiennych i kategorii od panelu obszaru roboczego w celu ich rozszerzenia lub zwężenia. Wielkość obszaru roboczego zmienia się automatycznie, dostosowując się do ilości pozostałego miejsca.

Tabele użytkownika: karta Opcje

Wygląd komórek danych

Określa elementy wyświetlane w pustych komórkach i komórkach, dla których obliczenie statystyk jest niemożliwe.

Puste komórki
W przypadku komórek tabeli niezawierających obserwacji (liczebność komórek równa 0) można wybrać jedną z trzech opcji wyświetlania: zero, puste lub określoną wartość tekstową. Wartość tekstowa może składać się maksymalnie z 255 znaków.

Statystyki, które nie mogą zostać wyliczone
Tekst wyświetlany w przypadku, kiedy nie można wyliczyć statystyki (np. średnia dla kategorii niezawierającej obserwacji). Wartość tekstowa może składać się maksymalnie z 255 znaków. Domyślna wartość to kropka (.).

Szerokość kolumn danych
Określa maksymalną i minimalną szerokość kolumn danych. Ustawienie to nie wpływa na szerokości kolumn zawierających etykiety wierszy.

Ustawienia szablonu TableLook
Powoduje zastosowanie specyfikacji szerokości kolumn danych bieżącym domyślnym szablonie TableLook. Można utworzyć własny domyślny szablon TableLook, który będzie wykorzystywany przy tworzeniu nowych tabel i będzie określać szerokość kolumn zawierających etykiety wierszy oraz kolumn danych.

Użytkownika
Powoduje zignorowanie domyślnych ustawień szerokości kolumn danych określonych w szablonie TableLook. Umożliwia określenie minimalnych i maksymalnych szerokości kolumn danych w tabeli oraz jednostki pomiarowej: punktów, cali lub centymetrów.

Braki danych dla zmiennych ilościowych
W przypadku tabel zawierających dwie lub więcej zmiennych ilościowych, opcja ta określa sposób postępowania z brakami danych przy wyliczaniu statystyk dla takich zmiennych.

Maksymalizuj użycie dostępnych danych (usuwanie według zmiennych)
Wszystkie obserwacje zawierające prawidłowe wartości dla każdej zmiennej ilościowej zostają uwzględnione w statystykach podsumowujących dla takiej zmiennej.

Użyj jednorodnej bazy obserwacji dla zmiennych ilościowych (usuwanie obserwacjami)
Obserwacje zawierające braki danych dla dowolnej zmiennej ilościowej w tabeli zostają wykluczone ze statystyk podsumowujących dla wszystkich zmiennych ilościowych występujących w tabeli.

Rzeczywista podstawa
Jeśli mamy zmienną reprezentującą wagi korekt, a nie wagi częstości, można użyć takiej zmiennej jako rzeczywistej wagi podstawowej. Koncepcja ważenia według podstawy rzeczywistej lub rzeczywistej wielkości próby oparta jest na metodach analizowania danych z prób złożonych. Ważenie według podstawy rzeczywistej umożliwia przybliżone wywodzenie statystyczne w ramach analizy obejmującej wprowadzane ad hoc korekty danych pochodzących z prostych prób losowych w oparciu o wagi korekt.

- Waga według podstawy rzeczywistej wpływa na wartości ważonych statystyk podsumowujących oraz testy istotności średnich kolumnowych i proporcji kolumnowych.
- Jeśli dla zbioru danych włączone jest ważenie, zmienna wagi zbioru danych jest ignorowana, a wyniki są ważone według zmiennej będącej rzeczywistą podstawą.
- Zmienna rzeczywistej podstawy wagi musi być liczbowa.
- Obserwacje z wagami ujemnymi, równymi 0 lub bez wag są wykluczane ze wszystkich wyników.

Zlicz powtarzające się odpowiedzi dla zestawów wielokrotnych kategorii
Powtórzona odpowiedź jest taką samą odpowiedzią dla dwóch lub więcej zmiennych w zestawie wielokrotnych kategorii. Domyślnie zduplikowane odpowiedzi nie są zliczana.

Ukryj małe liczebności
Możesz zdecydować o tym, aby ukryć liczebności, które są mniejsze od określonej liczby całkowitej.

- Ukryte wartości są wyświetlane jako **<N**, gdzie **N** jest określoną liczbą całkowitą.
- Określona liczba całkowita musi być większa lub równa 2.

- Jeśli dla zbioru danych włączone jest ważenie lub określona jest zmienna będąca rzeczywistą podstawą ważenia, to używana jest wartość ważona.

Wagi a zaokrąglenia

- Jeśli do ważenia używana jest opcja **Dane** > **Ważenie obserwacji**, to w testach istotności i przy wyznaczaniu przedziałów ufności oraz błędów standardowych wagi niecałkowitoliczbowe są zaokrąglane na poziomie komórki lub kategorii.
- W wypadku wybrania opcji **Użyj rzeczywistej podstawy zmiennej ważącej** wagi niecałkowitoliczbowe w wybranej zmiennej ważącej nie są zaokrąglane.
- Jeśli wybrane są obie opcje, to używana jest rzeczywista podstawa, a niecałkowitoliczbowe wagi nie są zaokrąglane.

Tabele użytkownika: karta Tytuły

Karta Tytuły umożliwia określenie sposobu wyświetlania tytułów, nagłówków i narożników.

Tytuł. Tekst wyświetlany nad tabelą.

Nagłówek. Tekst wyświetlany pod tabelą, a nad przypisami.

Narożnik. Tekst wyświetlany w lewym górnym rogu tabeli. Tekst narożnika jest wyświetlany tylko wówczas, gdy tabela zawiera zmienne w wierszach i jeśli właściwość etykiety wymiaru wierszowego w tabeli przestawnej jest ustawiona na **Zagnieżdżona**. *Nie* jest to domyślne ustawienie szablonu TableLook.

W tytule, nagłówku i narożniku tabeli można umieszczać następujące automatycznie generowane wartości:

Data. Bieżący rok, miesiąc i dzień wyświetlany w formacie zgodnym z ustawieniami regionalnymi w Panelu sterowania systemu Windows.

Czas. Bieżąca godzina, minuta i sekunda wyświetlana w formacie zgodnym z ustawieniami regionalnymi w Panelu sterowania systemu Windows.

Wyrażenie tabelowe. Zmienne wykorzystywane w tabeli i sposób ich wykorzystania. Jeśli zmienna posiada zdefiniowaną etykietę, jest ona wyświetlana. W wygenerowanej tabeli następujące symbole oznaczają sposób wykorzystywania zmiennych w tabeli:

- **+** oznacza zmienne zestawione.
- **>** oznacza zagnieżdżenie.
- **BY** oznacza tabelę krzyżową lub warstwy.

Tabele użytkownika: karta Testy

Karta Testy udostępnia testy istotności dotyczące tabel użytkownika.

Testy te nie są dostępne dla tabel, w których etykiety kategorii zostały przeniesione poza swój domyślny wymiar w tabeli lub do przeliczonych kategorii.

Testy średnich kolumnowych i proporcji kolumnowych

Testy średnich kolumnowych są dostępne w przypadku zmiennych ilościowych. Testy proporcji kolumnowych są dostępne w przypadku zmiennych jakościowych.

Porównaj średnie kolumn

Przeprowadzane parami testy równości średnich kolumnowych. Tabela musi zawierać jedną zmienną jakościową w kolumnach, a jako najbardziej wewnętrzny poziom wierszy musi zawierać zmienną ilościową. Tabela musi zawierać średnią jako statystykę podsumowania.

W przypadku zwykłych zmiennych jakościowych wariancję można oszacować na podstawie wszystkich kategorii lub tylko kategorii, które są porównywane. W przypadku zmiennych wielokrotnych odpowiedzi wariancja w teście średnich zawsze oparta jest wyłącznie na kategoriach, które są porównywane.

Porównaj proporcje kolumn
Testy równości proporcji kolumnowych. Tabela musi zawierać co najmniej jedną zmienną jakościową zarówno w kolumnach, jak i wierszach. Tabela musi obejmować liczebności lub wartości procentowe kolumn.

Identyfikacja istotnych różnic
W testach średnich kolumnowych i proporcji kolumnowych można uwidocznić istotne wyniki w oddzielnej tabeli lub w głównej tabeli.

Przedstaw w oddzielnej tabeli
Wyniki testu istotności są wyświetlane w oddzielnej tabeli. Jeżeli dwie wartości znacząco się różnią, komórka zwierająca większą wartość wyświetla klucz określający kolumnę mniejszej wartości.

Pokaż wartości istotności
Wartości istotności są wyświetlane w nawiasach po każdej wartości klucza w komórce. Ta opcja jest dostępna tylko wtedy, gdy istotne wyniki są wyświetlane w oddzielnej tabeli.

Przedstaw w tabeli głównej
Wyniki testu istotności są wyświetlane w tabeli głównej. Każda kategoria kolumn w tabeli jest oznaczona kluczem alfabetycznym. Dla każdej istotnej pary klucz kategorii z mniejszą średnią lub proporcją kolumn uwidoczniony jest w kategorii z większą średnią lub proporcją kolumn.

- Zatrzymanie wskaźnika myszy nad kluczem w komórce etykiety kolumny wewnątrz tabeli przestawnej powoduje podświetlenie wszystkich komórek tabeli z tym kluczem istotności. W przypadku tabeli z więcej niż jedną zmienną w wymiarze kolumn podświetlane są tylko komórki w tej podtabeli.
- Aby zaznaczyć wszystkie komórki w tabeli (lub podtabeli) mające ten sam klucz istotności, prawym przyciskiem myszy kliknij komórkę z etykieta kolumny i wybierz kolejno opcje **Wybierz** > **Zaznacz wszystkie komórki z tym kluczem istotności**.

Indeksy dolne w stylu APA
Istnieje możliwość oznaczania istotnych różnic w stylu APA, przy użyciu indeksów dolnych. Jeśli dwie wartości są istotnie różne, to będą miały różne litery w indeksach dolnych. Te indeksy dolne nie są odsyłaczami do przypisów. Gdy działa niniejsza opcja, styl zdefiniowanego przypisu w w istniejącej TableLook jest nadpisywany i przypisy są wyświetlane jako liczby z indeksem górnym. Aby zaznaczyć wszystkie komórki znajdujące się w tym samym wierszu i mające ten sam klucz istotności, kliknij prawym przyciskiem myszy komórkę z kluczem istotności i wybierz opcję **Zaznacz komórki o podobnym znaczeniu**

Poziomy istotności
Poziom istotności dla testów średnich kolumnowych i proporcji kolumnowych.

- Wartość musi być liczbą większą od 0 i mniejszą od 1.
- W przypadku określenia dwóch poziomów istotności wartości istotności mniejsze lub równe mniejszemu poziomowi są oznaczane wielkimi literami. Wartości istotności mniejsze lub równe większemu poziomowi są oznaczane małymi literami.
- W razie wybrania opcji **Użyj indeksów typu APA** druga wartość jest ignorowana.

Skoryguj wartości p dla wielokrotnych porównań
Korekta metodą **Bonferroniego** koryguje wskaźnik błędu w rodzinie (FWER). Metoda **Benjaminiego-Hochberga** to korekta wskaźnika fałszywych wykryć (FDR). Metoda ta jest mniej konserwatywna niż korekta Bonferroniego.

Testy niezależności (chi-kwadrat)
Powoduje wykonanie testu niezależności chi-kwadrat dla tabel, w których zarówno w wierszach, jak i kolumnach występuje przynajmniej jedna zmienna jakościowa.

W miejsce kategorii z sumami pośrednimi użyj sum pośrednich
Wszystkie sumy pośrednie zastępują kategorie testów istotności. W innym wypadku tylko sumy pośrednie, dla których kategorie z sumami pośrednimi zostały ukryte, zastępują kategorie do testowania.

Uwzględnij zmienne wielokrotnych odpowiedzi
Powoduje, że kategorie zestawów wielokrotnych odpowiedzi są uwzględnione w testach istotności. W przeciwnym razie kategorie zestawów wielokrotnych odpowiedzi są uwzględnione w testach istotności.

Proste tabele dla zmiennych kategorialnych

Proste tabele dla zmiennych kategorialnych

Większość tabel, które użytkownik będzie chciał utworzyć, prawdopodobnie zawierać będzie przynajmniej jedną **zmienną kategorialną**. Zmienna kategorialne to taka zmienna, która posiada ograniczoną liczbę odrębnych wartości lub kategorii (np. płeć czy religia). Zmienne kategorialne mogą być **nominalne** lub **porządkowe**.

- *Nominalny*. Zmienna może być traktowana jako nominalna, gdy jej wartości reprezentują kategorie bez wewnętrznego rangowania; na przykład wydział, na którym są zatrudnieni pracownicy. Przykładami zmiennych nominalnych są: region, kod pocztowy lub wyznanie.
- *Porządkowy*. Zmienna może być traktowana jako porządkowa, gdy jej wartości reprezentują kategorię z wewnętrznym rangowaniem, na przykład poziomy zadowolenia z usługi od bardzo niezadowolonego do bardzo zadowolonego). Przykładami zmiennych porządkowych mogą być oceny opinii reprezentujące stopień satysfakcji lub przekonania oraz oceny preferencji.

Ikona obok każdej zmiennej na liście zmiennych określa jej rodzaj.

Tabela 7. Ikony poziomu pomiaru

	Liczba	**Łańcuch**	**Data**	**Czas**
Zmienna (ilościowa)		n/a		
Porządkowa				
Nominalna				

Moduł Tabele użytkownika jest zoptymalizowany do użytku ze zmiennymi kategorialnymi posiadającymi zdefiniowane **etykiety wartości**. Więcej informacji zawiera temat “Tworzenie tabel” na stronie 1.

Przykładowy plik danych

W przykładach podanych w tym rozdziale wykorzystano plik danych *survey_sample.sav*. Więcej informacji zawiera temat “Pliki przykładowe” na stronie 77.

We wszystkich przedstawionych tutaj przykładach w oknach dialogowych wyświetlane są etykiety zmiennych, posortowane w porządku alfabetycznym. Opcje wyświetlania listy zmiennych ustawiane są na karcie Ogólne, w oknie dialogowym Opcje (menu Edycja, Opcje).

Pojedyncza zmienna kategorialna

Pomimo faktu, iż tabela z jedną zmienną kategorialną może być jedną z najprostszych tabel do utworzenia, często w zupełności wystarcza.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do obszaru Wiersze w panelu obszaru roboczego.

W panelu obszaru roboczego wyświetlony zostanie podgląd tabeli. Na podglądzie nie są wyświetlane rzeczywiste wartości danych, a jedynie symbole zastępcze w ich miejscu.

3. Kliknij przycisk **OK**, aby utworzyć tabelę.

   Tabela zostaje wyświetlona w oknie Edytora raportów.

		Count
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	45 to 54	481
	55 to 64	320
	65 or older	479

Rysunek 2. Pojedyncza zmienna kategorialna w wierszach

W tej przykładowej tabeli nagłówek kolumny *Liczebność* jest właściwie zbędny i można utworzyć tabelę bez niego.

4. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
5. Zaznacz (kliknij) opcję **Ukryj** dla elementu Pozycja.
6. Kliknij przycisk **OK**, aby utworzyć tabelę.

Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	45 to 54	481
	55 to 64	320
	65 or older	479

Rysunek 3. Pojedyncza zmienna kategorialna bez etykiety kolumny dla statystyk podsumowujących

Procenty

Oprócz liczebności można również wyświetlić procenty. W przypadku prostej tabeli z jedną zmienną kategorialną, jeśli zmienna jest wyświetlana w wierszach, użytkownika prawdopodobnie interesują procenty w kolumnie. I odwrotnie, jeśli zmienna jest wyświetlana w kolumnach, prawdopodobnie interesują go procenty w wierszu.

1. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. W grupie Statystyki podsumowujące usuń zaznaczenie opcji **Ukryj** dla elementu Pozycja. Ponieważ ta tabela zawierać będzie dwie kolumny, korzystne jest wyświetlenie ich etykiet, aby wiedzieć, co dana kolumna reprezentuje.
3. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
4. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję **% z N w kolumnie** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
5. Z komórki Etykieta na liście Pokaż w tabeli usuń domyślną etykietę i wpisz `Procent`.
6. Kliknij przycisk **Zastosuj do wybranej**, a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Count	Percent
Age category	Less than 25	242	8.6%
	25 to 34	627	22.2%
	35 to 44	679	24.0%
	45 to 54	481	17.0%
	55 to 64	320	11.3%
	65 or older	479	16.9%

Rysunek 4. Liczebności i procenty w kolumnie

Podsumowania

Podsumowania nie są automatycznie dodawane do tabel użytkownika, lecz jest to czynność nieskomplikowana.

1. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego Zaznacz opcję **Kategorie i podsumowania**.
3. W oknie dialogowym Kategorie i podsumowania zaznacz (kliknij) opcję **Podsumowanie ogółem**.
4. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Count	Percent
Age category	Less than 25	242	8.6%
	25 to 34	627	22.2%
	35 to 44	679	24.0%
	45 to 54	481	17.0%
	55 to 64	320	11.3%
	65 or older	479	16.9%
	Total	2828	100.0%

Rysunek 5. Liczebności, procenty w kolumnie i podsumowania

Więcej informacji zawiera temat "Podsumowania i sumy pośrednie dla zmiennych kategorialnych" na stronie 30.

Tabela krzyżowa

Tabela krzyżowa umożliwia analizę relacji pomiędzy dwiema zmiennymi kategorialnymi. Na przykład wykorzystując *Kategorię wieku* jako zmienną w wierszu i *Płeć* jako zmienną w kolumnie można utworzyć dwuwymiarową tabelę krzyżową pokazującą liczbę kobiet i mężczyzn w każdej kategorii.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.
3. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do obszaru Wiersze w panelu obszaru roboczego.
4. Przeciągnij zmienną *Płeć* z listy zmiennych do pola Kolumny w panelu obszaru roboczego (w celu znalezienia tej zmiennej może zajść konieczność przewinięcia listy zmiennych).
5. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Gender	
		Male	Female
		Count	Count
Age category	Less than 25	108	134
	25 to 34	276	351
	35 to 44	309	370
	45 to 54	221	260
	55 to 64	136	184
	65 or older	178	301

Rysunek 6. Tabela krzyżowa dla zmiennych Kategoria wieku i Płeć

Procenty w tabelach krzyżowych

W dwuwymiarowej tabeli krzyżowej procenty w wierszach i kolumnach mogą okazać się przydatną informacją.

1. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy zmienną *Płeć* w panelu obszaru roboczego.

   Można zauważyć, że opcja **Statystyki podsumowujące** jest nieaktywna w podręcznym menu kontekstowym. Wynika to z faktu, że statystyki podsumowujące można wybrać tylko dla zmiennej najbardziej zagnieżdżonej w wymiarze statystyki źródłowej. Domyślny wymiar statystyki źródłowej (wiersz lub kolumna) w przypadku zmiennych kategorialnych zależy od kolejności przeciągania

zmiennych do panelu obszaru roboczego. W powyższym przykładzie do wymiaru wierszy przeciągnięta została w pierwszej kolejności zmienna *Kategoria wieku*, a ponieważ w wymiarze wierszy nie występują inne zmienne, *Kategoria wieku* staje się zmienną statystyki źródłowej. Wymiar statystyki źródłowej można zmienić, jednak w tym przykładzie nie jest to konieczne. Więcej informacji zawiera temat "Statystyki podsumowujące" na stronie 6.

3. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
4. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję **% z N w kolumnie** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
5. Na liście Statystyki zaznacz pozycję **% z N w wierszu** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
6. Kliknij przycisk **Zastosuj do wybranej**, a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Gender					
		Male			Female		
		Count	Column N %	Row N %	Count	Column N %	Row N %
Age category	Less than 25	108	8.8%	44.6%	134	8.4%	55.4%
	25 to 34	276	22.5%	44.0%	351	21.9%	56.0%
	35 to 44	309	25.2%	45.5%	370	23.1%	54.5%
	45 to 54	221	18.0%	45.9%	260	16.3%	54.1%
	55 to 64	136	11.1%	42.5%	184	11.5%	57.5%
	65 or older	178	14.5%	37.2%	301	18.8%	62.8%

Rysunek 7. Tabela krzyżowa z procentami w wierszu i w kolumnie

Zmiana formatu wyświetlania

Możliwa jest zmiana formatu wyświetlania, w tym liczby miejsc dziesiętnych wyświetlanych w statystykach podsumowujących. Na przykład procenty są domyślnie wyświetlane z jednym miejscem dziesiętnym i symbolem procentów. Co zrobić, aby wartości komórek wyświetlane były z dwoma miejscami dziesiętnymi i bez symbolu procentów?

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
3. Dla dwóch zaznaczonych procentowych statystyk podsumowujących (**% z N w kolumnie** i **% z N w wierszu**) z rozwijanej listy Format wybierz **nnnn.n** i w komórce Dziesiętne wpisz dla obu wartości 2.
4. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Gender					
		Male			Female		
		Count	Column N %	Row N %	Count	Column N %	Row N %
Age category	Less than 25	108	8.79	44.63	134	8.38	55.37
	25 to 34	276	22.48	44.02	351	21.94	55.98
	35 to 44	309	25.16	45.51	370	23.13	54.49
	45 to 54	221	18.00	45.95	260	16.25	54.05
	55 to 64	136	11.07	42.50	184	11.50	57.50
	65 or older	178	14.50	37.16	301	18.81	62.84

Rysunek 8. Sformatowane komórki wyświetlające procenty w wierszach i kolumnach

Sumy brzegowe

W tabelach krzyżowych dość powszechne jest wyświetlanie **sum brzegowych** — podsumowania każdego wiersza i kolumny. Ponieważ nie są one domyślnie umieszczane w tabelach użytkownika, należy to zrobić własnoręcznie.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.

3. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do obszaru Wiersze w panelu obszaru roboczego.
4. Przeciągnij zmienną *Płeć* z listy zmiennych do pola Kolumny w panelu obszaru roboczego (w celu znalezienia tej zmiennej może zajść konieczność przewinięcia listy zmiennych).
5. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego Zaznacz opcję **Kategorie i podsumowania**.
6. W oknie dialogowym Kategorie i podsumowania zaznacz (kliknij) opcję **Podsumowanie ogółem**, a następnie kliknij przycisk **Zastosuj**.
7. Kliknij prawym przyciskiem myszy zmienną *Płeć* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
8. W oknie dialogowym Kategorie i podsumowania zaznacz (kliknij) opcję **Podsumowanie ogółem**, a następnie kliknij przycisk **Zastosuj**.
9. Zaznacz (kliknij) opcję **Ukryj** dla elementu Pozycja. (ponieważ wyświetlane są tylko liczebności nie zachodzi konieczność identyfikacji „statystyk" wyświetlanych w komórkach danych tabeli).
10. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Gender		
		Male	Female	Total
Age category	Less than 25	108	134	242
	25 to 34	276	351	627
	35 to 44	309	370	679
	45 to 54	221	260	481
	55 to 64	136	184	320
	65 or older	178	301	479
	Total	1228	1600	2828

Rysunek 9. Tabela krzyżowa z sumami brzegowymi

Sortowanie i wykluczanie kategorii

Domyślnie kategorie są wyświetlane w rosnącym porządku wartości danych, które reprezentują etykiety wartości kategorii. Na przykład chociaż dla zmiennej określającej kategorie są wyświetlane etykiety wartości *Poniżej 25*, *25 — 34*, *35 — 44*, ... itd., faktyczne wartości danych wynoszą 1, 2, 3, ... itd. i to one decydują o domyślnym porządku wyświetlania kategorii.

W prosty sposób można zmienić porządek kategorii i wykluczyć te, które mają nie być wyświetlane w tabeli.

Sortowanie kategorii

Można ręcznie zmienić ustawienie kategorii lub posortować je w porządku malejącym lub rosnącym:

- Wartości danych.
- Etykiety wartości.
- Liczby komórek.
- Statystyki podsumowujące. Dostępność statystyk podsumowujących sortowanie zależy od statystyk podsumowujących, które zostały wybrane z przeznaczeniem do wyświetlania w tabeli.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Jeśli w polu Wiersze w panelu obszaru roboczego nie jest jeszcze wyświetlana zmienna *Kategoria wieku* , przeciągnij ją tam.
3. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego Zaznacz opcję **Kategorie i podsumowania**.

   Zarówno wartości danych, jak i powiązane z nimi etykiety wartości, są wyświetlane zgodnie z bieżącym porządkiem wyświetlania, czyli w tym przypadku nadal w porządku rosnącym wartości danych.
4. W grupie Sortuj kategorie, z rozwijanej listy Porządek wybierz opcję **Malejąco**.

   Porządek sortowania został odwrócony.

5. Z rozwijanej listy Według wybierz opcję **Etykiety**.

   Kategorie zostaną posortowane w malejącym porządku alfabetycznym etykiet wartości.

   Należy zauważyć, że kategoria z etykietą *Poniżej 25* znajduje się u góry listy. Przy sortowaniu w porządku alfabetycznym litery występują po cyfrach. Ponieważ jest to jedyna etykieta rozpoczynająca się od litery i ponieważ lista jest sortowana w porządku malejącym (odwrotnym), kategoria ta zostaje umieszczona na górze listy.

   Aby dana kategoria występowała w innymi miejscu listy, można ją w prosty sposób przemieścić.
6. Kliknij kategorię z etykietą *Poniżej 25* na liście etykiet.
7. Kliknij przycisk ze strzałką skierowaną w dół, znajdujący się z prawej strony listy. Kategoria zostanie przesunięta o jeden wiersz w dół listy.
8. Klikaj przycisk ze strzałką skierowaną w dół do momentu, gdy kategoria znajdzie się u dołu listy.

Wykluczanie kategorii

Jeśli niektóre kategorie mają nie występować w tabeli, można je wykluczyć.

1. Kliknij kategorię z etykietą *Poniżej 25* na liście etykiet.
2. Kliknij przycisk ze strzałką, znajdujący się z lewej strony listy.
3. Kliknij kategorię z etykietą *65+* na liście etykiet.
4. Ponownie kliknij przycisk ze strzałką, znajdujący się z lewej strony listy.

   Dwie kategorie zostaną przeniesione z listy Pokaż w tabeli na listę Wyklucz kategorie. W przypadku zmiany zdania można je w prosty sposób przenieść z powrotem na listę Pokaż w tabeli.
5. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Gender		
		Male	Female	Total
Age category	55 to 64	136	184	320
	45 to 54	221	260	481
	35 to 44	309	370	679
	25 to 34	276	351	627
	Total	942	1165	2107

Rysunek 10. Tabela posortowana według etykiet wartości w porządku malejącym z wykluczonymi pewnymi kategoriami

Należy zauważyć, że podsumowania pokazują niższe wartości niż przed wykluczeniem dwóch kategorii. Wynika to z faktu, że podsumowania są obliczane na podstawie kategorii występujących w tabeli. Kategorie wykluczone nie są uwzględnione przy obliczaniu podsumowań. Więcej informacji zawiera temat “Podsumowania i sumy pośrednie dla zmiennych kategorialnych” na stronie 30.

Zestawianie, zagnieżdżanie oraz warstwy w przypadku zmiennych kategorialnych

Zestawianie, zagnieżdżanie oraz warstwy są metodami wyświetlania wielu zmiennych w jednej tabeli. W niniejszym rozdziale koncentrujemy się na stosowaniu tych technik do zmiennych kategorialnych, chociaż można ich używać również do zmiennych ilościowych.

Przykładowy plik danych

W przykładach podanych w tym rozdziale wykorzystano plik danych *survey_sample.sav*. Więcej informacji zawiera temat “Pliki przykładowe” na stronie 77.

We wszystkich przedstawionych tutaj przykładach w oknach dialogowych wyświetlane są etykiety zmiennych, posortowane w porządku alfabetycznym. Opcje wyświetlania listy zmiennych ustawiane są na karcie Ogólne, w oknie dialogowym Opcje (menu Edycja, Opcje).

Zestawianie zmiennych kategorialnych

Zestawianie polega na łączeniu oddzielnych tabel w celu jednoczesnego wyświetlenia informacji w nich zawartych. Można na przykład wyświetlić informacje na temat *płci* i *kategorii wiekowych* w oddzielnych sekcjach jednej tabeli.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. W oknie konstruktora tabel przeciągnij zmienną *Płeć* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
3. Przeciągnij zmienną *Kategoria wieku* z listy zmiennych do pola Wiersze pod zmienną *Płeć*.

   Te dwie zmienne są teraz zestawione ze sobą w wymiarze wierszowym.
4. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Count
Gender	Male	1232
	Female	1600
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	45 to 54	481
	55 to 64	320
	65 or older	479

Rysunek 11. Tabela zmiennych kategorialnych zestawionych w wierszach

W podobny sposób można również zestawiać zmienne w kolumnach.

Zestawianie w tabelach krzyżowych

Tabela zestawiona może zawierać różne zmienne w różnych wymiarach. Można, na przykład umieścić dwie zestawione zmienne w wierszach, a trzecią w kolumnie.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Jeśli zmienne *Kategoria wieku* i *Płeć* nie są jeszcze zestawione w wierszach, uczyń to korzystając z powyższych wskazówek.
3. Przeciągnij zmienną *Informacje czerpie z Internetu* z listy zmiennych do pola Kolumny w panelu obszaru roboczego.
4. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Get news from internet	
		No	Yes
		Count	Count
Gender	Male	873	359
	Female	1092	508
Age category	Less than 25	146	96
	25 to 34	368	259
	35 to 44	435	244
	45 to 54	346	135
	55 to 64	252	68
	65 or older	416	63

Rysunek 12. Dwie zestawione zmienne w wierszach umieszczone w tabeli krzyżowej ze zmienną w kolumnie

Występuje kilka zmiennych z etykietami rozpoczynającymi się od Informacje czerpie z..., więc rozróżnienie ich na liście zmiennych może sprawiać trudność (ponieważ mogą nie być wyświetlane na liście zmiennych w całości). Istnieją dwa sposoby pokazania całej etykiety danej zmiennej:

- Umieść wskaźnik myszy na zmiennej umieszczonej na liście, a wówczas w wyskakującej podpowiedzi zostanie wyświetlona cała jej etykieta.
- Kliknij i przeciągnij pasek pionowy oddzielający listę zmiennych i kategorii od panelu obszaru roboczego w celu ich poszerzenia.

Zagnieżdżanie zmiennych kategorialnych

Zagnieżdżanie, podobnie jak umieszczanie w tabelach krzyżowych, umożliwia pokazanie zależności pomiędzy dwiema zmiennymi kategorialnymi, z tą różnicą, że jedna zmienna jest zagnieżdżona w drugiej w tym samym wymiarze. Można na przykład zagnieździć zmienną *Płeć* w zmiennej *Kategoria wieku* w wymiarze wierszowym, pokazując liczbę kobiet i mężczyzn w każdej kategorii wiekowej.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.
3. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do obszaru Wiersze w panelu obszaru roboczego.
4. Przeciągnij zmienną *Płeć* z listy zmiennych i umieść ją z prawej strony zmiennej *Kategoria wieku* w polu Wiersze.

   Na podglądzie w panelu obszaru roboczego widać teraz, że zagnieżdżona tabela zawierać będzie jedną kolumnę liczebności, a jej każda komórka liczbę kobiet i mężczyzn w poszczególnych kategoriach wiekowych.

   Można zauważyć, że etykieta zmiennej *Płeć* jest wyświetlana kilkakrotnie, po jednym razie dla każdej kategorii wiekowej. Liczbę takich powtórzeń można zminimalizować umieszczając na najniższym poziomie zagnieżdżenia zmienną z najmniejszą liczbą kategorii.
5. Kliknij zmienną *Płeć* w panelu obszaru roboczego.
6. Przeciągnij zmienną do lewej krawędzi pola Wiersze.

   Teraz zamiast sześciokrotnego wyświetlania etykiety zmiennej *Płeć*, dwukrotnie zostaje powtórzona etykieta zmiennej *Kategoria wieku*. Tabela jest bardziej uporządkowana, a jednocześnie pokazuje identyczne wyniki.
7. Kliknij przycisk **OK**, aby utworzyć tabelę.

				Count
Gender	Male	Age category	Less than 25	108
			25 to 34	276
			35 to 44	309
			45 to 54	221
			55 to 64	136
			65 or older	178
	Female	Age category	Less than 25	134
			25 to 34	351
			35 to 44	370
			45 to 54	260
			55 to 64	184
			65 or older	301

Rysunek 13. Tabela ze zmienną Kategoria wieku zagnieżdżoną w zmiennej Płeć

Uwaga: Tabele użytkownika nie uwzględniają przetwarzania plików powstałych w wyniku podzielenia danych wg warstw. Aby uzyskać taki sam wynik, jak dla plików powstałych w wyniku podzielenia danych wg warstw, umieść zmienne podziału danych w warstwach na skrajnym zewnętrznym poziomie zagnieżdżenia w tabeli.

Ukrywanie etykiet zmiennych

Innym sposobem eliminacji zbędnych etykiet zmiennych w tabelach zagnieżdżonych jest po prostu ukrycie nazw lub etykiet zmiennych. Ponieważ etykiety wartości zmiennych *Płeć* i *Kategoria wieku* są prawdopodobnie wystarczająco opisowe bez etykiet zmiennych, można wyeliminować takie etykiety dla obu zmiennych.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i w podręcznym menu kontekstowym usuń zaznaczenie opcji **Pokaż etykietę zmiennej**.
3. Powtórz czynność dla zmiennej *Płeć*.

Etykiety zmiennych są nadal wyświetlane na podglądzie tabeli, lecz nie zostaną dołączone do samej tabeli.

4. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Count
Male	Less than 25	108
	25 to 34	276
	35 to 44	309
	45 to 54	221
	55 to 64	136
	65 or older	178
Female	Less than 25	134
	25 to 34	351
	35 to 44	370
	45 to 54	260
	55 to 64	184
	65 or older	301

Rysunek 14. Tabela zagnieżdżona bez etykiet zmiennych

Aby etykiety zmiennych były dołączone do tabeli — bez wielokrotnego wyświetlania ich w jej części głównej — można je umieścić w tytule tabeli lub w narożniku.

5. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
6. Kliknij kartę **Tytuły**.
7. Kliknij w dowolnym miejscu pola Tytuł.
8. Kliknij przycisk **Wyrażenie tabelowe**. W polu Tytuł wyświetlony zostanie tekst *&[Table Expression]*. Spowoduje to wygenerowanie tytułu tabeli zawierającego etykiety zmiennych wykorzystanych w tabeli.
9. Kliknij przycisk **OK**, aby utworzyć tabelę.

Gender > Age category

		Count
Male	Less than 25	108
	25 to 34	276
	35 to 44	309
	45 to 54	221
	55 to 64	136
	65 or older	178
Female	Less than 25	134
	25 to 34	351
	35 to 44	370
	45 to 54	260
	55 to 64	184
	65 or older	301

Rysunek 15. Etykiety zmiennych w tytule tabeli

Symbol „większe niż” (>) w tytule wskazuje, że zmienna *Kategoria wieku* jest zagnieżdżona w zmiennej *Płeć*.

Zagnieżdżone tabele krzyżowe

Tabela zagnieżdżona może zawierać różne zmienne w różnych wymiarach. Można na przykład zagnieździć zmienną *Kategoria wieku* w zmiennej *Płeć* w wymiarze wierszy i umieścić zagnieżdżone wiersze w tabeli krzyżowej z trzecią zmienną w wymiarze kolumn.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Jeśli zmienna *Kategoria wieku* nie jest jeszcze zagnieżdżona w zmiennej *Płeć*, uczyń to korzystając ze wskazówek z poprzedniego przykładu.
3. Przeciągnij zmienną *Informacje czerpie z Internetu* z listy zmiennych do pola Kolumny w panelu obszaru roboczego.

Można zauważyć, że tabela nie mieści się w panelu obszaru roboczego. Aby zobaczyć inne fragmenty tabeli na jej podglądzie użyj pasków przewijania w górę/w dół i w lewo/w prawo lub:

- Kliknij przycisk **Kompaktowy** w konstruktorze tabel, aby wyświetlić tabelę w widoku kompaktowym. Pokazywane są w nim same etykiety zmiennych, bez informacji na temat kategorii czy statystyk podsumowujących.
- Powiększ okno konstruktora tabel klikając i przeciągając jego krawędzie lub narożniki.

4. Kliknij przycisk **OK**, aby utworzyć tabelę.

				Get news from internet	
				No	Yes
				Count	Count
Gender	Male	Age category	Less than 25	59	49
			25 to 34	159	117
			35 to 44	217	92
			45 to 54	169	52
			55 to 64	112	24
			65 or older	155	23
	Female	Age category	Less than 25	87	47
			25 to 34	209	142
			35 to 44	218	152
			45 to 54	177	83
			55 to 64	140	44
			65 or older	261	40

Rysunek 16. Zagnieżdżone tabele krzyżowe

Zamiana miejscami wierszy i kolumn

Co ma zrobić użytkownik, który poświęcił mnóstwo czasu na utworzenie złożonej tabeli i uznał, że jest idealna poza tym, że chce zmienić jej orientację, zamieniając miejscami zmienne w wierszach i kolumnach? Na przykład utworzona została zagnieżdżona tabela krzyżowa ze zmiennymi *Kategoria wieku* i *Płeć* zagnieżdżonymi w wierszach lecz teraz te dwie zmienne demograficzne mają być zagnieżdżone w kolumnach.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy w dowolnym miejscu panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Zamień wiersze z kolumnami**.

   Zmienne w wierszach i kolumnach zostały zamienione miejscami.

   Przed utworzeniem tabeli zostanie wprowadzonych kilka modyfikacji w celu jej uporządkowania.
3. Zaznacz opcję **Ukryj**, aby zapobiec wyświetlaniu etykiety kolumny statystyk podsumowujących.
4. Kliknij prawym przyciskiem myszy zmienną *Płeć* w panelu obszaru roboczego i usuń zaznaczenie opcji **Pokaż etykietę zmiennej**.
5. Następnie kliknij przycisk **OK**, aby utworzyć tabelę.

		Male						Female					
		Age category						Age category					
		Less than 25	25 to 34	35 to 44	45 to 54	55 to 64	65 or older	Less than	25 to 34	35 to 44	45 to 54	55 to 64	65 or older
Get news from internet	No	59	159	217	169	112	155	87	209	218	177	140	261
	Yes	49	117	92	52	24	23	47	142	152	83	44	40

Rysunek 17. Tabela krzyżowa ze zmiennymi demograficznymi zagnieżdżonymi w kolumnach

Warstwy

Warstwy umożliwiają dodanie do tabel wymiaru głębokości, przekształcając je w trójwymiarowe „kostki” danych. Warstwy przypominają, w zasadzie, zagnieżdżanie lub zestawianie; podstawową różnicą jest fakt, że jednocześnie widoczna jest kategoria tylko jednej warstwy. Na przykład wykorzystując *Kategoria wieku* jako zmienną w wierszu i *Płeć* jako zmienną w warstwie można utworzyć tabelę, w której informacje dotyczące kobiet i mężczyzn wyświetlane są w różnych warstwach.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.
3. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do obszaru Wiersze w panelu obszaru roboczego.
4. Kliknij przycisk **Warstwy** w górnej części konstruktora tabel, aby wyświetlić listę warstw.
5. Przeciągnij zmienną *Płeć* z listy zmiennych do listy warstw.

   Można zauważyć, że dodanie zmiennej w warstwie nie ma widocznego wpływu na podgląd wyświetlany w panelu obszaru roboczego. Zmienne w warstwach nie wpływają na podgląd w panelu obszaru roboczego, chyba że zmienna taka jest zmienną statystyki źródłowej, a statystyki podsumowujące zostaną zmienione.
6. Kliknij przycisk **OK**, aby utworzyć tabelę.

Gender Male

		Count
Age category	Less than 25	108
	25 to 34	276
	35 to 44	309
	45 to 54	221
	55 to 64	136
	65 or older	178

Rysunek 18. Prosta tabela warstwowa

Na pierwszy rzut oka tabela ta nie różni się od prostej tabeli z jedną zmienną kategorialną. Jedyną różnicę stanowi obecność etykiety *Płeć respondenta Mężczyzna* w górnej części tabeli.

7. Kliknij dwukrotnie tabelę w oknie Edytora raportów, aby ją uaktywnić.
8. Widać teraz, że etykieta *Płeć respondenta Mężczyzna* stanowi element listy rozwijanej.
9. Kliknij strzałkę z prawej strony listy rozwijanej, aby wyświetlić pełną listę warstw.

   W tej tabeli lista zawiera jeszcze tylko jeden element.
10. Z rozwijanej listy wybierz *Płeć respondenta Kobieta*.

Gender Female

		Count
Age category	Less than 25	134
	25 to 34	351
	35 to 44	370
	45 to 54	260
	55 to 64	184
	65 or older	301

Rysunek 19. Prosta tabela warstwowa z wyświetlonymi różnymi warstwami

Dwie zestawione zmienne jakościowe w warstwach

Jeśli w warstwach występuje kilka zmiennych kategorialnych, można je zestawić lub zagnieździć. Domyślnie zmienne w warstwach są zestawione. (*Uwaga*: Jeśli w warstwach występują zmienne ilościowe, mogą one zostać wyłącznie zestawione).

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Jeśli zmienna *Kategoria wieku* nie występuje jeszcze w wierszach, a zmienna *Płeć* w warstwach, utwórz tabelę warstwową korzystając ze wskazówek z poprzedniego przykładu.
3. Przeciągnij zmienną *Poziom wykształcenia respondenta* z listy zmiennych do listy warstw, pod zmienną *Płeć*.

   Dwa przyciski opcji pod listą warstw w grupie Warstwy tworzą są teraz aktywne. Domyślnie zaznaczoną opcją jest **Oddzielne kategorie (zestawienie)**. Opcja ta jest równoważna zestawieniu.
4. Kliknij przycisk **OK**, aby utworzyć tabelę.
5. Kliknij dwukrotnie tabelę w oknie Edytora raportów, aby ją uaktywnić.

6. Kliknij strzałkę z prawej strony listy rozwijanej, aby wyświetlić pełną listę warstw.

Tabela zawiera siedem warstw: dwie warstwy dla dwóch kategorii *Płeć* i pięć warstw dla pięciu kategorii *Poziom wykształcenia respondenta*. W przypadku warstw zestawionych całkowita liczba warstw jest sumą liczby kategorii zmiennych w warstwach (łącznie z kategoriami podsumowań lub sum pośrednich utworzonymi dla zmiennych w warstwach).

Dwie zagnieżdżone zmienne jakościowe w warstwach

Zagnieżdżenie zmiennych jakościowych w warstwach powoduje utworzenie odrębnej warstwy dla każdej kombinacji kategorii takich zmiennych.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Jeśli tabela z zestawionymi warstwami nie została jeszcze utworzona, utwórz ją wykorzystując wskazówki z poprzedniego przykładu.
3. W grupie Warstwy tworzą zaznacz opcję **Kombinacje kategorii (zagnieżdżanie)**. Opcja ta jest równoważna zagnieżdżaniu.
4. Kliknij przycisk **OK**, aby utworzyć tabelę.
5. Kliknij dwukrotnie tabelę w oknie Edytora raportów, aby ją uaktywnić.
6. Kliknij strzałkę z prawej strony listy rozwijanej, aby wyświetlić pełną listę warstw.

W tabeli występuje 10 warstw (aby zobaczyć je wszystkie, należy przewinąć listę), po jednej dla każdej kombinacji kategorii zmiennych *Płeć* i *Poziom wykształcenia respondenta*. W przypadku warstw zagnieżdżonych całkowita liczba warstw stanowi *iloczyn* liczby kategorii każdej zmiennej w warstwie i liczby zmiennych (w podanym przykładzie 5 x 2 = 10).

Drukowanie tabel warstwowych

Domyślnie drukowana jest tylko aktualnie widoczna warstwa. Drukowanie wszystkich warstw tabeli:

1. Kliknij dwukrotnie tabelę w oknie Edytora raportów, aby ją uaktywnić.
2. Z menu Edytora raportów wybierz kolejno następujące pozycje:

   Format > **Właściwości tabeli...**
3. Kliknij kartę **Drukowanie**.
4. Zaznacz opcję **Drukuj wszystkie warstwy**.

Możesz również zapisać to ustawienie w szablonie TableLook, w tym domyślny szablon TableLook.

Podsumowania i sumy pośrednie dla zmiennych kategorialnych

W tabelach użytkownika można umieszczać zarówno podsumowania, jak i sumy pośrednie. Można je stosować do zmiennych kategorialnych na dowolnym poziomie zagnieżdżenia, w dowolnym wymiarze — wierszach, kolumnach i warstwach.

Przykładowy plik danych

W przykładach podanych w tym rozdziale wykorzystano plik danych *survey_sample.sav*. Więcej informacji zawiera temat “Pliki przykładowe” na stronie 77.

We wszystkich przedstawionych tutaj przykładach w oknach dialogowych wyświetlane są etykiety zmiennych, posortowane w porządku alfabetycznym. Opcje wyświetlania listy zmiennych ustawiane są na karcie Ogólne, w oknie dialogowym Opcje (menu Edycja, Opcje).

Proste podsumowanie dla jednej zmiennej

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**

2. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do obszaru Wiersze w panelu obszaru roboczego.
3. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
4. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję **% z N w kolumnie** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
5. Z komórki Etykieta na liście Pokaż w tabeli usuń domyślną etykietę i wpisz `Procent`.
6. Kliknij przycisk **Zastosuj do wybranej**.
7. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
8. W oknie dialogowym Kategorie i podsumowania zaznacz (kliknij) opcję **Podsumowanie ogółem**.
9. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Count	Percent
Age category	Less than 25	242	8.6%
	25 to 34	627	22.2%
	35 to 44	679	24.0%
	45 to 54	481	17.0%
	55 to 64	320	11.3%
	65 or older	479	16.9%
	Total	2828	100.0%

Rysunek 20. Proste podsumowanie dla jednej zmiennej kategorialnej

Sumowane są tylko wyświetlane kategorie

Podsumowania uwzględniają kategorie wyświetlane w tabeli. Jeśli pewne kategorie zostaną wykluczone z tabeli, obserwacje do nich należące nie są uwzględniane w obliczeniach podsumowań.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
3. Kliknij kategorię z etykietą *Poniżej 25* na liście etykiet.
4. Kliknij przycisk ze strzałką, znajdujący się z lewej strony listy.
5. Kliknij kategorię z etykietą *65+* na liście etykiet.
6. Ponownie kliknij przycisk ze strzałką, znajdujący się z lewej strony listy.

   Dwie kategorie zostaną przeniesione z listy Pokaż w tabeli na listę Wyklucz kategorie.
7. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Count	Percent
Age category	25 to 34	627	29.8%
	35 to 44	679	32.2%
	45 to 54	481	22.8%
	55 to 64	320	15.2%
	Total	2107	100.0%

Rysunek 21. Podsumowanie w tabeli z wykluczonymi kategoriami

Całkowita liczba obserwacji w tej tabeli wynosi tylko 2107, dużo mniej niż 2828, gdyby wszystkie kategorie były uwzględniane. W podsumowaniu są uwzględniane wyłącznie kategorie umieszczone w tabeli. (suma procentów nadal wynosi 100%, ponieważ wszystkie wartości procentowe są obliczane na podstawie całkowitej liczby obserwacji wykorzystanych w tabeli, a nie całkowitej liczby obserwacji zawartych w pliku danych).

Miejsce wyświetlania podsumowań

Podsumowania są domyślnie wyświetlane pod sumowanymi kategoriami. Miejsce ich wyświetlania można zmienić tak, aby były wyświetlane nad sumowanymi kategoriami.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
3. W grupie Łączne i Podgrupy zaznacz opcję **Powyżej kategorii**.
4. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Count	Percent
Age category	Total	2107	100.0%
	25 to 34	627	29.8%
	35 to 44	679	32.2%
	45 to 54	481	22.8%
	55 to 64	320	15.2%

Rysunek 22. Podsumowanie wyświetlane nad sumowanymi kategoriami

Podsumowania dla tabel zagnieżdżonych

Ponieważ podsumowania można stosować do zmiennych kategorialnych na dowolnym poziomie zagnieżdżenia, możliwe jest utworzenie tabel zawierających podsumowania ogółem w grupie na kilku poziomach zagnieżdżenia.

Podsumowania ogółem w grupie

Podsumowania dla zmiennych kategorialnych zagnieżdżonych w innych zmiennych kategorialnych reprezentują podsumowania ogółem w grupie.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Przeciągnij zmienną *Płeć* i umieść po lewej stronie zmiennej *Kategoria wieku* w panelu obszaru roboczego.
3. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.

   Przed utworzeniem tabeli przenieś podsumowania z powrotem pod sumowane kategorie.
4. W grupie Łączne i Podgrupy zaznacz opcję **Powyżej kategorii**.
5. Kliknij przycisk **Zastosuj**, aby zapisać ustawienie i powrócić do konstruktora tabel.
6. Kliknij przycisk **OK**, aby utworzyć tabelę.

				Count	Percent
Gender	Male	Age category	25 to 34	276	29.3%
			35 to 44	309	32.8%
			45 to 54	221	23.5%
			55 to 64	136	14.4%
			Total	942	100.0%
	Female	Age category	25 to 34	351	30.1%
			35 to 44	370	31.8%
			45 to 54	260	22.3%
			55 to 64	184	15.8%
			Total	1165	100.0%

Rysunek 23. Podsumowania kategorii zmiennej Kategoria wieku zagnieżdżone w kategoriach zmiennej Płeć

Tabela zawiera teraz dwa podsumowania ogółem w grupie: jedno dla mężczyzn i jedno dla kobiet.

Podsumowanie całości

Podsumowania zastosowane do zagnieżdżonych zmiennych są zawsze podsumowaniami ogółem w grupie, a nie podsumowaniami całości. Aby uzyskać podsumowanie całej tabeli należy zastosować podsumowania do zmiennej na skrajnym zewnętrznym poziomie zagnieżdżenia.

1. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy zmienną *Płeć* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
3. W oknie dialogowym Kategorie i podsumowania zaznacz (kliknij) opcję **Podsumowanie ogółem**.
4. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

				Count	Percent
Gender	Male	Age category	25 to 34	276	29.3%
			35 to 44	309	32.8%
			45 to 54	221	23.5%
			55 to 64	136	14.4%
			Total	942	100.0%
	Female	Age category	25 to 34	351	30.1%
			35 to 44	370	31.8%
			45 to 54	260	22.3%
			55 to 64	184	15.8%
			Total	1165	100.0%
	Total	Age category	25 to 34	627	29.8%
			35 to 44	679	32.2%
			45 to 54	481	22.8%
			55 to 64	320	15.2%
			Total	2107	100.0%

Rysunek 24. Podsumowania całości dla zagnieżdżonej tabeli

Należy zauważyć, że podsumowanie całości wynosi tylko 2107, a nie 2828. Dwie kategorie wieku są nadal wykluczone z tabeli, więc obserwacje należące do tych kategorii nie są uwzględniane w podsumowaniach.

Podsumowania dla zmiennych w warstwach

Podsumowania dla zmiennych w warstwach są wyświetlane w tabeli jako oddzielne warstwy.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Warstwy** w konstruktorze tabel, aby wyświetlić listę warstw.
3. Przeciągnij zmienną *Płeć* z pola wierszy w panelu obszaru roboczego na listę Warstwy.

   Uwaga: Ponieważ już określono podsumowania dla zmiennej *Płeć*, nie trzeba tego robić teraz. Przenoszenie zmiennej pomiędzy wymiarami nie wpływa na ustawienia dla tej zmiennej.
4. Kliknij przycisk **OK**, aby utworzyć tabelę.
5. Kliknij dwukrotnie tabelę w Edytorze raportów, aby ją uaktywnić.
6. Kliknij strzałkę skierowaną w dół z prawej strony listy rozwijanej Warstwa, aby wyświetlić wszystkie warstwy występujące w tabeli.

Tabela zawiera trzy warstwy: *Płeć męska*, *Płeć żeńska* i *Płeć ogółem*.

Miejsce wyświetlania podsumowań warstw

W przypadku podsumowań zmiennej warstwowej miejsce ich wyświetlania (poniżej lub powyżej) określa położenie warstwy podsumowań ogółem. Jeśli na przykład dla podsumowania zmiennej warstwowej określone zostanie położenie **Powyżej**, warstwa podsumowań ogółem jest pierwszą wyświetlaną warstwą.

Sumy pośrednie

Do tabeli można włączyć sumy pośrednie dla podzbiorów kategorii zmiennej. Można, na przykład włączyć sumy pośrednie dla kategorii wiekowych reprezentujących wszystkich respondentów przykładowej ankiety w wieku poniżej i powyżej 45 lat.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.
3. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do obszaru Wiersze w panelu obszaru roboczego.

4. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
5. Wybierz z listy Wartości **3,00**.
6. Kliknij polecenie **Dodaj sumę pośrednią**, aby wyświetlić okno dialogowe Definiuj sumę pośrednią.
7. W polu tekstowym etykieta wpisz `Suma pośrednia < 45`.
8. Następnie kliknij przycisk **Kontynuuj**.

   Spowoduje to wstawienie wiersza zawierającego sumę pośrednią dla pierwszych trzech kategorii wiekowych.
9. Wybierz z listy Wartości **6,00**.
10. Kliknij polecenie **Dodaj sumę pośrednią**, aby wyświetlić okno dialogowe Definiuj sumę pośrednią.
11. W polu tekstowym etykieta wpisz `Suma pośrednia 45+`.
12. Następnie kliknij przycisk **Kontynuuj**.

    Ważna uwaga: Miejsce wyświetlania podsumowań i sum pośrednich (**Powyżej kategorii których dotyczą** lub **Poniżej kategorii których dotyczą**) należy wybrać przed zdefiniowaniem jakiejkolwiek z nich. Zmiana położenia ma wpływ na wszystkie sumy pośrednie (a nie tylko aktualnie zaznaczone) i powoduje również *zmianę kategorii uwzględnionych w sumach pośrednich.*
13. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Count
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	Subtotal < 45	1548
	45 to 54	481
	55 to 64	320
	65 or older	479
	Subtotal 45+	1280

Rysunek 25. Sumy pośrednie dla kategorii wiekowej

W sumach pośrednich są uwzględniane tylko wyświetlane kategorie

Podobnie jak podsumowania, sumy pośrednie uwzględniają kategorie zawarte w tabeli.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.

   Uwaga:Wartość (nie etykieta wartości) wyświetlana dla pierwszej sumy pośredniej wynosi **1.00...3.00**, co oznacza, że suma pośrednia uwzględnia wszystkie wartości na liście od 1 do 3.
3. Wybierz z listy Wartości **1,00** (lub kliknij etykietę *Poniżej 25*).
4. Kliknij przycisk ze strzałką, znajdujący się z lewej strony listy.

Pierwsza kategoria wiekowa zostaje wykluczona, a wartość wyświetlana dla pierwszej sumy pośredniej zmienia się na **2,00...3,00**, co oznacza, że wykluczona kategoria nie zostanie uwzględniona w sumie pośredniej, gdyż sumy pośrednie uwzględniają tylko kategorie zawarte w tabeli. Wykluczenie kategorii powoduje automatyczne wykluczenie jej z sum pośrednich, więc nie można, na przykład wyświetlić samych sum pośrednich bez kategorii, na podstawie których są obliczane.

Ukrywanie kategorii z sumami pośrednimi

Można ukryć wyświetlane kategorie, definiujące sumy pośrednie i wyświetlać tylko sumy pośrednie, grupując (zwijając) kategorie bez naruszania danych.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.

3. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do obszaru Wiersze w panelu obszaru roboczego.
4. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
5. Wybierz z listy Wartości **3,00**.
6. Kliknij polecenie **Dodaj sumę pośrednią**, aby wyświetlić okno dialogowe Definiuj sumę pośrednią.
7. W polu tekstowym Etykieta wpisz `Poniżej 45`.
8. Wybierz (zaznacz) pole **Ukrywanie kategorii z sumami pośrednimi w tabeli**.
9. Następnie kliknij przycisk **Kontynuuj**.

   Spowoduje to wstawienie wiersza zawierającego sumę pośrednią dla pierwszych trzech kategorii wiekowych.
10. Wybierz z listy Wartości **6,00**.
11. Kliknij polecenie **Dodaj sumę pośrednią**, aby wyświetlić okno dialogowe Definiuj sumę pośrednią.
12. W polu tekstowym Etykieta wpisz `45 lub starsze`.
13. Wybierz (zaznacz) pole **Ukrywanie kategorii z sumami pośrednimi**.
14. Następnie kliknij przycisk **Kontynuuj**.
15. Aby włączyć sumę całkowitą z pośrednimi wybierz (zaznacz) **Łącznie** w grupie Pokaż.
16. Kliknij przycisk **Zastosuj**.

    W obszarze roboczym uwzględniono fakt, że sumy pośrednie będą wyświetlane, ale kategorie definiujące sumy pośrednie zostaną pominięte.
17. Kliknij przycisk **OK**, aby utworzyć tę tabelę.

		Count
Age category	Less than 45	1548
	45 or older	1280
	Total	2828

Rysunek 26. Tabela wyświetlająca tylko podsumowania i sumy pośrednie

Sumy pośrednie dla zmiennych w warstwach

Podobnie jak podsumowania, sumy pośrednie dla zmiennych w warstwach są wyświetlane w tabeli jako oddzielne warstwy. Sumy pośrednie są w zasadzie traktowane jako kategorie. Każda kategoria stanowi odrębną warstwę w tabeli, a kolejność wyświetlania kategorii warstwowych zależy od ustawień w oknie dialogowym Kategorie i podsumowania, które określają również miejsce wyświetlania kategorii sum pośrednich.

Przeliczone kategorie dla zmiennych kategorialnych

Do tabel użytkownika można włączyć przeliczone kategorie. Są to nowe kategorie, które są obliczane na podstawie kategorii w ten sam sposób na każdym poziomie zagnieżdżenia i dowolnym wymiarze--wierszu, kolumnie lub warstwie. Można na przykład włączyć obliczoną kategorię, która pokazuje różnicę między dwiema kategoriami.

Przykładowy plik danych

W przykładach podanych w tym rozdziale wykorzystano plik danych *survey_sample.sav*. Więcej informacji zawiera temat “Pliki przykładowe” na stronie 77.

Prosta przeliczona kategoria

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**

2. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do obszaru Wiersze w panelu obszaru roboczego.
3. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
4. Wybierz z listy Wartości **3,00**.
5. Kliknij polecenie **Dodaj kategorię**, aby wyświetlić okno dialogowe Zdefiniuj przeliczoną kategorię.
6. W polu tekstowym Etykieta przeliczonej kategorii wpisz `Poniżej 45`.
7. Wybierz wartość **Mniej niż 25 (1.00)** z Listy kategorii i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii. `[1]` wyświetla się w wyrażeniu.
8. Kliknij przycisk operatora plus (+) w oknie dialogowym (lub naciśnij klawisz + na klawiaturze).
9. Wybierz wartość **Od 25 do (34)** z Listy kategorii i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii.
10. Kliknij przycisk operatora plus (+) w oknie dialogowym (lub naciśnij klawisz + na klawiaturze).
11. Wybierz wartość **Od 35 do 44 (3.0)** z Listy kategorii i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii.
12. Następnie kliknij przycisk **Kontynuuj**.

    Spowoduje to wstawienie wiersza zawierającego sumę pośrednią dla pierwszych trzech kategorii wiekowych.
13. Wybierz z listy Wartości **5.00**.
14. Kliknij polecenie **Dodaj sumę pośrednią**, aby wyświetlić okno dialogowe Definiuj sumę pośrednią.
15. W polu tekstowym Etykieta wpisz `Poniżej 65`.
16. Następnie kliknij przycisk **Kontynuuj**.

    Spowoduje to wstawienie wiersza zawierającego sumę pośrednią dla pierwszych pięciu kategorii wiekowych.
17. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Count
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	Less than 45	1548
	45 to 54	481
	55 to 64	320
	Less than 65	2349
	65 or older	479

Rysunek 27. Przeliczona kategoria z sumą pośrednią

tabela zawiera przeliczoną kategorię (*Poniżej 45*) oraz sumę pośrednia (*Poniżej 65*). Suma pośrednia obejmuje kategorie, które się również zawierają w przeliczonej kategorii. Nie można utworzyć tej samej tabeli z samymi sumami pośrednimi, ponieważ sumy pośrednie nie mogą mieć tych samych kategorii.

Ukrywanie kategorii w przeliczonej kategorii

Podobnie jak w przypadku sum pośrednich, możliwe jest zapobiegnięcie wyświetlaniu kategorii, które są użyte w wyrażeniu przeliczonej kategorii i wyświetlają tylko samą przeliczoną kategorię. Poniższy przykład bazuje na poprzednim przykładzie.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
3. Wybierz przeliczoną kategorię *Poniżej 45* w liście Wartości.

4. Kliknij polecenie **Edytuj**, aby wyświetlić okno dialogowe Zdefiniuj przeliczoną kategorię.
5. Wybierz opcję **Ukryj kategorie używane w wyrażeniu z tabeli**.
6. Następnie kliknij przycisk **Kontynuuj**.
7. Wybierz sumę pośrednią *Poniżej 65* w liście Wartości.
8. Kliknij polecenie **Edytuj**, aby wyświetlić okno dialogowe Zdefiniuj sumę pośrednią.
9. Wybierz pole **Ukrywanie kategorii z sumami pośrednimi w tabeli**.
10. Następnie kliknij przycisk **Kontynuuj**.
11. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Count
Age category	Less than 45	1548
	Less than 65	2349
	65 or older	479

Rysunek 28. Przeliczona kategoria z sumą pośrednią i ukrytymi kategoriami

Jak w poprzednim przykładzie, tabela obejmuje przeliczoną kategorię i sumę pośrednią. Jednak w tym przypadku kategorie w każdej z nich są ukryte tak, aby wyświetlane były te podsumowania.

Sumy pośrednie odniesień w przeliczonej kategorii

Można włączyć sumy pośrednie do wyrażenia przeliczonej kategorii

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.
3. W oknie konstruktora tabel przeciągnij zmienną *Obecna sytuacja zawodowa respondenta* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
4. Przeciągnij zmienną *Stan cywilny* z listy zmiennych do pola Kolumny.
5. Kliknij prawym przyciskiem myszy zmienną *Obecna sytuacja zawodowa* w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Kategorie i podsumowania**.
6. Wybierz z listy Wartości **2**.
7. Kliknij polecenie **Dodaj sumę pośrednią**, aby wyświetlić okno dialogowe Definiuj sumę pośrednią.
8. W polu tekstowym Etykieta wpisz `Pracuje`.
9. Wybierz pole **Ukrywanie kategorii z sumami pośrednimi w tabeli**.
10. Następnie kliknij przycisk **Kontynuuj**.

    Spowoduje to wstawienie wiersza zawierającego sumę pośrednią dla pierwszych dwóch kategorii statusu zatrudnienia.
11. Wybierz z listy Wartości **8**.
12. Kliknij polecenie **Dodaj sumę pośrednią**, aby wyświetlić okno dialogowe Definiuj sumę pośrednią.
13. W polu tekstowym Etykieta wpisz `Nie pracuje`.
14. Wybierz pole **Ukrywanie kategorii z sumami pośrednimi**.
15. Następnie kliknij przycisk **Kontynuuj**.

    Spowoduje to wstawienie wiersza zawierającego sumę pośrednią dla innych kategorii statusu zatrudnienia.
16. Wybierz sumę pośrednią *Nie pracuje* w liście Wartości.
17. Kliknij polecenie **Dodaj kategorię**, aby wyświetlić okno dialogowe Zdefiniuj przeliczoną kategorię.
18. W polu tekstowym Etykieta przeliczonej kategorii wpisz `Pracuje / Nie pracuje`.
19. Wybierz **Pracuje (Pracuje #1)** z Listy podsumowań i sum pośrednich i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii.

20. Kliknij przycisk operatora dzielenia (/) w oknie dialogowym (lub naciśnij klawisz / na klawiaturze).
21. Wybierz **Nie pracuje (Nie pracuje #2)** z Listy podsumowań i sum pośrednich i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii..

    Przeliczana kategoria domyślnie wykorzystuje ten sam format co statystyka zmiennej, która w tym przypadku jest Liczebnością. Ponieważ miejsca dziesiętne wynikające z dzielenia mają być widoczne w wyrażeniu przeliczonej kategorii, a domyślny format Liczebności nie zawiera miejsc dziesiętnych, trzeba zmienić format.
22. Kliknij kartę Wyświetl formaty.
23. Zmień ustawienie Dziesiętne dla Liczebności na wartość **2**.
24. Następnie kliknij przycisk **Kontynuuj**.
25. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		Count	Count	Count	Count	Count
Labor force status	Working	916	64	330	67	494
	Not Working	429	219	116	26	169
	Working / Not Working	2.14	.29	2.84	2.58	2.92

Rysunek 29. Przeliczana kategoria pokazująca iloraz sum pośrednich

Tabela obejmuje dwie sumy pośrednie oraz przeliczoną kategorię. Przeliczona kategoria pokazuje iloraz sum pośrednich, tak aby można było łatwo porównać grupy reprezentowane przez każdą sumę pośrednią. Iloraz osób pracujących jest znacznie niższy niż iloraz niepracujących wdów/wdowców w porównaniu z innymi grupami. Występuje również nieznacznie niższy iloraz zamężnych/żonatych, prawdopodobnie spowodowane jest to przerwaniem pracy przez małżonków w celu opieki nad dzieckiem.

Używanie przeliczonych kategorii do wyświetlania niewyczerpujących się sum pośrednich

Sumy pośrednie wyczerpują się. To znaczy, że wszystkie sumy pośrednie w tabeli obejmują wszystkie wartości powyżej lub poniżej ich pozycji w tabeli. Z drugiej strony, przeliczane kategorie nie wyczerpują się i umożliwiają zsumowanie różnych kategorii w tabeli.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.
3. W oknie konstruktora tabel przeciągnij zmienną *Myśli o sobie jak o osobie liberalnej lub konserwatywnej* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
4. Kliknij prawym przyciskiem myszy zmienną *Myśli o sobie jak o osobie liberalnej lub konserwatywnej* w panelu obszaru roboczego i z podręcznego menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
5. Wybierz z listy Wartości **3**.
6. Kliknij polecenie **Dodaj kategorię**, aby wyświetlić okno dialogowe Zdefiniuj przeliczoną kategorię.
7. W polu tekstowym Etykieta przeliczonej kategorii wpisz `Suma pośrednia liberałów`. Należy zwrócić uwagę, że przed tym tekstem znajdują się cztery spacje. Spacje te służą do zazębienia w tabeli wyniku.
8. Wybierz wartość **Skrajnie liberalny(a) 1** z Listy kategorii i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii.
9. Kliknij przycisk operatora plus (+) w oknie dialogowym (lub naciśnij klawisz + na klawiaturze).
10. Wybierz wartość **Liberalny(a) 2** z Listy kategorii i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii.
11. Kliknij przycisk operatora plus (+) w oknie dialogowym (lub naciśnij klawisz + na klawiaturze).

12. Wybierz wartość **Lekko liberalny(a) 3** z Listy kategorii i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii.
13. Kliknij przycisk **Kontynuuj**.

    Spowoduje to wstawienie wiersza zawierającego sumę pośrednią dla kategorii liberałów.
14. Wybierz z listy Wartości **7**.
15. Kliknij polecenie **Dodaj kategorię**, aby wyświetlić okno dialogowe Zdefiniuj przeliczoną kategorię.
16. W polu tekstowym Etykieta przeliczonej kategorii wpisz `Suma pośrednia konserwatystów`. Należy zwrócić uwagę, że przed tym tekstem znajdują się cztery spacje. Spacje te służą do zazębienia w tabeli wyniku.
17. Wybierz wartość **Lekko konserwatywny(a) 5** z Listy kategorii i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii.
18. Kliknij przycisk operatora plus (+) w oknie dialogowym (lub naciśnij klawisz + na klawiaturze).
19. Wybierz wartość **Konserwatywny(a) 6** z Listy kategorii i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii.
20. Kliknij przycisk operatora plus (+) w oknie dialogowym (lub naciśnij klawisz + na klawiaturze).
21. Wybierz wartość **Skrajnie konserwatywny(a) 7** z Listy kategorii i kliknij na przycisk strzałki, aby ją skopiować do pola tekstowego Wyrażenie dla przeliczanej kategorii.
22. Kliknij przycisk **Kontynuuj**.

    Spowoduje to wstawienie wiersza zawierającego sumę pośrednią dla kategorii konserwatystów.
23. Kliknij przycisk **Zastosuj** , a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Count
Think of self as liberal or conservative	Extremely liberal	64
	Liberal	357
	Slightly liberal	351
	Liberal Subtotal	772
	Moderate	986
	Slightly conservative	432
	Conservative	415
	Extremely conservative	86
	Conservative Subtotal	933

Rysunek 30. Przeliczane kategorie wyświetlające niewyczerpujące się sumy pośrednie

Tabela ta obejmuje dwie przeliczone kategorie, które nie obejmują wszystkich kategorii wyświetlonych w tabeli. Kategoria *Umiarkowany* nie zawiera się w żadnej przeliczonej kategorii. Nie można utworzyć tej samej tabeli z sumami pośrednimi, ponieważ sumy pośrednie wyczerpują się.

Tabele zmiennych zawierających wspólne kategorie

Ankiety często zawierają wiele pytań ze wspólnym zestawem możliwych odpowiedzi. Nasza przykładowa ankieta zawiera szereg zmiennych dotyczących zaufania do różnych instytucji oraz usług publicznych i prywatnych. Każda zmienna ma ten sam zestaw kategorii odpowiedzi: 1 = *Duże*, 2 = *Niewielkie* i 3 = *Prawie żadne*. Wykorzystując zestawianie zmiennych można wyświetlić tak powiązane zmienne w jednej tabeli, a w jej kolumnach wyświetlić wspólne kategorie odpowiedzi. Te funkcje są dostępne także, jeśli używasz przeliczanych kategorii, pod warunkiem, że jakiekolwiek etykiety i wyrażenia przeliczanych kategorii są takie same dla wszystkich zmiennych.

	A great deal	Only some	Hardly any
Confidence in banks & financial institutions	490	1068	306
Confidence in education	511	1055	315
Confidence in major companies	500	1078	243
Confidence in medicine	844	864	167
Confidence in press	176	878	808
Confidence in television	196	936	744

Rysunek 31. Tabela zmiennych zawierających wspólne kategorie

Uwaga: W poprzedniej wersji modułu Tabele użytkownika tabela taka zwana była „tabelą częstości”.

Przykładowy plik danych

W przykładach podanych w tym rozdziale wykorzystano plik danych *survey_sample.sav*. Więcej informacji zawiera temat “Pliki przykładowe” na stronie 77.

We wszystkich przedstawionych tutaj przykładach w oknach dialogowych wyświetlane są etykiety zmiennych, posortowane w porządku alfabetycznym. Opcje wyświetlania listy zmiennych ustawiane są na karcie Ogólne, w oknie dialogowym Opcje (menu Edycja, Opcje).

Tabela liczebności

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**

2. Na liście zmiennych w konstruktorze tabel kliknij zmienną *Zaufanie do banków...*, a następnie trzymając klawisz Shift kliknij zmienną *Zaufanie do telewizji*, aby zaznaczyć wszystkie zmienne związane z zaufaniem. (*Uwaga*: Przyjmujemy założenie, że na liście zmiennych etykiety zmiennych wyświetlane są w porządku alfabetycznym, a nie w kolejności, w jakiej występują w pliku).
3. Przeciągnij sześć zaznaczonych zmiennych do pola Wiersze w panelu obszaru roboczego.

   Powoduje to zestawienie zmiennych w wymiarze wierszowym. Domyślnie etykiety kategorii dla każdej zmiennej są również wyświetlane w wierszach, czego efektem jest bardzo długa, wąska tabela (6 zmiennych x 3 kategorie = 18 wierszy), jednak ponieważ wszystkie sześć zmiennych wykorzystuje te same zdefiniowane etykiety kategorii (etykiety wartości), takie etykiety kategorii można umieścić w wymiarze kolumn.
4. Z rozwijanej listy Pozycja kategorii wybierz opcję **Etykiety wierszy w kolumnach**.

   Tabela zawiera teraz tylko sześć wierszy, po jednym dla każdej z zestawionych zmiennych, a zdefiniowane kategorie stają się kolumnami tabeli.
5. Przed utworzeniem tabeli odznacz (kliknij) przycisk **Ukryj** dla elementu Pozycja w grupie Statystyki podsumowujące. Etykieta *Liczebność* jest właściwie zbędna.
6. Kliknij przycisk **OK**, aby utworzyć tabelę.

	A great deal	Only some	Hardly any
Confidence in banks & financial institutions	490	1068	306
Confidence in education	511	1055	315
Confidence in major companies	500	1078	243
Confidence in medicine	844	864	167
Confidence in press	176	878	808
Confidence in television	196	936	744

Rysunek 32. Tabela zestawionych zmiennych w wierszach z etykietami wspólnych kategorii w kolumnach

Zamiast wyświetlać zmienne w wierszach i kategorie w kolumnach można utworzyć tabelę ze zmiennymi zestawionymi w kolumnach i kategoriami wyświetlanymi w wierszach. Stanowi to lepsze rozwiązanie, w przypadku, kiedy jest więcej kategorii niż zmiennych, lecz w przedstawionym przykładzie jest więcej zmiennych.

Tabela procentów

W przypadku tabel ze zmiennymi zestawionymi w wierszach i kategoriami wyświetlanymi w kolumnach najistotniejsze (lub najłatwiejsze w zrozumieniu) są procenty w wierszach. (W przypadku tabeli ze zmiennymi zestawionymi w kolumnach i kategoriami wyświetlanymi w wierszach prawdopodobnie pożądane byłyby procenty w kolumnach).

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy dowolną zmienną dotyczącą zaufania na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.
3. Na liście Statystyki zaznacz pozycję **% z N w wierszu** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
4. Kliknij dowolną komórkę w wierszu *Liczebność* na liście Pokaż w tabeli, a następnie kliknij przycisk ze strzałką, aby przenieść ją z powrotem z listy Pokaż w tabeli na listę Statystyki.
5. Kliknij przycisk **Zastosuj do wszystkich**, aby zastosować zmianę statystyki podsumowującej do zmiennych zestawionych w tabeli.

   Uwaga: Jeśli podgląd tabeli nie wygląda, jak na rysunku, prawdopodobnie zamiast przycisku **Zastosuj do wszystkich** kliknięty został przycisk **Zastosuj do wybranej**, który powoduje zastosowanie nowej statystyki podsumowującej tylko do wybranej zmiennej. W tym przykładzie uzyskalibyśmy po dwie kolumny dla każdej kategorii: jedną z zastępczymi liczebnościami wyświetlanymi dla wszystkich pozostałych zmiennych i jedną z zastępczą liczebnością wyświetlaną dla wybranej zmiennej. Dokładnie taka tabela zostałaby utworzona, jednak w tym przykładzie *nie* o to chodziło.
6. Kliknij przycisk **OK**, aby utworzyć tabelę.

	A great deal	Only some	Hardly any
Confidence in banks & financial institutions	26.3%	57.3%	16.4%
Confidence in education	27.2%	56.1%	16.7%
Confidence in major companies	27.5%	59.2%	13.3%
Confidence in medicine	45.0%	46.1%	8.9%
Confidence in press	9.5%	47.2%	43.4%
Confidence in television	10.4%	49.9%	39.7%

Rysunek 33. Tabela procentów w wierszach dla zmiennych zestawionych w wierszach, kategorie wyświetlane w kolumnach

Uwaga: W tabeli zmiennych zawierających wspólne kategorie można umieścić dowolną liczbę statystyk podsumowujących. Dla uproszczenia w każdym przykładzie pokazywana jest tylko jedna statystyka.

Element sterujący Podsumowania i kategorie

Tabelę zawierającą kategorie w przeciwnym wymiarze niż zmienne można utworzyć tylko, jeśli wszystkie zmienne tabeli posiadają takie same kategorie, wyświetlane w identycznym porządku. Obejmuje to podsumowania, sumy pośrednie i inne dokonane korekty kategorii. Oznacza to, że wszelkie modyfikacje dokonane w oknie dialogowym Kategorie i podsumowania, muszą obejmować wszystkie zmienne tabeli, wykorzystujące wspólne kategorie.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy pierwszą zmienną dotyczącą zaufania na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Kategorie i podsumowania**.

3. W oknie dialogowym Kategorie i podsumowania zaznacz opcję **Podsumowanie ogółem**, a następnie kliknij przycisk **Zastosuj**.

   Pierwsza rzecz rzucająca się w oczy to fakt, że etykiety kategorii zostały przeniesione z powrotem z kolumn do wierszy. Można również zauważyć, że element sterujący Pozycja kategorii jest teraz nieaktywny. Wynika to z faktu, że zmienne nie wykorzystują już identycznego zestawu „kategorii". Jedna ze zmiennych posiada teraz kategorię sumy.
4. Kliknij prawym przyciskiem myszy dowolną zmienną dotyczącą zaufania w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Zaznacz zmienne w wierszach** lub naciśnij klawisz Ctrl i kliknij każdą ze zmiennych zestawionych w panelu obszaru roboczego w celu ich zaznaczenia (niezbędne może być przewinięcie listy zmiennych lub maksymalizacja okna konstruktora tabel).
5. W grupie Definiuj kliknij pozycję **Kategorie i podsumowania**.
6. Jeśli w oknie dialogowym Kategorie i podsumowania opcja **Suma** nie jest jeszcze zaznaczona, zaznacz ją, a następnie kliknij przycisk **Zastosuj**.
7. Rozwijana lista Pozycja kategorii powinna być znowu aktywna, gdyż wszystkie zmienne posiadają teraz dodatkową kategorię sumy, więc można wybrać opcję **Etykiety wierszy w kolumnach**.
8. Kliknij przycisk **OK**, aby utworzyć tabelę.

	A great deal	Only some	Hardly any	Total
Confidence in banks & financial institutions	26.3%	57.3%	16.4%	100.0%
Confidence in education	27.2%	56.1%	16.7%	100.0%
Confidence in major companies	27.5%	59.2%	13.3%	100.0%
Confidence in medicine	45.0%	46.1%	8.9%	100.0%
Confidence in press	9.5%	47.2%	43.4%	100.0%
Confidence in television	10.4%	49.9%	39.7%	100.0%

Rysunek 34. Tabela procentów w wierszach dla zmiennych zestawionych w wierszach, kategorie i sumy wyświetlane w kolumnach

Zagnieżdżanie zmiennych w tabelach zawierających wspólne kategorie

W tabelach zagnieżdżonych, jeśli etykiety kategorii mają być wyświetlane w przeciwnym wymiarze, zmienne zestawione zawierające wspólne kategorie muszą znajdować się na najwyższym poziomie zagnieżdżenia swojego wymiaru.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Przeciągnij zmienną *Płeć* z listy zmiennych do lewej krawędzi pola Wiersze.

   Zmienne zestawione zawierające wspólne kategorie są teraz zagnieżdżone w kategoriach płci na podglądzie tabeli.
3. Następnie przeciągnij zmienną *Płeć* do prawej strony zestawionych zmiennych dotyczących zaufania na podglądzie tabeli.

Również w tym przypadku etykiety kategorii powróciły do wymiaru wierszowego, a element sterujący Pozycja kategorii jest nieaktywny. Obecnie istnieje jedna zmienna zestawiona z zagnieżdżoną w niej jedną zmienną *Płeć*, natomiast pozostałe zestawione zmienne nie zawierają zmiennych zagnieżdżonych. Można dodać zmienną *Płeć* jako zmienną zagnieżdżoną do każdej ze zmiennych zestawionych, lecz przeniesienie wówczas etykiet wierszy do kolumn spowoduje wyświetlanie w kolumnach etykiet kategorii dla zmiennej *Płeć*, a nie etykiet kategorii dla zestawionych zmiennych zawierających wspólne kategorie. Wynika to z faktu, że zmienna *Płeć* będzie wówczas najbardziej zagnieżdżoną zmienną, a zmiana położenia kategorii zawsze ma zastosowanie do takiej zmiennej.

Statystyki podsumowujące

Statystyki podsumowujące obejmują wszystko — od prostych liczebności zmiennych kategorialnych po miary rozproszenia, takie jak błąd standardowy średniej dla zmiennych ilościowych. *Nie* obejmują one testów istotności dostępnych na karcie Testy w oknie dialogowym Tabele użytkownika. Więcej informacji zawiera temat “Testy” na stronie 53.

Statystyki podsumowujące dla zmiennych kategorialnych i zestawów wielokrotnych odpowiedzi obejmują liczebności oraz szereg różnych obliczeń procentowych, w tym:

- procent w wierszu,
- procent w kolumnie,
- procent w podtabeli,
- procent w tabeli,
- procent N ważnych.

Oprócz statystyk podsumowujących dostępnych w przypadku zmiennych kategorialnych, dla zmiennych ilościowych oraz statystyk podsumowania wybieranych przez użytkownika dla zmiennych jakościowych dostępne są również następujące obliczenia:

- Średnia
- Mediana
- Percentyle
- Suma
- Odchylenie standardowe
- Przedział
- Wartości minimalne i maksymalne.

Dodatkowe statystyki podsumowujące dostępne są w przypadku zestawów wielokrotnych odpowiedzi. Dostępna jest również pełna lista statystyk podsumowujących. Więcej informacji zawiera temat “Statystyki podsumowujące” na stronie 6.

Przykładowy plik danych

W przykładach podanych w tym rozdziale wykorzystano plik danych *survey_sample.sav*. Więcej informacji zawiera temat “Pliki przykładowe” na stronie 77.

We wszystkich przedstawionych tutaj przykładach w oknach dialogowych wyświetlane są etykiety zmiennych, posortowane w porządku alfabetycznym. Opcje wyświetlania listy zmiennych ustawiane są na karcie Ogólne, w oknie dialogowym Opcje (menu Edycja, Opcje).

Zmienna podsumowująca statystyki źródłowej

To, jakie statystyki podsumowujące są dostępne, zależy od poziomu pomiaru zmiennej podsumowującej statystyki źródłowej. Źródło statystyk podsumowujących (zmienna, na podstawie której obliczane są statystyki podsumowujące) jest określone przez:

- **Poziom pomiaru**. Jeśli tabela (lub sekcja tabeli w przypadku tabeli zestawionej) zawiera zmienną ilościową, statystyki podsumowujące obliczane są na podstawie tej zmiennej.
- **Kolejność wybierania zmiennych**. Domyślny wymiar statystyki źródłowej (wiersz lub kolumna) w przypadku zmiennych kategorialnych zależy od kolejności przeciągania zmiennych do panelu obszaru roboczego. Jeśli najpierw, na przykład zmienna zostanie przeciągnięta do pola wierszy, wymiar wierszowy będzie domyślnym wymiarem statystyki źródłowej.
- **Zagnieżdżanie**. W przypadku zmiennych kategorialnych statystyki podsumowujące są obliczane na podstawie najbardziej wewnętrznej zmiennej w wymiarze statystyki źródłowej.

Tabela zestawiona może zawierać wiele zmiennych podsumowujących statystyki źródłowej (zarówno ilościowych, jak i kategorialnych), jednak każda sekcja tabeli zawiera tylko jedną taką zmienną.

Źródło statystyk podsumowujących dla zmiennych kategorialnych

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
3. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**. (ponieważ jest to jedyna zmienna w tabeli, będzie ona zmienną statystyki źródłowej).
4. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję *% z N w kolumnie* i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
5. Kliknij przycisk **Zastosuj do wybranej**.
6. W konstruktorze tabel przeciągnij zmienną *Informacje czerpie z Internetu* na prawo od zmiennej *Kategoria wieku* na panelu obszaru roboczego.
7. Ponownie kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* na panelu obszaru roboczego. Pozycja **Statystyki podsumowujące** w menu kontekstowym jest teraz wyłączona, ponieważ zmienna *Kategoria wieku* nie jest najbardziej zagnieżdżoną zmienną w wymiarze statystyki źródłowej.
8. Kliknij prawym przyciskiem myszy zmienną *Informacje czerpie z Internetu* na panelu obszaru roboczego. Pozycja **Statystyki podsumowujące** jest włączona, ponieważ kliknięta zmienna, będąc najbardziej zagnieżdżoną zmienną wymiaru statystyki źródłowej, jest zmienną statystyki źródłowej. (Ponieważ tabela ma tylko jeden wymiar (wiersze), właśnie on jest wymiarem statystyki źródłowej).
9. Przeciągnij zmienną *Informacje czerpie z Internetu* z pola Wiersze na panelu obszaru roboczego do pola Kolumny.
10. Ponownie prawym przyciskiem myszy kliknij zmienną *Informacje czerpie z Internetu* na panelu obszaru roboczego. Pozycja **Statystyki podsumowujące** w podręcznym menu kontekstowym jest teraz wyłączona, ponieważ zmienna ta nie jest już wymiarem statystyki źródłowej.

Zmienna *Kategoria wieku* ponownie jest zmienną statystyki źródłowej, ponieważ domyślnym wymiarem statystyki źródłowej dla zmiennych kategorialnych jest pierwszy wymiar, w którym umieszczono zmienne podczas tworzenia tabeli. W tym przykładzie pierwszą czynnością było umieszczenie zmiennych w wymiarze wierszowym. A zatem wymiar wierszowy jest domyślnym wymiarem statystyki źródłowej; ponieważ *Kategoria wieku* jest obecnie jedyną zmienną w tym wymiarze, właśnie ona będzie zmienną statystyki źródłowej.

Źródło statystyk podsumowujących dla zmiennych ilościowych

1. Przeciągnij zmienną ilościową *Liczba godzin oglądania telewizji dziennie* na lewo od zmiennej *Kategoria wieku* w polu Wiersze panelu obszaru roboczego.

   Pierwszą rzeczą wartą odnotowania jest fakt, że podsumowania *Liczebność* i *% z N w kolumnie* zostały zastąpione podsumowaniem *Średnia*, a po kliknięciu prawym przyciskiem myszy zmiennej *Liczba godzin oglądania telewizji dziennie* na panelu obszaru roboczego okazuje się, że właśnie ona jest zmienną podsumowującą statystyki źródłowej. W przypadku tabel ze zmienną ilościową, właśnie zmienna ilościowa jest zawsze zmienną statystyki źródłowej, niezależnie od jej poziomu zagnieżdżenia i wymiaru, zaś domyślną statystyką podsumowującą dla zmiennych ilościowych jest średnia.
2. Przeciągnij zmienną *Liczba godzin oglądania telewizji dziennie* z pola Wiersze do pola Kolumny nad zmienną *Informacje czerpie z Internetu*.
3. Kliknij prawym przyciskiem myszy zmienną *Liczba godzin oglądania telewizji dziennie* i z menu podręcznego wybierz polecenie **Statystyki podsumowujące**. (Jest ona nadal zmienną statystyki źródłowej, mimo że została przeniesiona do innego wymiaru).
4. W oknie dialogowym Statystyki podsumowujące, na liście Pokaż w tabeli, kliknij komórkę **Format** i z listy rozwijanej Format wybierz pozycję **nnnn** (być może konieczne będzie przewinięcie listy w celu odnalezienia tej pozycji).
5. W komórce Dziesiętne wpisz 2.
6. Kliknij przycisk **Zastosuj do wybranej**.

Na podglądzie tabeli na panelu obszaru roboczego widać teraz, że wartości średniej będą wyświetlane z dwoma miejscami dziesiętnymi.

7. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Hours per day watching TV	
		Get news from internet	
		No	Yes
		Mean	Mean
Age category	Less than 25	3.54	2.12
	25 to 34	3.42	2.14
	35 to 44	3.00	2.01
	45 to 54	2.83	2.06
	55 to 64	3.24	2.37
	65 or older	3.82	2.33

Rysunek 35. Zmienne ilościowe podsumowane w ramach zmiennych jakościowych umieszczonych w tabeli krzyżowej

Zmienne zestawione

Ponieważ tabela zestawiona może zawierać wiele zmiennych statystyki źródłowej i ponieważ możliwe jest określenie różnych statystyk podsumowujących dla każdej z tych zmiennych statystyki źródłowej, przy określaniu statystyk podsumowujących w tabelach zestawionych należy wziąć pod uwagę pewne szczególne uwarunkowania.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.
3. Kliknij zmienną *Informacje czerpie z czasopism* na liście zmiennych, a następnie, trzymając wciśnięty klawisz Shift, kliknij zmienną *Informacje czerpie z telewizji*, aby zaznaczyć wszystkie zmienne dotyczące wiadomości. (*Uwaga*: Przyjmujemy założenie, że na liście zmiennych etykiety zmiennych wyświetlane są w porządku alfabetycznym, a nie w kolejności, w jakiej występują w pliku).
4. Przeciągnij pięć zmiennych dotyczących wiadomości do pola Wiersze na panelu obszaru roboczego.

   Pięć zmiennych dotyczących wiadomości jest teraz zestawionych w wymiarze wierszowym.
5. Kliknij zmienną *Informacje czerpie z Internetu* na panelu obszaru roboczego; teraz zaznaczona będzie tylko ta zmienna.
6. Następnie kliknij prawym przyciskiem myszy zmienną *Informacje czerpie z Internetu* i wybierz polecenie **Statystyki podsumowujące** z menu kontekstowego.
7. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję *% z N w kolumnie* i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli. (Zaznaczone statystyki można przenieść z listy Statystyki do listy Pokaż w tabeli klikając przycisk ze strzałką lub przeciągając je).
8. Następnie kliknij przycisk **Zastosuj do wybranej**.

   Została dodana kolumna dla procentu w kolumnie, ale na podglądzie tabeli na panelu obszaru roboczego widać, że procent w kolumnie będzie wyświetlany tylko dla jednej zmiennej. Wynika to z faktu, że w tabeli zestawionej istnieje wiele zmiennych statystyki źródłowej, a każda z nich może mieć inne statystyki podsumowujące. Jednak w tym przykładzie mają być wyświetlone te same statystyki podsumowujące dla wszystkich zmiennych.
9. Kliknij prawym przyciskiem myszy zmienną *Informacje czerpie z gazet codziennych* w panelu obszaru roboczego i z menu kontekstowego wybierz polecenie **Statystyki podsumowujące**.
10. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję *% z N w kolumnie* i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
11. Następnie kliknij przycisk **Zastosuj do wszystkich**.

Teraz na podglądzie tabeli widać, że procenty w kolumnach będą wyświetlane dla wszystkich zmiennych zestawionych.

Statystyki podsumowania wybierane przez użytkownika dla zmiennych kategorialnych

W przypadku jakościowych zmiennych statystyki źródłowej możliwe jest uwzględnienie statystyk podsumowania wybieranych przez użytkownika, które różnią się od statystyk wyświetlanych dla kategorii zmiennej. Na przykład dla zmiennej porządkowej można wyświetlić procenty dla poszczególnych kategorii oraz średnią lub medianę jako statystykę podsumowania wybieraną przez użytkownika.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.
3. Kliknij zmienną *Zaufanie do prasy* na liście zmiennych, a następnie, trzymając naciśnięty klawisz Ctrl, kliknij zmienną *Zaufanie do telewizji*, aby zaznaczyć obie zmienne.
4. Przeciągnij te dwie zmienne do pola Wiersze na panelu obszaru roboczego. Spowoduje to zestawienie obu zmiennych w wymiarze wierszowym.
5. Kliknij prawym przyciskiem myszy zmienną w panelu obszaru roboczego i z menu kontekstowego wybierz polecenie **Zaznacz zmienne w wierszach** (być może obie zmienne są już zaznaczone, ale dla pewności należy wybrać to polecenie).
6. Ponownie kliknij zmienną prawym przyciskiem myszy i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
7. W oknie dialogowym Kategorie i podsumowania kliknij (zaznacz) opcję **Podsumowanie ogółem**, a następnie kliknij przycisk **Zastosuj**.

   Na podglądzie tabeli na panelu obszaru roboczego widoczny będzie teraz wiersz podsumowania ogółem dla obu zmiennych. Aby wyświetlić statystyki podsumowania wybierane przez użytkownika, należy określić dla tabeli podsumowania ogółem i sumy pośrednie.
8. Kliknij prawym przyciskiem myszy dowolną zmienną w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
9. W oknie dialogowym Statystyki podsumowujące zaznacz pozycję *Liczebność* na liście Pokaż w tabeli, a następnie kliknij przycisk ze strzałką, aby przenieść ją z powrotem z listy Pokaż w tabeli do listy Statystyki.
10. Na liście Statystyki zaznacz pozycję *% z N w kolumnie* i kliknij przycisk ze strzałką, aby przenieść ją do listy Pokaż w tabeli.
11. Kliknij (zaznacz) opcję **Statystyki użytkownika dla podsumowań ogółem i sum pośrednich**.
12. Kliknij pozycję *Liczebność* na liście Pokaż w tabeli dotyczącej podsumowań użytkownika, a następnie kliknij przycisk ze strzałką, aby przenieść ją z powrotem z listy Pokaż w tabeli na listę Statystyki podsumowań użytkownika.
13. Na liście Statystyki podsumowań użytkownika zaznacz pozycję *Średnia* i kliknij przycisk ze strzałką, aby przenieść ją na listę Pokaż w tabeli.
14. Kliknij komórkę **Format** dla średniej na liście Pokaż w tabeli i z rozwijanej listy formatów wybierz pozycję **nnnn** (być może konieczne będzie przewinięcie listy w celu odnalezienia tej pozycji).
15. W komórce Dziesiętne wpisz 2.
16. Kliknij przycisk **Zastosuj do wszystkich**, aby zastosować te ustawienia do obu zmiennych w tabeli.

    Została dodana nowa kolumna przeznaczona na statystyki podsumowania wybierane przez użytkownika, jednak prawdopodobnie nie jest to optymalne rozwiązanie, ponieważ — jak wynika z podglądu na panelu obszaru roboczego — wynikowa tabela zawierać będzie wiele pustych komórek.
17. W konstruktorze tabel, w grupie Statystyki podsumowujące, wybierz opcję **Wiersze** z listy rozwijanej Pozycja.

    Spowoduje to przeniesienie wszystkich statystyk podsumowujących do wymiaru wierszowego, dzięki czemu wszystkie statystyki podsumowujące będą wyświetlane w jednej kolumnie tabeli.
18. Kliknij przycisk **OK**, aby utworzyć tabelę.

Confidence in press	A great deal	Column N %	9.5%
	Only some	Column N %	47.2%
	Hardly any	Column N %	43.4%
	Total	Mean	2.34
Confidence in television	A great deal	Column N %	10.4%
	Only some	Column N %	49.9%
	Hardly any	Column N %	39.7%
	Total	Mean	2.29

Rysunek 36. Zmienne kategorialne ze statystykami podsumowania wybieranymi przez użytkownika

Wyświetlanie wartości kategorii

W poprzedniej tabeli występuje tylko jeden drobny problem — interpretacja wartości średniej może sprawiać trudności bez znajomości wartości kategorii, na podstawie których została obliczona. Czy średnia równa 2,34 przypada pomiędzy kategorią *Bardzo duże* a *Częściowe*, czy też może pomiędzy kategorią *Częściowe* a *Prawie żadne*?

Pomimo tego, iż nie można rozwiązać tego problemu bezpośrednio w oknie dialogowym Tabele użytkownika, można znaleźć dla niego rozwiązanie bardziej ogólne.

1. Z menu wybierz kolejno następujące pozycje:

   Edycja > **Opcje...**
2. W oknie dialogowym Opcje kliknij kartę **Etykietowanie**.
3. W grupie Opisywanie tabel przestawnych wybierz pozycję **Wartości i etykiety** z listy rozwijanej **Wartości zmiennych w opisach pokazywane są jako**.
4. Kliknij przycisk **OK**, aby zapisać to ustawienie.
5. Otwórz kreatora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika) i kliknij przycisk **OK**, aby ponownie utworzyć tabelę.

Confidence in press	1 A great deal	Column N %	9.5%
	2 Only some	Column N %	47.2%
	3 Hardly any	Column N %	43.4%
	Total	Mean	2.34
Confidence in television	1 A great deal	Column N %	10.4%
	2 Only some	Column N %	49.9%
	3 Hardly any	Column N %	39.7%
	Total	Mean	2.29

Rysunek 37. Wartości i etykiety wyświetlane dla kategorii zmiennej

Wartości kategorii pozwalają zorientować się, że wartość średniej 2,34 przypada pomiędzy kategorią *Częściowe* a *Prawie żadne*. Wyświetlanie wartości kategorii w tabeli znacznie ułatwia interpretację wartości statystyk podsumowania wybieranych przez użytkownika, takich jak średnia.

To ustawienie wyświetlania jest ustawieniem globalnym i ma wpływ na wszystkie tabele przestawne generowane przez wszystkie procedury, a ponadto jest zachowywane we wszystkich sesjach, dopóki nie zostanie zmienione. Aby przywrócić ustawienie nakazujące wyświetlanie tylko etykiet wartości:

6. Z menu wybierz kolejno następujące pozycje:

   Edycja > **Opcje...**
7. W oknie dialogowym Opcje kliknij kartę **Etykietowanie**.
8. W grupie Opisywanie tabel przestawnych wybierz pozycję **Etykiety** z listy rozwijanej **Wartości zmiennych w opisach pokazywane są jako**.
9. Kliknij przycisk **OK**, aby zapisać to ustawienie.

Podsumowywanie zmiennych ilościowych

Podsumowywanie zmiennych ilościowych

Dla zmiennych ilościowych dostępnych jest szeroki zakres statystyk podsumowujących. Oprócz liczebności i procentów dostępnych dla zmiennych kategorialnych, do statystyk podsumowujących dla zmiennych ilościowych należą:

- Średnia
- Mediana
- Percentyle
- Suma
- Odchylenie standardowe
- Przedział
- Wartości minimalne i maksymalne.

Więcej informacji zawiera temat "Statystyki podsumowujące dla zmiennych ilościowych i podsumowania użytkownika dla zmiennych jakościowych" na stronie 9.

Przykładowy plik danych

W przykładach podanych w tym rozdziale wykorzystano plik danych *survey_sample.sav*. Więcej informacji zawiera temat "Pliki przykładowe" na stronie 77.

We wszystkich przedstawionych tutaj przykładach w oknach dialogowych wyświetlane są etykiety zmiennych, posortowane w porządku alfabetycznym. Opcje wyświetlania listy zmiennych są ustawiane na karcie Ogólne, w oknie dialogowym Opcje (menu Edycja, Opcje).

Zestawione zmienne ilościowe

W jednej tabeli można podsumować wiele zmiennych ilościowych zestawiając je ze sobą.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. W konstruktorze tabel kliknij na liście zmiennych zmienną *Wiek respondenta*, naciśnij klawisz Ctrl i kliknij zmienną *Lata nauki szkolnej respondenta*, a następnie *Liczba godzin oglądania telewizji dziennie*, aby zaznaczyć wszystkie trzy zmienne.
3. Przeciągnij trzy zaznaczone zmienne do pola Wiersze panelu obszaru roboczego.

   Trzy zmienne są zestawione ze sobą w wymiarze wierszowym. Ponieważ wszystkie trzy są zmiennymi ilościowymi, nie są pokazywane żadne kategorie, a domyślną statystyką podsumowującą jest średnia.
4. Kliknij przycisk **OK**, aby utworzyć tabelę.

	Mean
Age of respondent	46
Highest year of school completed	13
Hours per day watching TV	3

Rysunek 38. Tabela wartości średnich zestawionych zmiennych ilościowych

Wielokrotne statystyki podsumowujące

Domyślnie dla zmiennych ilościowych wyświetlana jest średnia. Można jednak wybrać dla nich inną statystykę podsumowującą i wyświetlić więcej niż jedną statystykę.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy jedną z trzech zmiennych ilościowych na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.

3. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję *Mediana* i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli. (Zaznaczone statystyki można przenieść z listy Statystyki do listy Pokaż w tabeli klikając przycisk ze strzałką lub przeciągając je).
4. Na liście Pokaż w tabeli kliknij komórkę **Format** dla mediany i z rozwijanej listy formatów wybierz format **nnnn**.
5. W komórce Dziesiętne wpisz 1.
6. Wprowadź identyczne zmiany dla średniej na liście Pokaż w tabeli.
7. Kliknij przycisk **Zastosuj do wszystkich**, aby zastosować zmiany do wszystkich trzech zmiennych ilościowych.
8. Kliknij przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

	Mean	Median
Age of respondent	45.6	42.0
Highest year of school completed	13.3	13.0
Hours per day watching TV	2.9	2.0

Rysunek 39. Średnia i mediana wyświetlana w tabeli zestawionych zmiennych ilościowych

Liczebność, N ważnych i Braki danych

Często zachodzi potrzeba wyświetlenia liczby obserwacji wykorzystanych do wyliczenia statystyk podsumowujących, takich jak średnia. Można założyć (nie bez powodu), że informację taką zawiera statystyka podsumowująca *Liczebność*. Jeśli jednak występują braki danych, nie stanowi ona dokładnej liczby obserwacji służącej do wyliczenia innych statystyk. Dokładną liczbę obserwacji służących do wyliczenia innych statystyk zawiera statystyka *N ważnych*.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy jedną z trzech zmiennych ilościowych na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.
3. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję **Liczebność** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
4. Na liście Statystyki zaznacz następnie pozycję **N ważnych** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
5. Kliknij przycisk **Zastosuj do wszystkich**, aby zastosować zmiany do wszystkich trzech zmiennych ilościowych.
6. Kliknij przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

	Mean	Median	Count	Valid N
Age of respondent	45.6	42.0	2832	2828
Highest year of school completed	13.3	13.0	2832	2820
Hours per day watching TV	2.9	2.0	2832	2337

Rysunek 40. Liczebność a N ważnych

Dla wszystkich trzech zmiennych *Liczebność* jest taka sama: 2832. Nie przypadkowo — jest to całkowita liczba obserwacji zawartych w pliku danych. Ponieważ zmienne ilościowe nie są zagnieżdżone w żadnych zmiennych kategorialnych, *Liczebność* jest po prostu całkowitą liczbą obserwacji występujących w pliku danych.

Jednak statystyka *N ważnych* jest inna dla każdej zmiennej i znacznie odbiega od statystyki *Liczebność* dla zmiennej *Liczba godzin oglądania telewizji dziennie*. Wynika to z faktu, że w przypadku tej zmiennej występuje duża ilość **braków danych**, tzn. obserwacji bez zarejestrowanych wartości dla danej zmiennej lub z wartościami zdefiniowanymi jako braki danych (np. kod 99 reprezentujący wartość *nie dotyczy* w przypadku ciąży u mężczyzn).

7. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
8. Kliknij prawym przyciskiem myszy jedną z trzech zmiennych ilościowych na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.

9. W oknie dialogowym Statystyki podsumowujące, na liście Pokaż w tabeli, zaznacz pozycję **N ważnych** i kliknij przycisk ze strzałką, aby przenieść ją z powrotem z listy Pokaż w tabeli do listy Statystyki.
10. Na liście Pokaż w tabeli, zaznacz pozycję **Liczebność** i kliknij przycisk ze strzałką, aby przenieść ją z powrotem z listy Pokaż w tabeli do listy Statystyki.
11. Na liście Statystyki zaznacz pozycję **Brakujące** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
12. Kliknij przycisk **Zastosuj do wszystkich**, aby zastosować zmiany do wszystkich trzech zmiennych ilościowych.
13. Kliknij przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

	Mean	Median	Missing
Age of respondent	45.6	42.0	4
Highest year of school completed	13.3	13.0	12
Hours per day watching TV	2.9	2.0	495

Rysunek 41. Liczba braków danych wyświetlanych w tabeli statystyk podsumowujących dla zmiennych ilościowych

W tabeli wyświetlana jest teraz liczba braków danych dla każdej zmiennej ilościowej. Wyraźnie widać, że zmienna *Liczba godzin oglądania telewizji dziennie* zawiera dużą liczbę braków danych, natomiast pozostałe dwie zmienne zawierają ich niewiele. Fakt ten należy uwzględnić przy ocenie wiarygodności statystyk podsumowujących dla tej zmiennej.

Różne podsumowania dla różnych zmiennych

Oprócz wyświetlenia wielu statystyk podsumowujących można wyświetlić różne statystyki podsumowujące dla różnych zmiennych w zestawionej tabeli. Na przykład z poprzedniej tabeli wynikało, że tylko jedna z trzech zmiennych zawiera dużą ilość braków danych, więc celowe może być pokazanie liczby braków danych tylko dla tej zmiennej.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij zmienną *Wiek respondenta* na podglądzie tabeli w panelu obszaru roboczego, a następnie naciśnij klawisz Ctrl i kliknij zmienną *Lata nauki szkolnej respondenta*, aby zaznaczyć obie zmienne.
3. Kliknij prawym przyciskiem myszy jedną z dwóch zaznaczonych zmiennych i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
4. W oknie dialogowym Statystyki podsumowujące, na liście Pokaż w tabeli, zaznacz pozycję **Brakujące** i kliknij przycisk ze strzałką, aby przenieść ją z powrotem z listy Pokaż w tabeli do listy Statystyki.
5. Kliknij przycisk **Zastosuj do wybranej**, aby zastosować zmianę tylko do dwóch zaznaczonych zmiennych.

   Symbole zastępcze w komórkach danych tabeli wskazują, że liczba braków danych będzie wyświetlana wyłącznie dla zmiennej *Liczba godzin oglądania telewizji dziennie*.
6. Kliknij przycisk **OK**, aby utworzyć tabelę.

	Mean	Median	Missing
Age of respondent	45.6	42.0	
Highest year of school completed	13.3	13.0	
Hours per day watching TV	2.9	2.0	495

Rysunek 42. Tabela z różnymi statystykami podsumowującymi dla różnych zmiennych

Chociaż ta tabela zawiera żądane informacje, jej układ może utrudniać interpretację wyników. Osoba odczytująca tabelę mogłaby pomyśleć, że puste komórki w kolumnie *Brakujące* oznaczają zerową ilość braków danych dla tych zmiennych.

7. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
8. W grupie Statystyki podsumowujące w konstruktorze tabel, z rozwijanej listy Pozycja wybierz **Wiersze**.

9. Kliknij przycisk **OK**, aby utworzyć tabelę.

Age of respondent	Mean	45.6
	Median	42.0
Highest year of school completed	Mean	13.3
	Median	13.0
Hours per day watching TV	Mean	2.9
	Median	2.0
	Missing	495

Rysunek 43. Statystyki podsumowujące i zmienne wyświetlane w wymiarze wierszy

Teraz widać wyraźnie, że tabela podaje liczbę braków danych tylko dla jednej zmiennej.

Podsumowania grupowe w kategoriach

Zmienne kategorialne można wykorzystać jako zmienne grupujące do wyświetlenia podsumowań zmiennych ilościowych w grupach zdefiniowanych według kategorii zmiennej jakościowej.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Przeciągnij zmienną *Płeć* z listy zmiennych do pola Kolumny w panelu obszaru roboczego.

   Po kliknięciu prawym przyciskiem myszy zmiennej *Płeć* na podglądzie tabeli w panelu obszaru roboczego widać, że opcja **Statystyki podsumowujące** jest nieaktywna w podręcznym menu kontekstowym. Wynika to z faktu, że w tabeli ze zmiennymi ilościowymi są one zawsze zmiennymi statystyki źródłowej.
3. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Gender	
		Male	Female
Age of respondent	Mean	44.6	46.3
	Median	42.0	43.0
Highest year of school completed	Mean	13.4	13.2
	Median	13.0	13.0
Hours per day watching TV	Mean	2.8	2.9
	Median	2.0	2.0
	Missing	213	282

Rysunek 44. Podsumowania zmiennych ilościowych pogrupowane przy wykorzystaniu zmiennej jakościowej z kolumny

Tabela taka ułatwia porównywanie wartości średnich (średnia i mediana) dla mężczyzn i kobiet oraz wyraźnie pokazuje, że różnica między nimi jest niewielka — co nie jest może bardzo interesujące, ale może się czasami przydać.

Wiele zmiennych grupujących

Grupy można dodatkowo podzielić zagnieżdżając i / lub wykorzystując jakościowe zmienne grupujące występujące w wierszach i kolumnach.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Przeciągnij zmienną *Informacje czerpie z Internetu* z listy zmiennych do lewej krawędzi pola Wiersze w panelu obszaru roboczego. Umieść ją tak, aby wszystkie trzy zmienne ilościowe były w niej zagnieżdżone.

   Chociaż może się zdarzyć, że sytuacja pokazana w drugim przykładzie powyżej będzie pożądana, w tym przypadku tak nie jest.
3. Kliknij przycisk **OK**, aby utworzyć tabelę.

				Gender	
				Male	Female
Get news from internet	No	Age of respondent	Mean	47.0	48.8
			Median	45.0	46.0
		Highest year of school completed	Mean	13.4	13.1
			Median	13.0	12.0
		Hours per day watching TV	Mean	3.2	3.4
			Median	2.0	3.0
			Missing	213	282
	Yes	Age of respondent	Mean	38.7	41.1
			Median	35.0	38.0
		Highest year of school completed	Mean	13.2	13.3
			Median	13.0	13.0
		Hours per day watching TV	Mean	2.1	2.1
			Median	2.0	2.0
			Missing	0	0

Rysunek 45. Podsumowania zmiennych ilościowych pogrupowane według zmiennych jakościowych w wierszach i kolumnach

Zagnieżdżanie zmiennych kategorialnych w zmiennych ilościowych

Chociaż powyższa tabela może zawierać żądane informacje, jej układ może utrudniać interpretację wyników. Można na przykład porównać średni wiek mężczyzn korzystających z Internetu jako źródła wiadomości i tych, którzy z niego nie korzystają — jednak byłoby to łatwiejsze, gdyby wartości znajdowały się obok siebie, a nie były oddzielone. Zamiana położenia dwóch zmiennych w wierszach, zagnieżdżenie jakościowej zmiennej grupującej w trzech zmiennych ilościowych może poprawić układ tabeli. W przypadku zmiennych ilościowych poziom zagnieżdżenia nie ma wpływu na zmienną statystyki źródłowej. Zmienna ilościowa jest zawsze zmienną statystyki źródłowej, bez względu na poziom zagnieżdżenia.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Na podglądzie tabeli w panelu obszaru roboczego kliknij zmienną *Wiek respondenta*, naciśnij klawisz Ctrl i kliknij zmienną *Lata nauki szkolnej respondenta*, a następnie *Liczba godzin oglądania telewizji dziennie*, aby zaznaczyć wszystkie trzy zmienne ilościowe.
3. Przeciągnij trzy zmienne ilościowe do lewej krawędzi pola Wiersze, zagnieżdżając w każdej z nich zmienną kategorialną *Informacje czerpie z Internetu*.
4. Kliknij przycisk **OK**, aby utworzyć tabelę.

				Gender	
				Male	Female
Age of respondent	Get news from internet	No	Mean	47.0	48.8
			Median	45.0	46.0
		Yes	Mean	38.7	41.1
			Median	35.0	38.0
Highest year of school completed	Get news from internet	No	Mean	13.4	13.1
			Median	13.0	12.0
		Yes	Mean	13.2	13.3
			Median	13.0	13.0
Hours per day watching TV	Get news from internet	No	Mean	3.2	3.4
			Median	2.0	3.0
			Missing	213	282
		Yes	Mean	2.1	2.1
			Median	2.0	2.0
			Missing	0	0

Rysunek 46. Zmienna kategorialna w wierszu zagnieżdżona w zestawionych zmiennych ilościowych

Porządek zagnieżdżenia zależy od relacji lub porównań, które mają zostać uwidocznione w tabeli. Zmiana porządku zagnieżdżenia zmiennych ilościowych nie zmienia wartości statystyk podsumowujących; zmienia jedynie ich względne położenie w tabeli.

Przedziały ufności

Dostępne są przedziały ufności i błędy standardowe wielu statystyk tabel.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. W konstruktorze tabel przesuń pozycję *Najwyższy stopień* do obszaru wierszy w panelu roboczym.
3. Kliknij opcję **Statystyki podsumowujące...**
4. Na liście **Statystyki** w oknie dialogowym **Statystyki podsumowujące** kliknij ikonę obok pozycji **Liczebność**, aby rozwinąć listę statystyk związanych z liczebnością.
5. Przesuń następujące statystyki do obszaru **Przedstaw**: Kolumna N%, Dolna granica ufności dla %, Górna granica ufności dla % oraz Błąd standardowy % z liczebności w kolumnie.
6. W grupie **Przedziały ufności**, w polu **Poziom (%)**, wprowadź 99.
7. Kliknij przycisk **Zastosuj do wybranej**, a następnie kliknij przycisk **Zamknij**.
8. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Count	Column N %	99.0% Lower CL for Column N %	99.0% Upper CL for Column N %	Standard Error of Column N %
Highest degree	LT High school	430	15.2%	13.6%	17.0%	0.7%
	High school	1500	53.2%	50.7%	55.6%	0.9%
	Junior college	209	7.4%	6.2%	8.7%	0.5%
	Bachelor	478	16.9%	15.2%	18.8%	0.7%
	Graduate	205	7.3%	6.1%	8.6%	0.5%

Rysunek 47. Tabela liczebności, wartości procentowych kolumn i przedziałów ufności

9. Wywołaj okno dialogowe **Custom Tables** i kliknij opcję **Statystyki podsumowujące...**
10. Dla dolnej granicy ufności zmień etykietę na "Dolna granica ufności (&[Poziom ufności])". Łańcuch "&[Poziom ufności]" zostanie zastąpiony wartością określonego poziomu ufności.
11. Kliknij przycisk **Zastosuj do wybranej**, a następnie kliknij przycisk **Zamknij**.
12. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Count	Column N %	Lower Confidence Limit (99.0%)	99.0% Upper CL for Column N %	Standard Error of Column N %
Highest degree	LT High school	430	15.2%	13.6%	17.0%	0.7%
	High school	1500	53.2%	50.7%	55.6%	0.9%
	Junior college	209	7.4%	6.2%	8.7%	0.5%
	Bachelor	478	16.9%	15.2%	18.8%	0.7%
	Graduate	205	7.3%	6.1%	8.6%	0.5%

Rysunek 48. Tabela ze zmodyfikowaną etykietą przedziału ufności

Testy

Testy

Do badania zależności pomiędzy zmiennymi w wierszach i kolumnach dostępne są trzy różne testy istotności. W niniejszym rozdziale omówiono wyniki każdego z takich testów, ze szczególnym uwzględnieniem skutków zagnieżdżania i zestawiania. Więcej informacji zawiera temat “Zestawianie, zagnieżdżanie oraz warstwy w przypadku zmiennych kategorialnych” na stronie 24.

Przykładowy plik danych

W przykładach podanych w niniejszym rozdziale wykorzystano plik danych *survey_sample.sav*. Więcej informacji zawiera temat “Pliki przykładowe” na stronie 77.

Testy niezależności (chi-kwadrat)

Test niezależności chi-kwadrat ma na celu ustalenie istnienia zależności pomiędzy dwiema zmiennymi kategorialnymi. Na przykład może wystąpić konieczność ustalenia, czy zmienna *Obecna sytuacja zawodowa respondenta* jest powiązana ze zmienną *Stan cywilny*.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. W oknie konstruktora tabel przeciągnij zmienną *Obecna sytuacja zawodowa respondenta* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
3. Przeciągnij zmienną *Stan cywilny* z listy zmiennych do pola Kolumny.
4. Wybierz **Wiersze** jako pozycję statystyk podsumowujących.
5. Zaznacz zmienną *Obecna sytuacja zawodowa respondenta* i w grupie Definiuj kliknij **Statystyki podsumowujące**.
6. Na liście Statystyki zaznacz pozycję **% w kolumnie N** i dodaj ją do listy Pokaż w tabeli.
7. Kliknij przycisk **Zastosuj do wybranej**.
8. W oknie dialogowym Tabele użytkownika kliknij kartę **Testy**.
9. Zaznacz opcję **Testy niezależności (Chi-kwadrat)**..
10. Kliknij przycisk **OK**, aby utworzyć tabelę i przeprowadzić test Chi-kwadrat.

			Marital status				
			Married	Widowed	Divorced	Separated	Never married
Labor force status	Working full time	Count	778	44	295	58	392
		Column %	57.8%	15.5%	66.1%	62.4%	59.1%
	Working part-time	Count	138	20	35	9	102
		Column %	10.3%	7.1%	7.8%	9.7%	15.4%
	Temporarily not working	Count	23	2	9	1	11
		Column %	1.7%	.7%	2.0%	1.1%	1.7%
	Unemployed, laid off	Count	13	3	10	0	32
		Column %	1.0%	1.1%	2.2%	.0%	4.8%
	Retired	Count	168	150	53	6	17
		Column %	12.5%	53.0%	11.9%	6.5%	2.6%
	School	Count	9	1	7	2	60
		Column %	.7%	.4%	1.6%	2.2%	9.0%
	Keeping house	Count	200	55	25	13	35
		Column %	14.9%	19.4%	5.6%	14.0%	5.3%
	Other	Count	16	8	12	4	14
		Column %	1.2%	2.8%	2.7%	4.3%	2.1%

Rysunek 49. Tabela krzyżowa zmiennej Obecna sytuacja zawodowa respondenta ze zmienną Stan cywilny

Otrzymana tabela jest tabelą krzyżową zmiennej *Obecna sytuacja zawodowa respondenta* ze zmienną *Stan cywilny* , z proporcjami kolumnowymi jako statystykami podsumowującymi. Proporcje kolumnowe są obliczane w taki sposób, że ich suma w każdej kolumnie wynosi 100%. Jeśli zmienne nie są ze sobą powiązane, wówczas w każdym wierszu proporcje powinny mieć podobną wartość dla wszystkich kolumn. Proporcje wydają się różnić, jednak test chi-kwadrat pozwala upewnić się o ich poprawności.

		Marital status
Labor force status	Chi-square	729.242
	df	28
	Sig.	.000*

*. The Chi-square statistic is significant at the 0.05 level.

Rysunek 50. Test chi-kwadrat Pearsona

Test niezależności zakłada, że zmienne *Obecna sytuacja zawodowa respondenta* i *Stan cywilny* są niezależne od siebie, tzn. proporcje są takie same dla wszystkich kolumn, a wszelkie obserwowane rozbieżności wynikają ze zmienności losowej. Statystyka chi-kwadrat mierzy całkowitą rozbieżność pomiędzy obserwowanymi liczbami komórek a liczbami oczekiwanymi przy założeniu, że proporcje kolumnowe we wszystkich kolumnach są równe. Większa wartość statystyki chi-kwadrat oznacza większą rozbieżność pomiędzy obserwowanymi a oczekiwanymi liczbami komórek — jest to lepszy dowód na to, że proporcje kolumnowe nie są równe i że hipoteza niezależności zmiennych jest błędna, a więc zmienne *Obecna sytuacja zawodowa respondenta* i *Stan cywilny* są ze sobą powiązane.

Obliczona wartość statystyki chi-kwadrat wynosi 729,242. W celu ustalenia, czy otrzymany dowód jest wystarczający do odrzucenia hipotezy niezależności zmiennych, obliczona zostaje wartość istotności. Wartość istotności stanowi prawdopodobieństwo, że wartość krytyczna statystyki z rozkładu chi-kwadrat przy 28 stopniach swobody jest mniejsza niż 729,242. Ponieważ wartość istotności jest mniejsza niż poziom alfa określony na karcie Testy, hipotezę niezależności zmiennych można odrzucić przy poziomie istotności 0,05. A więc zmienne *Obecna sytuacja zawodowa respondenta* i *Stan cywilny* są rzeczywiście ze sobą powiązane.

Skutki zagnieżdżania i zestawiania dla testów niezależności

W przypadku testów niezależności obowiązuje następująca zasada: dla każdej najbardziej wewnętrznej podtabeli wykonywany jest odrębny test. Wpływ zagnieżdżania na testy możemy ocenić wykorzystując poprzedni przykład, jednak ze zmienną *Stan cywilny* zagnieżdżoną w poziomach zmiennej *Płeć*.

1. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Przeciągnij zmienną *Płeć* z listy zmiennych do pola Kolumny w panelu obszaru roboczego powyżej zmiennej, *Stan cywilny*.
3. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Gender	
		Male	Female
		Marital status	Marital status
Labor force status	Chi-square	246.637	542.589
	df	28	28
	Sig.	.000*,1,2	.000*,1,2

*. The Chi-square statistic is significant at the 0.05 level.

1. More than 20% of cells in this sub-table have expected cell counts less than 5.
2. The minimum expected cell count in this sub-table is less than one.

Rysunek 51. Test chi-kwadrat Pearsona

W przypadku zmiennej *Stan cywilny* zagnieżdżonej w poziomach zmiennej *Płeć*, wykonywane są dwa testy — po jednym dla każdego poziomu zmiennej *Płeć*. Wartość istotności dla każdego testu wskazuje, że hipotezę wzajemnej niezależności zmiennych *Stan cywilny* i *Obecna sytuacja zawodowa respondenta* można odrzucić zarówno dla kobiet, jak i mężczyzn. Jednak uwagi do tabel wskazują, że ponad 20% komórek każdej z nich ma oczekiwane liczebności mniejsze niż 5, a minimalna oczekiwana liczba komórek jest mniejsza od 1. Oznacza to, że tabele te mogą nie spełniać założeń testu chi-kwadrat, a więc wyniki testów są niepewne.

Uwaga: Przypisy mogą zostać zakryte obramowaniem komórek. Można temu zapobiec zmieniając wyrównanie komórek w oknie dialogowym Właściwości komórki.

Ocena wpływu zestawiania na testy:

4. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
5. Przeciągnij zmienną *Poziom wykształcenia respondenta* z listy zmiennych do pola Wiersze pod zmienną *Obecna sytuacja zawodowa respondenta*.
6. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Gender	
		Male	Female
		Marital status	Marital status
Labor force status	Chi-square	246.637	542.589
	df	28	28
	Sig.	.000*,1,2	.000*,1,2
Highest degree	Chi-square	43.844	105.506
	df	16	16
	Sig.	.000*	.000*

*. The Chi-square statistic is significant at the 0.05 level.

1. More than 20% of cells in this sub-table have expected cell counts less than 5.
2. The minimum expected cell count in this sub-table is less than one.

Rysunek 52. Test chi-kwadrat Pearsona

W przypadku zestawienia zmiennej *Poziom wykształcenia respondenta* ze zmienną *Obecna sytuacja zawodowa respondenta* wykonywane są cztery testy — test niezależności zmiennych *Stan cywilny* i *Obecna sytuacja zawodowa respondenta*, oraz zmiennych *Stan cywilny* i *Poziom wykształcenia respondenta* dla każdego poziomu zmiennej *Płeć*. Wyniki testów dla zmiennych *Stan cywilny* i *Obecna sytuacja zawodowa respondenta* są identyczne, jak poprzednio. Wyniki testów dla zmiennych *Stan cywilny* i *Poziom wykształcenia respondenta* wskazują, że zmienne te nie są niezależne od siebie.

Porównywanie średnich kolumnowych

Testy średnich kolumnowych mają na celu ustalenie istnienia zależności pomiędzy zmienną kategorialną w kolumnach, a zmienną ilościową w wierszach. Wyniki testów można również wykorzystać do określenia względnego uporządkowania kategorii zmiennej kategorialnej według średniej wartości zmiennej ilościowej. Na przykład może wystąpić konieczność ustalenia, czy zmienna *Liczba godzin oglądania telewizji dziennie* jest powiązana ze zmienną *Informacje czerpie z gazet codziennych*.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. Kliknij przycisk **Resetuj**, aby przywrócić na wszystkich kartach domyślne ustawienia.
3. W oknie konstruktora tabel przeciągnij zmienną *Liczba godzin oglądania telewizji dziennie* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
4. Przeciągnij zmienną *Informacje czerpie z gazet codziennych* z listy zmiennych do pola Kolumny.
5. Zaznacz zmienną *Liczba godzin oglądania telewizji dziennie* i w grupie Definiuj kliknij **Statystyki podsumowujące**.
6. Jako format wybierz opcję **nnnn**.
7. Jako liczbę wyświetlanych miejsc dziesiętnych wybierz **2**. Warto zauważyć, że spowodowało to zmianę formatu na **nnnn.nn**.
8. Kliknij przycisk **Zastosuj do wybranej**.
9. W oknie dialogowym Tabele użytkownika kliknij kartę **Testy**.
10. Zaznacz opcję **Porównaj średnie kolumnowe (testy t)**.
11. Kliknij przycisk **OK**, aby utworzyć tabelę i wykonać testy średnich kolumnowych.

	Get news from newspapers	
	No	Yes
	Mean	Mean
Hours per day watching TV	2.92	2.74

Rysunek 53. Tabela krzyżowa zmiennej Informacje czerpie z gazet codziennych ze zmienną Liczba godzin oglądania telewizji dziennie

Otrzymana tabela pokazuje średnią *liczbę godzin oglądania telewizji dziennie* przez ludzi, którzy uzyskują i nie uzyskują wiadomości z gazet. Widoczna różnica między średnimi oznacza, że ludzie nie uzyskujący wiadomości z gazet spędzają o ok. 0,18 godziny więcej przed telewizorem niż ludzie, którzy uzyskują wiadomości z gazet. Aby sprawdzić, czy taka różnica nie wynika ze zmienności losowej, należy wykonać testy średnich kolumnowych.

	Get news from newspapers	
	No	Yes
	(A)	(B)
Hours per day watching TV		

Rysunek 54. Porównania średnich kolumnowych

W tabeli testu średnich kolumnowych każdej kategorii zmiennej w kolumnie przypisany zostaje klucz literowy. W przypadku zmiennej *Informacje czerpie z gazet codziennych* kategorii *Nie* przypisana zostaje litera A, a kategorii *Tak* litera B. Dla każdej pary kolumn średnie kolumnowe są porównywane za pomocą testu *t*. Ponieważ występują tylko dwie kolumny wykonywany jest tylko jeden test. Dla każdej istotnej pary klucz kategorii z mniejszą średnią jest umieszczany pod kategorią z większą średnią. Ponieważ w komórkach tabeli nie są raportowane żadne klucze, oznacza to, że średnie kolumnowe statystycznie się nie różnią.

Wyniki istotności w notacji stylu APA

Jeśli wyniki istotności nie mają się pojawiać w odrębnej tabeli, można spowodować, aby wyświetlały się w tabeli głównej. Wyniki istotności identyfikuje się przy użyciu notacji stylu APA z literami z indeksem dolnym. Zakończ poprzednie kroki porównywania średnich kolumnowych, ale dokonaj następującej zmiany w karcie Testy:

1. W obszarze Identyfikacja istotnych różnic, wybierz opcję **W tabeli głównej z użyciem indeksu dolnego APA**.
2. Kliknij przycisk **OK**, aby utworzyć tabelę i wykonać testy średnich kolumnowych przy pomocy notacji stylu APA.

	Get news from newspapers	
	No	Yes
	Mean	Mean
Hours per day watching TV	2.92_a	2.74_a

Rysunek 55. Porównania średnich kolumnowych przy pomocy notacji stylu APA

W tabeli testu średnich kolumnowych każdej kategorii zmiennej w kolumnie przypisana zostaje litera z indeksem dolnym. Dla każdej pary kolumn, porównywane są średnie kolumnowe przy pomocy testu *t*. Jeśli para wartości znacząco się od siebie różni, wartości te mają przypisane *różne* litery z indeksem dolnym. Ponieważ występują tylko dwie kolumny wykonywany jest tylko jeden test. Ponieważ średnie kolumnowe w tym przykładzie dzielą tę samą literę z indeksem dolnym, nie są one statystycznie.

Skutki zagnieżdżania i zestawiania dla testów średnich kolumnowych

W przypadku testów średnich kolumnowych obowiązuje następująca zasada: dla każdej najbardziej wewnętrznej podtabeli wykonywany jest osobny zestaw testów parami. Wpływ zagnieżdżania na testy

możemy ocenić wykorzystując poprzedni przykład, jednak ze zmienną *Liczba godzin oglądania telewizji dziennie* zagnieżdżoną w poziomach zmiennej *Obecna sytuacja zawodowa respondenta*.

1. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Przeciągnij zmienną *Obecna sytuacja zawodowa respondenta* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
3. Kliknij przycisk **OK**, aby utworzyć tabelę.

			Get news from newspapers	
			No	Yes
			(A)	(B)
Labor force status	Working full time	Hours per day watching TV	B	
	Working part-time	Hours per day watching TV		
	Temporarily not working	Hours per day watching TV		
	Unemployed, laid off	Hours per day watching TV		
	Retired	Hours per day watching TV		
	School	Hours per day watching TV		
	Keeping house	Hours per day watching TV		
	Other	Hours per day watching TV		

Rysunek 56. Porównania średnich kolumnowych

Przy zmiennej *Liczba godzin oglądania telewizji dziennie* zagnieżdżonej w poziomach zmiennej *Obecna sytuacja zawodowa respondenta* wykonywanych jest siedem zestawów testów średnich kolumnowych: jeden n każdy poziom zmiennej *Obecna sytuacja zawodowa respondenta*. Kategoriom zmiennej *Informacje czerpie z gazet codziennych* przypisane zostają takie same klucze literowe. Dla respondentów pracujących w pełnym wymiarze godzin (*Na pełnym etacie*) w kolumnie kluczy A pojawia się klucz B. Oznacza to, że dla osób pracujących na pełnym etacie średnia wartość zmiennej *Liczba godzin oglądania telewizji dziennie* jest niższa niż dla ludzi, którzy uzyskują wiadomości z gazet. W kolumnie nie występują inne klucze, można więc stwierdzić, że w średnich kolumnowych nie występują inne statystycznie istotne różnice.

Korekty Bonferroniego. Przy wykonywaniu kilku testów do testów średnich kolumnowych stosowana jest korekta Bonferroniego celem zapewnienia, że poziom alfa (czyli wskaźnik wyników fałszywie dodatnich) określony na karcie Testy ma zastosowanie do każdego *zestawu* testów. Jednak w przypadku tej tabeli nie zastosowano korekt Bonferroniego, ponieważ chociaż wykonano siedem zestawów testów, w każdym zestawie porównywana była tylko jedna para kolumn.

Ocena wpływu zestawiania na testy:

4. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
5. Przeciągnij zmienną *Informacje czerpie z Internetu* z listy zmiennych do pola Kolumny, znajdującą się po lewej stronie zmiennej *Informacje czerpie z gazet codziennych*.
6. Kliknij przycisk **OK**, aby utworzyć tabelę.

			Get news from internet		Get news from newspapers	
			No	Yes	No	Yes
			(A)	(B)	(A)	(B)
Labor force status	Working full time	Hours per day watching TV	B		B	
	Working part-time	Hours per day watching TV	B			
	Temporarily not working	Hours per day watching TV				
	Unemployed, laid off	Hours per day watching TV	B			
	Retired	Hours per day watching TV	B			
	School	Hours per day watching TV	B			
	Keeping house	Hours per day watching TV	B			
	Other	Hours per day watching TV	B			

Rysunek 57. Porównania średnich kolumnowych

Razem ze zmienną *Informacje czerpie z Internetu* zestawionej ze zmienną *Informacje czerpie z gazet codziennych*, wykonywanych jest 14 zestawów testów średnich kolumnowych — po jednym dla każdego poziomu zmiennej *Obecna sytuacja zawodowa respondenta* dla zmiennych*Informacje czerpie z Internetu* i *Informacje czerpie z gazet codziennych*. Również w tym przypadku nie są stosowane żadne korekty Bonferroniego, ponieważ w każdym zestawie porównywana była tylko jedna para kolumn. Testy dla zmiennej *Informacje czerpie z gazet codziennych* są identyczne,jak poprzednio. W przypadku zmiennej *Informacje czerpie z gazet codziennych*, kategorii *Nie* przypisana zostaje litera A, a kategorii *Tak* litera B. Klucz B jest raportowany w kolumnie A dla każdego zestawu testów średnich kolumnowych z wyjątkiem respondentów czasowo niezatrudnionych — Czasowo nie pracuje. Oznacza to, że średnia wartość zmiennej *Liczba godzin oglądania telewizji dziennie* jest niższa dla osób uzyskujących wiadomości z Internetu niż dla tych, którzy nie uzyskują wiadomości z gazet. Dla zestawu *Czasowo nie pracuje* nie są raportowane żadne klucze; a więc średnie kolumnowe nie różnią się statystycznie dla tych respondentów.

Porównywanie proporcji kolumnowych

Testy proporcji kolumnowych mają na celu ustalenie względnego uporządkowania kategorii zmiennej kategorialnej w kolumnach według proporcji kategorii zmiennej kategorialnej w wierszach. Na przykład po zastosowaniu testu chi-kwadrat w celu stwierdzenia, że zmienne *Obecna sytuacja zawodowa respondenta* i *Stan cywilny* nie są niezależne, można sprawdzić, które wiersze i kolumny powodują taką zależność.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. Kliknij przycisk **Resetuj**, aby przywrócić na wszystkich kartach domyślne ustawienia.
3. W oknie konstruktora tabel przeciągnij zmienną *Obecna sytuacja zawodowa respondenta* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
4. Przeciągnij zmienną *Stan cywilny* z listy zmiennych do pola Kolumny.
5. Zaznacz zmienną *Obecna sytuacja zawodowa respondenta* i w grupie Definiuj kliknij **Statystyki podsumowujące**.
6. Na liście Statystyki zaznacz pozycję **% w kolumnie N** i dodaj ją do listy Pokaż w tabeli.
7. Usuń zaznaczenie pozycji **Liczebność** z listy Pokaż.
8. Kliknij przycisk **Zastosuj do wybranej**.
9. W oknie dialogowym Tabele użytkownika kliknij kartę **Testy**.
10. Zaznacz opcję **Porównaj proporcje kolumnowe (testy z)**.

11. Kliknij przycisk **OK**, aby utworzyć tabelę i wykonać testy proporcji kolumnowych.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		Column %	Column %	Column %	Column %	Column %
Labor force status	Working full time	57.8%	15.5%	66.1%	62.4%	59.1%
	Working part-time	10.3%	7.1%	7.8%	9.7%	15.4%
	Temporarily not working	1.7%	.7%	2.0%	1.1%	1.7%
	Unemployed, laid off	1.0%	1.1%	2.2%	.0%	4.8%
	Retired	12.5%	53.0%	11.9%	6.5%	2.6%
	School	.7%	.4%	1.6%	2.2%	9.0%
	Keeping house	14.9%	19.4%	5.6%	14.0%	5.3%
	Other	1.2%	2.8%	2.7%	4.3%	2.1%

Rysunek 58. Tabela krzyżowa zmiennej Obecna sytuacja zawodowa respondenta ze zmienną Stan cywilny

Otrzymana tabela jest tabelą krzyżową zmiennej *Obecna sytuacja zawodowa respondenta* ze zmienną *Stan cywilny* , z proporcjami kolumnowymi jako statystykami podsumowującymi.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		(A)	(B)	(C)	(D)	(E)
Labor force status	Working full time	B		A B	B	B
	Working part-time					A B C
	Temporarily not working					
	Unemployed, laid off				.	A B
	Retired	E	A C D E	E		
	School					A B C
	Keeping house	C E	C E		C E	
	Other					

Rysunek 59. Porównania proporcji kolumnowych

W tabeli testu proporcji kolumnowych każdej kategorii zmiennej w kolumnie przypisany zostaje klucz literowy. W przypadku zmiennej *Stan cywilny* kategorii *Żonaty / Zamężna / W konkubinacie* przypisana zostaje litera A, a kategorii*Wdowiec / Wdowa*litera B, itd. aż do kategorii *Kawaler / Panna*, dla której przypisana jest litera E. Dla każdej pary kolumn proporcje kolumnowe są porównywane za pomocą testu *z*. Wykonywanych jest siedem zestawów testów proporcji kolumnowych, po jednym dla każdego poziomu zmiennej *Obecna sytuacja zawodowa respondenta*. Ponieważ mamy 5 poziomów zmiennej *Stan cywilny* w każdym zestawie testów porównywanych jest (5*4)/2 = 10 par kolumn, z wykorzystaniem korekt Bonferroniego do korekty wartości istotności. Dla każdej pary istotnej klucz kategorii z mniejszą proporcją jest umieszczany pod kategorią z większą proporcją.

W przypadku zestawu testów powiązanych ze zmienną *Na pełnym etacie*, klucz B występuje w każdej z kolumn. Klucz A występuje również w kolumnie kluczy C. W pozostałych kolumnach nie są raportowane żadne klucze. Można więc wyciągnąć wniosek, że proporcja osób rozwiedzionych pracujących w pełnym wymiarze godzin jest większa niż proporcja osób zamężnych/żonatych pracujących w pełnym wymiarze godzin, która z kolei jest większa niż proporcja wdów/wdowców pracujących w pełnym wymiarze godzin. Proporcje osób w separacji lub nigdy nie zamężnych/żonatych i pracujących w pełnym wymiarze godzin są bardzo zbliżone do proporcji osób rozwiedzionych lub zamężnych/żonatych i pracujących w pełnym wymiarze godzin, lecz są większe niż proporcje dla wdów/wdowców pracujących w pełnym wymiarze godzin.

W przypadku testów powiązanych z kategorią *Na niepełnym etacie* lub *Uczy się / studiuje* klucze A, B i C występują w kolumnie kluczy E. W pozostałych kolumnach nie są raportowane żadne klucze. A więc proporcje osób nigdy nie zamężnych/żonatych i uczących się lub pracujących w niepełnym wymiarze

godzin są większe niż proporcje osób zamężnych/żonatych, wdów/wdowców lub rozwiedzionych i uczących się lub pracujących w niepełnym wymiarze godzin.

W przypadku testów powiązanych z kategorią *Czasowo nie pracuje* lub *INNA SYTUACJA* we wszystkich kolumnach nie są raportowane żadne klucze. A więc nie występuje zauważalna różnica w proporcjach osób zamężnych/żonatych, wdów/wdowców, w separacji lub nigdy nie zamężnych/żonatych, którzy są tymczasowo bez pracy lub mają inny, nieusystematyzowany status zatrudnienia.

Testy powiązane z kategorią *Emerytura / Renta* pokazują, że proporcja wdów/wdowców na emeryturze jest większa niż proporcje dla wszystkich pozostałych kategorii stanu cywilnego, będących na emeryturze. Ponadto proporcja osób zamężnych/żonatych lub rozwiedzionych i będących na emeryturze jest większa niż proporcja osób nigdy nie zamężnych/żonatych będących na emeryturze.

Proporcje osób żonatych/zamężnych, wdów/wdowców lub w separacji, które należą do kategorii Zajmuje się domem są większe niż w przypadku osób rozwiedzionych lub nigdy nie zamężnych/żonatych.

Proporcja osób nigdy nie zamężnych/żonatych, należących do kategorii *Bezrobotny i poszukujący pracy* jest większa niż w przypadku osób żonatych/zamężnych lub wdów/wdowców. Należy również zauważyć, że kolumna *W separacji* jest oznaczona kropką ("."), co oznacza, że obserwowana proporcja osób w separacji w wierszu *Bezrobotny i poszukujący pracy* wynosi 0 lub 1, a w konsekwencji nie można dokonać porównań wykorzystując tę kolumnę dla respondentów bezrobotnych.

Scalanie wyników istotności w tabelę główną

Jeśli wyniki istotności nie mają się pojawiać w odrębnej tabeli, można spowodować, aby wyświetlały się w tabeli głównej. Zakończ poprzednie kroki porównywania proporcji kolumn, ale dokonaj następującej zmiany w karcie Testy:

1. W obszarze Identyfikacja istotnych różnic, wybierz opcję **W tabeli głównej**.
2. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		(a)	(b)	(c)	(d)	(e)
		Column N %	Column N %	Column N %	Column N %	Column N %
Labor force status	Working full time	57.8% b	15.5%	66.1% a b	62.4% b	59.1% b
	Working part-time	10.3%	7.1%	7.8%	9.7%	15.4% a b c
	Temporarily not working	1.7%	0.7%	2.0%	1.1%	1.7%
	Unemployed, laid off	1.0%	1.1%	2.2%	0.0%[1]	4.8% a b
	Retired	12.5% e	53.0% a c d e	11.9% e	6.5%	2.6%
	School	0.7%	0.4%	1.6%	2.2%	9.0% a b c
	Keeping house	14.9% c e	19.4% c e	5.6%	14.0% c e	5.3%
	Other	1.2%	2.8%	2.7%	4.3%	2.1%

Rysunek 60. Wyniki oceny istotności scalone w główną tabelę

3. Kliknij dwukrotnie tabelę w **oknie raportu**, aby ją aktywować.
4. Zatrzymaj wskaźnik nad jedną z etykiet kolumn zawierających klucze literowe. Na przykład zatrzymaj wskaźnik nad komórką z etykietą "(a)".

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		(a)	(b)	(c)	(d)	(e)
		Column N %	Column N %	Column N %	Column N %	Column N %
Labor force status	Working full time	57.8% b	15.5%	66.1% a b	62.4% b	59.1% b
	Working part-time	10.3%	7.1%	7.8%	9.7%	15.4% a b c
	Temporarily not working	1.7%	0.7%	2.0%	1.1%	1.7%
	Unemployed, laid off	1.0%	1.1%	2.2%	0.0%[1]	4.8% a b
	Retired	12.5% e	53.0% a c d e	11.9% e	6.5%	2.6%
	School	0.7%	0.4%	1.6%	2.2%	9.0% a b c
	Keeping house	14.9% c e	19.4% c e	5.6%	14.0% c e	5.3%
	Other	1.2%	2.8%	2.7%	4.3%	2.1%

Rysunek 61. Wyróżnianie wszystkich komórek zawierających wybrany klucz literowy

Wszystkie komórki zawierające ten klucz zostaną w tabeli wyróżnione.

5. Kliknij prawym przyciskiem myszy komórkę etykiety kolumny z kluczem literowym. Z menu kontekstowego wybierz opcję **Zaznacz** > **Zaznacz wszystkie komórki z tym kluczem istotności**.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		(a)	(b)	(c)	(d)	(e)
		Column N %	Column N %	Column N %	Column N %	Column N %
Labor force status	Working full time	57.8% b	15.5%	66.1% a b	62.4% b	59.1% b
	Working part-time	10.3%	7.1%	7.8%	9.7%	15.4% a b c
	Temporarily not working	1.7%	0.7%	2.0%	1.1%	1.7%
	Unemployed, laid off	1.0%	1.1%	2.2%	0.0%[1]	4.8% a b
	Retired	12.5% e	53.0% a c d e	11.9% e	6.5%	2.6%
	School	0.7%	0.4%	1.6%	2.2%	9.0% a b c
	Keeping house	14.9% c e	19.4% c e	5.6%	14.0% c e	5.3%
	Other	1.2%	2.8%	2.7%	4.3%	2.1%

Rysunek 62. Wybieranie wszystkich komórek zawierających wybrany klucz literowy

Wszystkie komórki zawierające ten klucz zostaną w tabeli zaznaczone.

Skutki zagnieżdżania i zestawiania w przypadku testów proporcji kolumnowych

W przypadku testów proporcji kolumnowych obowiązuje następująca zasada: dla każdej najbardziej wewnętrznej podtabeli wykonywany jest osobny zestaw testów parami. Wpływ zagnieżdżania na testy możemy ocenić wykorzystując poprzedni przykład, jednak ze zmienną *Obecna sytuacja zawodowa respondenta* zagnieżdżoną w poziomach zmiennej *Płeć*.

1. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Przeciągnij zmienną *Płeć* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
3. Kliknij przycisk **OK**, aby utworzyć tabelę.

				Marital status				
				Married	Widowed	Divorced	Separated	Never married
				(A)	(B)	(C)	(D)	(E)
Gender	Male	Labor force status	Working full time	B		B	B	B
			Working part-time					A
			Temporarily not working		.			
			Unemployed, laid off		.		.	A
			Retired	E	A C D E	E		
			School		.			A C
			Keeping house					
			Other				A	
	Female	Labor force status	Working full time	B		A B	B	B
			Working part-time	B				B
			Temporarily not working				.	
			Unemployed, laid off				.	A
			Retired	E	A C D E	E		
			School					A B C
			Keeping house	C E	C E		C	
			Other					

Rysunek 63. Porównania proporcji kolumnowych

W przypadku zmiennej *Obecna sytuacja zawodowa respondenta* zagnieżdżonej w poziomach zmiennej *Płeć*, wykonywanych jest 14 zestawów testów proporcji kolumnowych — po jednym dla każdego poziomu zmiennej *Obecna sytuacja zawodowa respondenta* dla każdego poziomu zmiennej *Płeć*. Kategoriom zmiennej *Stan cywilny* przypisane zostają te same klucze literowe.

W wynikach otrzymanych w tabeli można zauważyć klika prawidłowości:

- Przy większej liczbie testów występuje więcej kolumn z zerowymi proporcjami kolumnowymi. Najczęściej występują one wśród respondentów w separacji i wdowców.
- Wydaje się, że różnice w kolumnie poprzednio obserwowane wśród respondentów o kategorii *Zajmuje się domem* można całkowicie przypisać kobietom.

Ocena wpływu zestawiania na testy:

4. Otwórz ponownie konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
5. Przeciągnij zmienną *Poziom wykształcenia respondenta* z listy zmiennych do pola Wiersze pod zmienną *Płeć*.
6. Kliknij przycisk **OK**, aby utworzyć tabelę.

				Marital status				
				Married	Widowed	Divorced	Separated	Never married
				(A)	(B)	(C)	(D)	(E)
Gender	Male	Labor force status	Working full time	B		B	B	B
			Working part-time					A
			Temporarily not working		.			
			Unemployed, laid off		.		.	A
			Retired	E	A C D E	E		
			School		.			A C
			Keeping house					
			Other				A	
	Female	Labor force status	Working full time	B		A B	B	B
			Working part-time	B				B
			Temporarily not working				.	
			Unemployed, laid off				.	A
			Retired	E	A C D E	E		
			School					A B C
			Keeping house	C E	C E		C	
			Other					
Highest degree	LT High school				A C E			
	High school							
	Junior college			B		B		B
	Bachelor			B				B
	Graduate			B				

Rysunek 64. Porównania proporcji kolumnowych

W przypadku zestawienia zmiennej *Poziom wykształcenia respondenta* ze zmienną *Płeć*, wykonywanych jest 19 zestawów testów średnich kolumnowych — 14 omówionych powyżej oraz po jednym dla każdego poziomu zmiennej *Poziom wykształcenia respondenta*. Kategoriom zmiennej *Stan cywilny* przypisane zostają te same klucze literowe.

W wynikach otrzymanych w tabeli można zauważyć klika prawidłowości:

- Wyniki 14 testów są takie same, jak poprzednio.
- Wykształcenie Mniej niż HS jest powszechniejsze u wdowców niż u respondentów zamężnych/żonatych, rozwiedzionych i nigdy nie zamężnych/żonatych.
- Wykształcenie policealne wydaje się być częstsze wśród osób zamężnych/żonatych, rozwiedzionych i nigdy nie zamężnych/żonatych.

Uwaga na temat wag i zestawów wielokrotnych odpowiedzi

Wagi obserwacji zawsze bazują na liczebnościach, nie na odpowiedziach; nawet w przypadku gdy jedna ze zmiennych jest zmienną wielokrotnych odpowiedzi.

Zestawy wielokrotnych odpowiedzi

Moduł Tabele użytkownika i Kreator wykresów obsługuje również specjalny rodzaj „zmiennej” o nazwie **zestaw wielokrotnych odpowiedzi**. Zestawy wielokrotnych odpowiedzi nie są „zmiennymi” w normalnym tego słowa znaczeniu. Nie występują w Edytorze danych, a inne procedury nie rozpoznają ich. Zestawy wielokrotnych odpowiedzi wykorzystują wiele zmiennych do rejestracji odpowiedzi na pytania w przypadku, gdy respondent może udzielić więcej niż jednej odpowiedzi. Są one traktowane podobnie jak zmienne kategorialne i można je, w większości przypadków, poddawać podobnym operacjom.

Zestawy wielokrotnych odpowiedzi są tworzone z wielu zmiennych w pliku danych. Zestaw wielokrotnych odpowiedzi stanowi specjalną konstrukcję w pliku danych. W pliku danych programu IBM SPSS Statistics można zdefiniować i zapisać wiele zestawów wielokrotnych odpowiedzi, jednak nie można ich importować ani eksportować z/do innych formatów plików. Za pomocą opcji Skopiuj właściwości danych, dostępnej z menu Dane w oknie Edytora danych, można kopiować zestawy wielokrotnych odpowiedzi z innych plików danych programu IBM SPSS Statistics.

Przykładowy plik danych

W przykładach podanych w niniejszym rozdziale wykorzystano plik danych *survey_sample.sav*. Więcej informacji zawiera temat "Pliki przykładowe" na stronie 77.

We wszystkich przedstawionych tutaj przykładach w oknach dialogowych wyświetlane są etykiety zmiennych, posortowane w porządku alfabetycznym. Opcje wyświetlania listy zmiennych są ustawiane na karcie Ogólne, w oknie dialogowym Opcje (menu Edycja, Opcje).

Liczebności, odpowiedzi, procenty i podsumowania

Wszystkie statystyki podsumowujące dostępne dla zmiennych kategorialnych są również dostępne dla zestawów wielokrotnych odpowiedzi. Dostępnych jest również kilka dodatkowych statystyk.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**

2. Przeciągnij element *Źródła informacji* (jest to opisowa etykieta zestawu wielokrotnych odpowiedzi *$infozes*) z listy zmiennych do pola Wiersze w panelu obszaru roboczego.

   Ikona obok „zmiennej" na liście zmiennych identyfikuje ją jako zestaw wielokrotnych dychotomii.

   Rysunek 65. Ikona zestawu wielokrotnych dychotomii

   W zestawie wielokrotnych dychotomii każda „kategoria" jest w zasadzie odrębną zmienną, a etykietami kategorii są etykiety zmiennych (lub nazwy zmiennych w przypadku zmiennych bez zdefiniowanych etykiet). W tym przykładzie liczebności, które będą wyświetlane, reprezentują liczbę obserwacji z odpowiedzią *Tak* dla każdej zmiennej wchodzącej w skład zestawu.

3. Kliknij prawym przyciskiem myszy zmienną *Źródła informacji* na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Kategorie i podsumowania**.
4. W oknie dialogowym Kategorie i podsumowania zaznacz (kliknij) opcję **Podsumowanie ogółem**, a następnie kliknij przycisk **Zastosuj**.
5. Ponownie kliknij prawym przyciskiem myszy zmienną *Źródła informacji* i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
6. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję **% z N w kolumnie** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
7. Kliknij przycisk **Zastosuj do wybranej**, a następnie przycisk **OK**, aby utworzyć tabelę.

		Count	Column N %
News sources	Get news from internet	867	41.7%
	Get news from radio	551	26.5%
	Get news from television	1077	51.8%
	Get news from news magazines	294	14.1%
	Get news from newspapers	805	38.7%
	Total	2081	100.0%

Rysunek 66. Liczebności i procenty w kolumnie dla wielokrotnych dychotomii

Niezgadzające się sumy

Patrząc na liczby w tabeli można zauważyć, że występuje dosyć duża rozbieżność pomiędzy „podsumowaniami", a wartościami, które rzekomo są sumowane — sumy wydają się być dużo mniejsze

niż powinny. Wynika to z faktu, że liczebność dla każdej „kategorii” występującej w tabeli jest liczbą obserwacji z wartością 1 (odpowiedź *Tak*) dla danej zmiennej, a całkowita liczba odpowiedzi *Tak* dla wszystkich pięciu zmiennych, wchodzących w skład zestawu wielokrotnych dychotomii, może z łatwością przekroczyć całkowitą liczbę obserwacji zawartych w pliku danych.

„Liczebność” ogółem jest natomiast całkowitą liczbą obserwacji z odpowiedzią *Tak* dla przynajmniej jednej zmiennej występującej w zestawie, która nigdy nie może przekroczyć całkowitej liczby obserwacji zawartych w pliku danych. W niniejszym przykładzie liczebność ogółem wynosząca 2081 jest o prawie 800 jednostek mniejsza od całkowitej liczby obserwacji zawartych w pliku danych. Jeżeli żadna z tych zmiennych nie zawiera braków danych, oznacza to, że prawie 800 respondentów ankiety udzieliło odpowiedzi, że nie uzyskuje wiadomości z żadnego z podanych źródeł. Na podstawie liczebności ogółem obliczane są wartości procentowe w kolumnie, co oznacza, że suma wartości procentowych w kolumnie w tym przykładzie wynosi więcej niż 100% wyświetlone jako wartość procentowa w kolumnie ogółem.

Zgadzające się sumy

Chociaż termin „liczebność” jest zwykle dosyć jednoznaczny, powyższy przykład pokazuje, jak mylący może on być w kontekście podsumowań dla zestawów wielokrotnych odpowiedzi, w których żądaną statystyką podsumowującą są często *odpowiedzi*.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy zmienną *Źródła informacji* na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.
3. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję **Odpowiedzi** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
4. Na liście Statystyki zaznacz pozycję **% z odpowiedzi w kolumnie** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
5. Kliknij przycisk **Zastosuj do wybranej**, a następnie przycisk **OK**, aby utworzyć tabelę.

		Count	Column N %	Responses	Column Responses %
News sources	Get news from internet	867	41.7%	867	24.1%
	Get news from radio	551	26.5%	551	15.3%
	Get news from television	1077	51.8%	1077	30.0%
	Get news from news magazines	294	14.1%	294	8.2%
	Get news from newspapers	805	38.7%	805	22.4%
	Total	2081	100.0%	3594	100.0%

Rysunek 67. Odpowiedzi i procenty z odpowiedzi w kolumnie dla wielokrotnych dychotomii

Dla każdej „kategorii” w zestawie wielokrotnych dychotomii statystyka *Odpowiedzi* odpowiada zawsze statystyce *Liczebność*. Jednak podsumowania ogółem znacznie się różnią. Całkowita liczba odpowiedzi wynosi 3594 — ponad 1500 odpowiedzi więcej niż liczebność ogółem i o ponad 700 odpowiedzi więcej niż całkowita liczba obserwacji w pliku danych.

W przypadku wartości procentowych podsumowania ogółem dla statystyk *% z N w kolumnie* i *% z odpowiedzi w kolumnie* wynoszą 100%, lecz wartości procentowe z odpowiedzi w kolumnie dla poszczególnych kategorii w zestawie wielokrotnych dychotomii są dużo niższe. Wynika to z faktu, iż podstawa obliczania wartości procentowej w kolumnie obliczana jest na podstawie całkowitej liczby odpowiedzi, która w tym przypadku wynosi 3594, co daje dużo mniejsze wartości procentowe niż w przypadku obliczania procentu z liczby 2081.

Wartości procentowe przekraczające 100%

Wartości procentowe w kolumnie, jak i z odpowiedzi w kolumnie, wynoszą 100%, chociaż w przedstawionym przykładzie suma poszczególnych wartości w kolumnie *% z N w kolumnie* jest większa niż 100%. Powstaje pytanie, jak wyświetlić wartość procentową obliczaną na podstawie całkowitej liczby obserwacji, a nie całkowitej liczby odpowiedzi, wyświetlając jednocześnie sumę wartości procentowych w poszczególnych kategoriach.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).

2. Kliknij prawym przyciskiem myszy zmienną *Źródła informacji* na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.
3. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję **% z Odpowiedzi (Baza: liczebność)** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
4. Kliknij przycisk **Zastosuj do wybranej**, a następnie przycisk **OK**, aby utworzyć tabelę.

		Count	Column N %	Responses	Column Responses %	Column Responses % (Base: Count)
News sources	Get news from internet	867	41.7%	867	24.1%	41.7%
	Get news from radio	551	26.5%	551	15.3%	26.5%
	Get news from television	1077	51.8%	1077	30.0%	51.8%
	Get news from news magazines	294	14.1%	294	8.2%	14.1%
	Get news from newspapers	805	38.7%	805	22.4%	38.7%
	Total	2081	100.0%	3594	100.0%	172.7%

Rysunek 68. Procent z odpowiedzi w kolumnie obliczany na podstawie obliczania wartości procentowej

Stosowanie zestawów wielokrotnych odpowiedzi z innymi zmiennymi

W zasadzie zestawów wielokrotnych odpowiedzi można używać podobnie jak zmiennych kategorialnych. Można, na przykład umieścić zestaw wielokrotnych odpowiedzi w tabeli krzyżowej ze zmienną kategorialną lub zagnieździć go w niej.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Przeciągnij zmienną *Płeć* z listy zmiennych do lewej krawędzi pola Wiersze w panelu podglądu, zagnieżdżając zestaw wielokrotnych odpowiedzi *Źródła informacji* w kategoriach płci.
3. Kliknij prawym przyciskiem myszy zmienną *Płeć* na podglądzie tabeli w panelu obszaru roboczego i w podręcznym menu kontekstowym usuń zaznaczenie opcji **Pokaż etykietę zmiennej**.
4. Powtórz czynność dla zmiennej *Źródła informacji*.

   Spowoduje to usunięcie z tabeli kolumn zawierających etykiety zmiennych (ponieważ w tym przypadku są właściwie zbędne).
5. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Count	Column N %	Responses	Column Responses %	Column Responses % (Base: Count)
Male	Get news from internet	359	40.1%	359	23.3%	40.1%
	Get news from radio	233	26.0%	233	15.1%	26.0%
	Get news from television	451	50.3%	451	29.3%	50.3%
	Get news from news magazines	121	13.5%	121	7.9%	13.5%
	Get news from newspapers	375	41.9%	375	24.4%	41.9%
	Total	896	100.0%	1539	100.0%	171.8%
Female	Get news from internet	508	42.9%	508	24.7%	42.9%
	Get news from radio	318	26.8%	318	15.5%	26.8%
	Get news from television	626	52.8%	626	30.5%	52.8%
	Get news from news magazines	173	14.6%	173	8.4%	14.6%
	Get news from newspapers	430	36.3%	430	20.9%	36.3%
	Total	1185	100.0%	2055	100.0%	173.4%

Rysunek 69. Zestaw wielokrotnych odpowiedzi zagnieżdżony w zmiennej kategorialnej

Zmienna statystyki źródłowej oraz dostępne statystyki podsumowujące

Jeśli tabela nie zawiera zmiennej ilościowej, zmienne kategorialne i zestawy wielokrotnych odpowiedzi są traktowane tak samo pod względem zmiennej statystyki źródłowej: najbardziej wewnętrznie zagnieżdżona zmienna w wymiarze statystyki źródłowej jest uznawana za zmienną statystyki źródłowej. Ponieważ istnieją statystyki podsumowujące, które można przypisać wyłącznie zestawom wielokrotnych odpowiedzi, oznacza to, że aby uzyskać taką statystykę dla zestawu wielokrotnych odpowiedzi, musi on być najbardziej zagnieżdżoną zmienną w wymiarze statystyki źródłowej.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).

2. Na podglądzie tabeli w panelu obszaru roboczego przeciągnij zmienną *News sources* do lewej strony zmiennej *Gender*, zmieniając porządek zagnieżdżenia.

Wszystkie specjalne statystyki zestawów wielokrotnych odpowiedzi — odpowiedzi, procent z odpowiedzi w kolumnie — zostaną usunięte z podglądu tabeli, ponieważ zmienna kategorialna *Płeć* jest teraz najbardziej zagnieżdżoną zmienną, a w konsekwencji, zmienną statystyki źródłowej.

Na szczęście konstruktor tabel „zapamiętuje" ustawienia tego typu. Po przeniesieniu zmiennej *Źródła informacji* w poprzednie położenie, gdy była zagnieżdżona w zmiennej *Płeć*, wszystkie statystyki podsumowujące związane z odpowiedziami zostaną przywrócone do podglądu tabeli.

Zestawy wielokrotnych odpowiedzi a odpowiedzi powtórzone

Zestawy wielokrotnych kategorii oferują dodatkową możliwość niedostępną w przypadku zestawów wielokrotnych dychotomii: możliwość zliczania powtórzonych odpowiedzi. W wielu przypadkach powtórzone odpowiedzi w zestawach wielokrotnych kategorii wynikają prawdopodobnie z błędów kodowania. Na przykład w przypadku pytania „W jakich krajach, Twoim zdaniem, produkuje się najlepsze samochody?" odpowiedź *Szwecja, Niemcy i Szwecja* jest prawdopodobnie nieprawidłowa.

Jednakże w innych przypadkach, powtórzone odpowiedzi mogą być jak najbardziej prawidłowe. Na przykład w przypadku pytania „W jakich krajach wyprodukowano twoje trzy ostatnie samochody?" odpowiedź *Szwecja, Niemcy i Szwecja* ma sens.

Moduł Tabele użytkownika umożliwia powtarzanie odpowiedzi w zestawach wielokrotnych kategorii. Domyślnie, powtórzone odpowiedzi nie są zliczane, lecz można włączyć ich uwzględnianie.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia.
3. Przeciągnij zmienną *Kraj produkcji ostatnich posiadanych aut* z listy zmiennych do pola Wiersze panelu obszaru roboczego.

   Ikona obok „zmiennej" na liście zmiennych identyfikuje ją jako zestaw wielokrotnych kategorii.

   Rysunek 70. Ikona zestawu wielokrotnych kategorii

   W przypadku zestawów wielokrotnych kategorii wyświetlane kategorie reprezentują wspólny zestaw zdefiniowanych etykiet wartości dla wszystkich zmiennych wchodzących w jego skład (natomiast w przypadku zestawów wielokrotnych dychotomii „kategorie" są w istocie etykietami wartości dla każdej zmiennej wchodzącej w skład zestawu).
4. Kliknij prawym przyciskiem myszy zmienną *Kraj produkcji ostatnich posiadanych aut* na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Kategorie i podsumowania**.
5. W oknie dialogowym Kategorie i podsumowania zaznacz (kliknij) opcję **Podsumowanie ogółem**, a następnie kliknij przycisk **Zastosuj**.
6. Ponownie kliknij prawym przyciskiem myszy zmienną *Kraj produkcji ostatnich posiadanych aut* i z menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
7. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję **Odpowiedzi** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
8. Kliknij przycisk **Zastosuj do wybranej**, a następnie przycisk **OK**, aby utworzyć tabelę.

		Count	Responses
Car maker, most recent cars	American	1938	1938
	Japanese	1327	1327
	Korean	695	695
	German	693	693
	Swedish	360	360
	Other	343	343
	Total	2832	5356

Rysunek 71. Zestaw wielokrotnych kategorii: liczebności i odpowiedzi bez powtórzonych wartości

Domyślnie, powtórzone odpowiedzi nie są zliczane, a więc w tej tabeli wartości dla każdej kategorii w kolumnie *Liczebność* i *Odpowiedzi* są identyczne. Różnią się jedynie podsumowania ogółem.

9. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
10. Kliknij kartę **Opcje**.
11. Kliknij (zaznacz) opcję **Zlicz powtórzone odpowiedzi dla zestawów wielokrotnych kategorii**.
12. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Count	Responses
Car maker, most recent cars	American	1938	2797
	Japanese	1327	1717
	Korean	695	760
	German	693	754
	Swedish	360	383
	Other	343	359
	Total	2832	6770

Rysunek 72. Zestaw wielokrotnych odpowiedzi z powtórzonymi odpowiedziami

W tej tabeli występuje zauważalna różnica pomiędzy wartościami w kolumnie *Liczebność* i *Odpowiedzi*, zwłaszcza w przypadku samochodów amerykańskich, co oznacza, że wielu respondentów posiadało po kilka takich samochodów.

Testowanie istotności z zestawami wielokrotnych odpowiedzi

W testach istotności można korzystać z zestawów wielokrotnych odpowiedzi w taki sam sposób jak ze zmiennych kategorialnych.

- W przypadku testów niezależności (chi-kwadrat) lub porównywania proporcji kolumnowych (testy z) testy przeprowadzane są na liczebności, która musi stanowić jedną ze statystyk podsumowujących, wyświetlanych w tabeli.
- W przypadku zestawów wielokrotnych kategorii nie są wykonywane testy porównujące proporcje kolumny lub średnie kolumnowe (testy t), jeśli na karcie Opcje wybrano opcję **Zlicz powtórzone odpowiedzi dla zestawów wielokrotnych kategorii**. Więcej informacji zawiera temat “Tabele użytkownika: karta Opcje” na stronie 15.

Testy niezależności z zestawami wielokrotnych odpowiedzi

Ten przykład tworzy tabelę krzyżową ze zmiennej kategorialnej i zestawu wielokrotnych odpowiedzi oraz wykonuje test chi-kwadrat niezależności tabeli krzyżowej.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia.
3. Przeciągnij element *Źródła informacji* (jest to opisowa etykieta zestawu wielokrotnych dychotomii *$infozes*) z listy zmiennych do pola Kolumny w panelu obszaru roboczego.
4. Przeciągnij zmienną *Płeć* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
5. Kliknij kartę **Testy**.
6. Zaznacz opcję **Testy niezależności (chi-kwadrat)**..
7. Wybierz opcję **Uwzględnij zmienne wielokrotnych odpowiedzi w teście**.

8. Kliknij przycisk **OK**, aby wykonać procedurę.

		News sources
Gender	Chi-square	10.266
	df	5
	Sig.	.068

Rysunek 73. Wyniki chi-kwadrat

Poziom istotności testu chi-kwadra wynoszący 0,068 wskazuje na to, że mężczyźni i kobiety najprawdopodobniej nie różnią się znacznie w doborze źródeł wiadomości (zakładając, że jako kryterium wykazywania istotności statystycznej użyto wartości istotności równej lub mniejszej od 0,05).

Porównywanie średnich kolumnowych z zestawami wielokrotnych odpowiedzi

W tym przykładzie obliczana jest średnia zmiennej ilościowej zawarta w kategoriach określonych przez zestawy wielokrotnych odpowiedzi oraz porównywane są średnie kategorii z innymi średnimi kategorii w celu znalezienia znaczących różnic.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia.
3. Przeciągnij element *Źródła informacji* (jest to opisowa etykieta zestawu wielokrotnych dychotomii *$infozes*) z listy zmiennych do pola Kolumny w panelu obszaru roboczego.
4. Przeciągnij zmienną *Wiek respondenta* do pola Wiersze na panelu obszaru roboczego.
5. Kliknij kartę **Testy**.
6. Zaznacz opcję **Porównaj średnie kolumnowe (testy t)**.
7. Wybierz opcję **Uwzględnij zmienne wielokrotnych odpowiedzi w teście**.
8. Kliknij przycisk **OK**, aby wykonać procedurę.

	News sources				
	Get news from newspapers	Get news from news magazines	Get news from television	Get news from radio	Get news from internet
	Mean	Mean	Mean	Mean	Mean
Age of respondent	52	40	48	40	40

Comparisons of Column Means

	News sources				
	Get news from newspapers	Get news from news magazines	Get news from television	Get news from radio	Get news from internet
	(A)	(B)	(C)	(D)	(E)
Age of respondent	B C D E		B D E		

Results are based on two-sided tests assuming equal variances with significance level 0.05. For each significant pair, the key of the smaller category appears under the category with larger mean.

Rysunek 74. Wyniki testu istotności

- Wszystkie kategorie zestawów wielokrotnych odpowiedzi opatrzone są literami (A, B, C, D, E) i dla każdej kategorii, dla której średnia innej kategorii jest zarówno niższa, jak i różni się znacząco od średniej danej kategorii, wyświetlana jest litera reprezentująca kategorię z niższą średnią.
- *Informacje czerpie z gazet codziennych* (A) ma najwyższą średnią wieku i wszystkie inne średnie kategorii znacząco się od niego różnią.
- *Informacje czerpie z telewizji* (C) ma niższą średnią wieku niż (A) i wszystkie inne średnie kategorii (B, D, i E) znacząco się od niego różnią. (C również znacząco różni się od A, co zaznaczono wcześniej).
- Średnie wieku dla *Informacje czerpie z magazynów* (B), *Informacje czerpie z radia* (D) i *Informacje czerpie z Internetu* (E) nie różnią się znacznie od siebie.

Braki danych

Wiele plików danych zawiera pewną ilość braków danych. Szeroki zakres czynników może powodować braki danych. Na przykład osoby ankietowane mogą nie odpowiedzieć na każde pytanie, pewne zmienne mogą nie mieć zastosowania do danych obserwacji, a błędy w kodowaniu mogą powodować odrzucanie niektórych wartości.

W programie IBM SPSS Statistics istnieją dwa rodzaje braków danych:

- **Zdefiniowane braki danych.** Wartości zdefiniowane jako zawierające braki danych. Do takich wartości można przypisać etykiety, określające powód braku danych (np. kod 99 i etykieta *Nie dotyczy* w przypadku pytania o ciążę u mężczyzn).
- **Systemowe braki danych.** Jeżeli dana zmienna numeryczna nie posiada wartości, zostanie jej przypisany systemowy brak danych. Jest on oznaczany kropką w widoku Dane w Edytorze danych.

Program udostępnia szereg funkcji ułatwiających łagodzenie skutków braku danych, a nawet umożliwiających analizę wzorów występujących w brakach danych. W tym rozdziale przyjęto znacznie prostszy cel, a mianowicie opisanie sposobu, w jaki moduł Tabele użytkownika obsługuje braki danych, i jak braki danych wpływają na obliczanie statystyk podsumowujących.

Przykładowy plik danych

W przykładach podanych w tym rozdziale wykorzystano plik danych *missing_values.sav*. Więcej informacji zawiera temat “Pliki przykładowe” na stronie 77. Jest to bardzo prosty, przykładowy plik danych, zawierający tylko jedną zmienną i dziesięć obserwacji, mający na celu zilustrowanie podstawowych zagadnień związanych z brakiem danych.

Tabele bez braków danych

Domyślnie kategorie zdefiniowanych braków danych nie są wyświetlane w tabelach użytkownika (podobnie jak systemowe braki danych).

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. W oknie konstruktora tabel przeciągnij zmienną *Zmienna z brakami danych* (jedyna zmienna w pliku) z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
3. Kliknij prawym przyciskiem myszy zmienną w panelu obszaru roboczego i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
4. W oknie dialogowym Kategorie i podsumowania kliknij (zaznacz) opcję **Podsumowanie ogółem**, a następnie kliknij przycisk **Zastosuj**.
5. Kliknij prawym przyciskiem myszy zmienną *Zmienna z brakami danych* na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.
6. W oknie dialogowym Statystyki podsumowujące, na liście Statystyki, zaznacz pozycję **% z N w kolumnie** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
7. Kliknij przycisk **Zastosuj do wybranej**.

   Pomiędzy kategoriami pokazywanymi na podglądzie tabeli w panelu obszaru roboczego a kategoriami wyświetlanymi na liście Kategorie (pod listą zmiennych, z lewej strony konstruktora tabel) można zauważyć niewielką rozbieżność. Lista Kategorie zawiera kategorię z etykietą *Braki danych* , która nie została uwzględniona na podglądzie tabeli, ponieważ kategorie z brakami danych są domyślnie wyłączone. Ponieważ słowo „danych” występuje w etykiecie w liczbie mnogiej, można wnioskować, że zmienna należy do dwóch lub więcej kategorii zdefiniowanych braków danych.
8. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Count	Column N %
Variable with missing values	Low	2	28.6%
	Medium	3	42.9%
	High	2	28.6%
	Total	7	100.0%

Rysunek 75. Tabela bez braków danych

W tej tabeli wszystko się zgadza. Wartości kategorii po zsumowaniu zgadzają się z wynikiem, a procenty dokładnie odpowiadają wartościom, które zostałyby otrzymane biorąc za podstawę obliczania wartości procentowej całkowitą liczbę obserwacji (np. 3/7= 0,429, czyli 42,9%). Jednak całkowita liczba obserwacji nie jest równa liczbie obserwacji z pliku danych; jest równa liczbie obserwacji **bez braków danych**, czyli liczbie obserwacji nie zawierających zdefiniowanych lub systemowych braków danych dla danej zmiennej.

Uwzględnianie braków danych w tabelach

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij prawym przyciskiem myszy zmienną *Zmienna z brakami danych* na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Kategorie i podsumowania**.
3. W oknie dialogowym Kategorie i podsumowania kliknij (zaznacz) opcję **Braki danych**, a następnie kliknij przycisk **Zastosuj**.

   Na podglądzie tabeli występuje teraz kategoria *Braki danych*. Chociaż na podglądzie tabeli widoczna jest tylko jedna kategoria braków danych, w tabeli wyświetlone zostają wszystkie zdefiniowane kategorie braków danych.
4. Kliknij prawym przyciskiem myszy zmienną *Zmienna z brakami danych* na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.
5. W oknie dialogowym Statystyki podsumowujące kliknij (zaznacz) opcję **Statystyki użytkownika dla podsumowań ogółem i sum pośrednich**.
6. Na liście podsumowujących statystyk użytkownika zaznacz pozycję **N ważnych** i kliknij przycisk ze strzałką, aby dodać ją do listy Wyświetlanie.
7. Powtórz czynność dla pozycji **Ogółem**.
8. Kliknij przycisk **Zastosuj do wybranej**, a następnie przycisk **OK** w konstruktorze tabel, aby utworzyć tabelę.

		Count	Column N %	Valid N	Total N
Variable with missing values	Low	2	22.2%		
	Medium	3	33.3%		
	High	2	22.2%		
	Don't know	1	11.1%		
	Not applicable	1	11.1%		
	Total	9	100.0%	7	10

Rysunek 76. Tabela z brakami danych

W tabeli wyświetlane są teraz dwie zdefiniowane kategorie braków danych--*Nie wiem* oraz *Nie dotyczy*, a całkowita liczba obserwacji wynosi 9 zamiast 7, ponieważ zostały uwzględnione dwie dodane obserwacje ze zdefiniowanymi brakami danych (po jednej w każdej z kategorii zdefiniowanych braków danych). Procenty w kolumnie również się różnią, ponieważ są oparte na liczbie wartości bez braków danych oraz ze zdefiniowanymi brakami danych. W obliczeniach procentów nie są jedynie uwzględnione systemowe braki danych.

Wartość *N ważnych* pokazuje całkowitą liczbę obserwacji bez braków danych (7), a wartość *Ogółem* całkowitą liczbę obserwacji, wraz z obserwacjami ze zdefiniowanymi i systemowymi brakami danych. Całkowita liczba obserwacji wynosi 10, o jedną więcej niż liczba obserwacji bez braków danych i ze zdefiniowanymi brakami danych, wyświetlana jako liczba całkowita w kolumnie *Liczebność*. Wynika to z faktu, że występuje jedna obserwacja z systemowym brakiem danych.
9. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).

10. Ponownie kliknij prawym przyciskiem myszy zmienną *Zmienna z brakami danych* na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.
11. Na górnej liście statystyk (a nie statystyk użytkownika dla podsumowań ogółem i sum pośrednich) zaznacz pozycję **% z N w kolumnie** i kliknij przycisk ze strzałką, aby dodać ją do listy Wyświetlanie.
12. Powtórz czynność dla pozycji **% z N ogółem w kolumnie**.
13. Można je również dodać do listy statystyk użytkownika dla podsumowań ogółem i sum pośrednich.
14. Kliknij przycisk **Zastosuj do wybranej**, a następnie przycisk **OK**, aby utworzyć tabelę.

		Count	Column N %	Column Valid N %	Column Total N %	Valid N	Total N
Variable with missing values	Low	2	22.2%	28.6%	20.0%		
	Medium	3	33.3%	42.9%	30.0%		
	High	2	22.2%	28.6%	20.0%		
	Don't know	1	11.1%	.0%	10.0%		
	Not applicable	1	11.1%	.0%	10.0%		
	Total	9	100.0%	100.0%	100.0%	7	10

Rysunek 77. Tabela z brakami danych oraz procentami obserwacji ważnych i ogółem

- Wartość *% N w kolumnie* stanowi procent obserwacji w każdej kategorii, z uwzględnieniem liczby obserwacji bez braków danych i ze zdefiniowanymi brakami danych (ponieważ zdefiniowane braki danych zostały dołączone do tabeli).
- *% z N ważnych w kolumnie* stanowi procent obserwacji w każdej kategorii, uwzględniający tylko ważne obserwacje bez braków danych. Wartości takie są równe procentom w kolumnie pierwotnej tabeli, nie uwzględniającej obserwacji ze zdefiniowanymi brakami danych.
- *% z N Ogółem w kolumnie* stanowi procent w każdej kategorii, uwzględniający obserwacje ze zdefiniowanymi i systemowymi brakami danych. Po dodaniu poszczególnych procentów w tej kategorii wynik wynosi tylko 90%, ponieważ jedna z 10 obserwacji (10%) zawiera systemowy brak danych. Pomimo faktu, iż obserwacja ta zostaje uwzględniona w bazie przy obliczaniu wartości procentowych, w tabeli nie jest przewidziana żadna kategoria dla obserwacji z systemowymi brakami danych.

Formatowanie i dostosowywanie tabel

Formatowanie i dostosowywanie tabel

Moduł Tabele użytkownika umożliwia definiowanie w procesie konstruowania tabel wielu właściwości ich formatowania, w tym:

- Formatu wyświetlania i etykiet statystyk podsumowujących
- Minimalnej i maksymalnej szerokości kolumny danych
- Tekstu lub wartości wyświetlanych w pustych komórkach

Ustawienia takie przechowywane są w obrębie interfejsu konstruktora tabel (do momentu ich zmiany, wyzerowania ustawień konstruktora lub otwarcia innego pliku danych), co umożliwia utworzenie wielu tabel o takich samych ustawieniach formatowania bez konieczności ich ręcznej edycji po utworzeniu. Ustawienia takie, wraz z wszystkimi pozostałymi parametrami tabel, można również zapisać, wklejając odpowiednią składnię komendy do okna edytora komend za pomocą przycisku Wklej, dostępnego w interfejsie konstruktora tabel, a następnie zapisując je w postaci pliku.

Wiele ustawień formatowania tabel można również zmienić po ich utworzeniu, wykorzystując wszystkie opcje formatowania dostępne w Edytorze raportów dla tabel przestawnych. Niniejszy rozdział poświęcony jest definiowaniu ustawień formatowania tabel przed ich utworzeniem. Więcej informacji na temat tabel przestawnych można uzyskać korzystając z karty Indeks w systemie pomocy i wpisując `tabele przestawne` jako słowo kluczowe.

Przykładowy plik danych

W przykładach podanych w niniejszym rozdziale wykorzystano plik danych *survey_sample.sav*. Więcej informacji zawiera temat “Pliki przykładowe” na stronie 77.

We wszystkich przedstawionych tutaj przykładach w oknach dialogowych wyświetlane są etykiety zmiennych, posortowane w porządku alfabetycznym. Opcje wyświetlania listy zmiennych ustawiane są na karcie Ogólne, w oknie dialogowym Opcje (menu Edycja, Opcje).

Format wyświetlania statystyk podsumowujących

Moduł Tabele użytkownika stara się przedstawiać statystyki podsumowujące w najodpowiedniejszych formatach domyślnych, tym niemniej w niektórych przypadkach warto je zmienić.

1. Z menu wybierz kolejno następujące pozycje:

   Analiza > **Tabele** > **Tabele użytkownika...**
2. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
3. Przeciągnij zmienną *Zaufanie do telewizji* pod zmienną *Kategoria wieku* w polu Wiersze, ustawiając je jedna na drugiej w wymiarze wierszowym.
4. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Zaznacz zmienne w wierszach**.
5. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* i z menu kontekstowego wybierz opcję **Kategorie i podsumowania**.
6. W oknie dialogowym Kategorie i podsumowania wybierz (zaznacz) opcję **Podsumowanie ogółem**, a następnie kliknij przycisk **Zastosuj**.
7. Kliknij prawym przyciskiem myszy dowolną zmienną na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.
8. Na liście Statystyki zaznacz pozycję **% z N w kolumnie** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
9. Zaznacz opcję **Statystyki użytkownika dla podsumowań ogółem i sum pośrednich**.
10. Na liście Statystyki dla statystyk podsumowujących użytkownika zaznacz pozycję **% z N w kolumnie** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
11. Powtórz czynność dla pozycji **Średnia**.
12. Następnie kliknij przycisk **Zastosuj do wszystkich**.

    Wartości zastępcze na podglądzie tabeli określają domyślny format wyświetlania każdej statystyki podsumowującej.

    - W przypadku liczb domyślnym formatem wyświetlania jest **nnnn**--liczby całkowite bez miejsc dziesiętnych.
    - W przypadku procentów domyślnym formatem wyświetlania jest **nnnn.n%**--liczby z jednym miejscem dziesiętnym i symbolem procentu.
    - W przypadku średnich domyślny format wyświetlania jest *inny* dla każdej z dwóch zmiennych.

    W przypadku statystyk podsumowujących innych niż liczebności (w tym N ważnych i N Ogółem) czy procenty, domyślnym formatem wyświetlania jest format zdefiniowany dla danej zmiennej w Edytorze danych. Patrząc na zmienne w widoku Zmienne w Edytorze danych widzimy, że zmienna *Kategoria wieku* (zmienna *katwiek*) jest zdefiniowana jako zmienna posiadająca dwa miejsca dziesiętne, natomiast zmienna *Zaufanie do telewizji* (zmienna *zauftv*) nie ma miejsc dziesiętnych.

    Jest to jeden z przypadków, gdy format domyślny jest prawdopodobnie nieodpowiedni, gdyż lepiej jest, kiedy obydwie wartości średnie są wyświetlane z taką samą liczbą miejsc dziesiętnych.
13. Kliknij prawym przyciskiem myszy dowolną zmienną na podglądzie tabeli w panelu obszaru roboczego i z menu podręcznego wybierz opcję **Statystyki podsumowujące**.

    W przypadku średniej komórka Format na liście Pokaż w tabeli wskazuje, że formatem jest *Auto*, co oznacza, że zastosowany zostanie format zdefiniowany dla zmiennej, a komórka Dziesiętne jest

nieaktywna. W celu określenia liczby miejsc dziesiętnych należy w pierwszej kolejności wybrać inny format.

14. Na liście Pokaż w tabeli dla statystyk użytkownika kliknij komórkę Format dla średniej i z rozwijanej listy formatów wybierz **nnnn**.
15. W komórce Dziesiętne wprowadź wartość 1.
16. Następnie kliknij przycisk **Zastosuj do wszystkich**, aby zastosować to ustawienie do obu zmiennych.

Na podglądzie tabeli widać, że wartości obu średnich będą wyświetlane z jednym miejscem dziesiętnym. (Można by teraz utworzyć tę tabelę, jednak wartość „średniej" dla zmiennej *Kategoria wieku* może okazać się trudna do interpretacji, gdyż kody liczbowe dla tej zmiennej mają zakres tylko od 1 do 6).

Etykiety statystyk podsumowujących

Oprócz formatów wyświetlania statystyk podsumowujących można również zmienić treść ich etykiet opisowych.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij przycisk **Resetuj**, aby usunąć poprzednie ustawienia w konstruktorze tabel.
3. W oknie konstruktora tabel przeciągnij zmienną *Kategoria wieku* z listy zmiennych do pola Wiersze w panelu obszaru roboczego.
4. Przeciągnij zmienną *Wypłata za ubiegły tydzień* z listy zmiennych do pola Kolumny w panelu obszaru roboczego.
5. Kliknij prawym przyciskiem myszy zmienną *Kategoria wieku* na podglądzie tabeli w panelu obszaru roboczego i z podręcznego menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
6. Na liście Statystyki zaznacz pozycję **% z N w kolumnie** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
7. Dwukrotnie kliknij wyrażenie *w kolumnie* w kolumnie Etykieta na liście Pokaż w tabeli, aby zmodyfikować zawartość komórki. Usuń z etykiety wyrażenie *w kolumnie*, pozostawiając jedynie *%*.
8. W podobny sposób zmodyfikuj zawartość komórki *Liczebność* w kolumnie Etykieta, zmieniając ją na *N*.

   Przy okazji można zmienić format statystyki % w kolumnie, usuwając zbędny symbol procentów (gdyż sama etykieta kolumny wskazuje, że zawiera ona procenty).
9. Kliknij komórkę Format dla statystyki *% z N w kolumnie* i z rozwijanej listy formatów wybierz **nnnn.n**.
10. Następnie kliknij przycisk **Zastosuj do wybranej**.

    Podgląd tabeli pokazuje zmodyfikowany format wyświetlania i zmodyfikowane etykiety.
11. Kliknij przycisk **OK**, aby utworzyć tabelę.

		How get paid last week											
		Hourly wage		Daily wage		Weekly wage		Monthly salary		Annual salary		Other pay rate	
		N	%	N	%	N	%	N	%	N	%	N	%
Age category	Less than 25	91	14.0	0	.0	12	9.7	3	2.0	7	3.1	14	7.7
	25 to 34	175	26.9	5	29.4	33	26.6	37	24.8	63	28.0	31	17.1
	35 to 44	185	28.5	5	29.4	42	33.9	45	30.2	66	29.3	61	33.7
	45 to 54	124	19.1	5	29.4	25	20.2	38	25.5	58	25.8	41	22.7
	55 to 64	52	8.0	0	.0	10	8.1	23	15.4	29	12.9	19	10.5
	65 or older	23	3.5	2	11.8	2	1.6	3	2.0	2	.9	15	8.3

Rysunek 78. Tabela ze zmodyfikowanymi etykietami statystyk podsumowujących

Szerokość kolumny

Nietrudno zauważyć, że tabela w poprzednim przykładzie jest dosyć szeroka. Jednym z możliwych rozwiązań tego problemu jest zamiana zmiennych w wierszach i kolumnach. Można również zwęzić kolumny, gdyż wydają się zbyt szerokie. (W tym celu skróciliśmy etykiety statystyk podsumowujących).

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij kartę **Opcje**.

3. W grupie Szerokość dla kolumn danych wybierz opcję **Użytkownika**.
4. W polu Maksimum wpisz 36. (Upewnij się, czy ustawienie Jednostki ma wartość **Punkty**).
5. Kliknij przycisk **OK**, aby utworzyć tabelę.

		How get paid last week											
		Hourly wage		Daily wage		Weekly wage		Monthly salary		Annual salary		Other pay rate	
		N	%	N	%	N	%	N	%	N	%	N	%
Age category	Less than 25	91	14.0	0	.0	12	9.7	3	2.0	7	3.1	14	7.7
	25 to 34	175	26.9	5	29.4	33	26.6	37	24.8	63	28.0	31	17.1
	35 to 44	185	28.5	5	29.4	42	33.9	45	30.2	66	29.3	61	33.7
	45 to 54	124	19.1	5	29.4	25	20.2	38	25.5	58	25.8	41	22.7
	55 to 64	52	8.0	0	.0	10	8.1	23	15.4	29	12.9	19	10.5
	65 or older	23	3.5	2	11.8	2	1.6	3	2.0	2	.9	15	8.3

Rysunek 79. Tabela ze zmniejszoną szerokością kolumn

Tabela jest teraz bardziej zwarta.

Wartości wyświetlane w pustych komórkach

Domyślnie w pustych komórkach (komórkach nie zawierających obserwacji) wyświetlana jest wartość 0. Można jednak sprawić, aby komórki takie pozostawały puste lub zdefiniować łańcuch tekstowy, który będzie w nich wyświetlany.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Kliknij kartę **Opcje**.
3. W grupie Wygląd komórek danych dla pustych komórek wybierz **Tekst** i wpisz `Brak`.
4. Kliknij przycisk **OK**, aby utworzyć tabelę.

		How get paid last week											
		Hourly wage		Daily wage		Weekly wage		Monthly salary		Annual salary		Other pay rate	
		N	%	N	%	N	%	N	%	N	%	N	%
Age category	Less than 25	91	14.0	None	None	12	9.7	3	2.0	7	3.1	14	7.7
	25 to 34	175	26.9	5	29.4	33	26.6	37	24.8	63	28.0	31	17.1
	35 to 44	185	28.5	5	29.4	42	33.9	45	30.2	66	29.3	61	33.7
	45 to 54	124	19.1	5	29.4	25	20.2	38	25.5	58	25.8	41	22.7
	55 to 64	52	8.0	None	None	10	8.1	23	15.4	29	12.9	19	10.5
	65 or older	23	3.5	2	11.8	2	1.6	3	2.0	2	.9	15	8.3

Rysunek 80. Tabela, w której pustych komórkach wyświetlane jest słowo „Brak”

Teraz w czterech pustych komórkach tabeli zamiast wartości 0 wyświetlane jest słowo *Brak*.

Wartości wyświetlane w przypadku braku statystyk

Jeśli nie można obliczyć statystyki, wartością wyświetlaną domyślnie jest kropka (.), będąca symbolem wskazującym systemowy brak danych. Przypadek ten różni się od przypadku „pustej” komórki i z tego względu wartość zastępująca brakujące statystyki jest ustawiana niezależnie od wartości wyświetlanych w komórkach nie zawierających obserwacji.

1. Otwórz konstruktora tabel (menu Analiza, Tabele specjalne, Tabele użytkownika).
2. Przeciągnij zmienną *Liczba godzin oglądania telewizji dziennie* z listy zmiennych i upuść u góry pola Kolumny, nad zmienną *Wypłata za ubiegły tydzień*.

   Ponieważ zmienna *Liczba godzin oglądania telewizji dziennie* jest zmienną ilościową, automatycznie staje się zmienną statystyki źródłowej, a statystyka podsumowująca zmienia się na średnią.
3. Kliknij prawym przyciskiem myszy zmienną *Liczba godzin oglądania telewizji dziennie* na podglądzie tabeli w panelu obszaru roboczego i z podręcznego menu kontekstowego wybierz opcję **Statystyki podsumowujące**.
4. Na liście Statystyki zaznacz pozycję **N ważnych** i kliknij przycisk ze strzałką, aby dodać ją do listy Pokaż w tabeli.
5. Kliknij przycisk **Zastosuj do wybranej**.
6. Kliknij kartę **Opcje**.
7. W polu tekstowym Oznaczenie statystyk, które nie mogą być wyliczone, wpisz `NA`.

8. Kliknij przycisk **OK**, aby utworzyć tabelę.

		Hours per day watching TV											
		How get paid last week											
		Hourly wage		Daily wage		Weekly wage		Monthly salary		Annual salary		Other pay rate	
		Mean	Valid N	Mean	Valid N	Mean	Valid N	Mean	Valid N	Mean	Valid N	Mean	Valid N
Age category	Less than 25	3	71	NA	None	3	10	2	3	2	6	2	8
	25 to 34	3	134	5	2	2	30	2	29	2	52	2	22
	35 to 44	3	136	2	5	3	30	2	34	2	47	3	46
	45 to 54	2	90	2	4	2	22	2	36	2	45	2	34
	55 to 64	3	40	NA	None	3	7	2	15	2	23	3	15
	65 or older	3	18	2	2	1	1	NA	0	1	2	3	11

Rysunek 81. Tabela z brakującymi statystykami oznaczonymi jako „NA”

Tekst *NA* jest wyświetlany dla średniej w trzech komórkach tabeli. W każdym przypadku odpowiednia wartość zmiennej *N ważnych* zawiera wyjaśnienie, dlaczego: Brak obserwacji, na podstawie których można obliczyć średnią.

Można jednak zauważyć pewną rozbieżność — jedna z trzech wartości statystyki N ważnych jest wyświetlana jako 0, a nie jako etykieta *Brak*, która powinna być wyświetlana w komórkach bez obserwacji. Wynika to z faktu, iż chociaż nie istnieją ważne obserwacje, z których można by wyliczyć średnią, kategoria nie jest właściwie pusta. Po powrocie do oryginalnej tabeli zawierającej tylko dwie zmienne kategorialne, widać, że w tej kategorii z tabeli krzyżowej występują trzy obserwacje. Nie ma jednak obserwacji ważnych, gdyż wszystkie trzy zawierają braki danych dla zmiennej ilościowej *Liczba godzin oglądania telewizji dziennie*.

Pliki przykładowe

Przykładowe pliki zainstalowane wraz z produktem można znaleźć w podkatalogu *Samples* w katalogu instalacyjnym. W podkatalogu Samples znajdują się osobne foldery dla każdego z następujących języków: angielskiego, francuskiego, niemieckiego, włoskiego, japońskiego, koreańskiego, polskiego, rosyjskiego, chińskiego uproszczonego, hiszpańskiego i chińskiego tradycyjnego.

Nie wszystkie przykładowe pliki są dostępne we wszystkich językach. Jeżeli przykładowy plik nie jest dostępny w danym języku, ten folder językowy zawiera angielską wersję przykładowego pliku.

Opisy

Poniżej znajdują się krótkie opisy przykładowych plików wykorzystywanych w różnych przykładach zamieszczonych w dokumentacji.

- **accidents.sav.** Jest to plik danych hipotetycznych dotyczący przedsiębiorstwa ubezpieczeniowego, które analizuje związane z wiekiem i płcią czynniki ryzyka wypadków samochodowych w danym regionie. Każda obserwacja odpowiada klasyfikacji krzyżowej kategorii wieku i płci.
- **adl.sav.** Jest to plik danych hipotetycznych dotyczący starań służących określeniu korzyści płynących z proponowanego typu terapii dla pacjentów po udarze. Lekarze losowo przypisali pacjentów po udarze do jednej z dwóch grup. W pierwszej grupie przeprowadzono standardową fizjoterapię, w drugiej zaś także dodatkową terapię psychologiczną. Trzy miesiące po leczeniu zdolności każdego pacjenta do wykonywania codziennych czynności zostały ocenione jako zmienne porządkowe.
- **advert.sav.** Jest to plik danych hipotetycznych dotyczący badań firmy handlowej nad relacją między nakładami finansowymi na reklamę a wynikami sprzedaży. W tym celu zgromadzono wyniki sprzedaży z przeszłości i powiązane z nimi koszty reklamy.
- **aflatoxin.sav.** Jest to plik danych hipotetycznych dotyczący badań nasion kukurydzy pod kątem obecności aflatoksyny, trucizny, której stężenie różni się znacznie między plonami zbóż i w ich obrębie. Zakład przetwórstwa zbóż otrzymał 16 próbek z każdego z 8 plonów i zmierzył poziom aflatoksyny w cząsteczkach na miliard (PPB).
- **anorectic.sav.** Pracując nad ustandaryzowaną symptomatologią zachowań anorektycznych i bulimicznych, [1] przebadali 55 nastolatków ze stwierdzonymi zaburzeniami odżywiania. Każdy pacjent był badany cztery razy w ciągu czterech lat, co dało łącznie 220 obserwacji. W każdej obserwacji u

pacjentów oceniano każdy z 16 symptomów. Brakuje ocen symptomów pacjenta 71 przy 2. badaniu, pacjenta 76 przy 2. badaniu i pacjenta 47 przy 3. badaniu, co daje 217 ważnych obserwacji.

- **bankloan.sav.** Ten plik danych hipotetycznych dotyczy dążeń banku do zmniejszenia wskaźnika niespłaconych kredytów. Plik zawiera informacje finansowe i demograficzne dotyczące 850 przeszłych i obecnych klientów. Pierwsze 700 obserwacji to klienci, którym przyznano już kredyty. Ostatnie 150 obserwacji to przyszli, potencjalni klienci, których bank musi sklasyfikować jako mających dobrą lub złą sytuację kredytową.
- **bankloan_binning.sav.** Jest to plik danych hipotetycznych, zawierający informacje finansowe i demograficzne o 5000 byłych klientów.
- **behavior.sav.** W klasycznym przykładzie [2] 52 studentów poproszono o ocenę kombinacji 15 sytuacji i 15 zachowań za pomocą 10-punktowej skali, od 0 = „wyjątkowo właściwe" do 9 = „wyjątkowo niewłaściwe." Wartości zostały uśrednione i zarejestrowane jako niepodobieństwa.
- **behavior_ini.sav.** Ten plik danych zawiera konfigurację początkową dwuwymiarowego rozwiązania zagadnienia z pliku *behavior.sav*.
- **brakes.sav.** Jest to plik danych hipotetycznych, który dotyczy kontroli jakości w zakładzie produkującym hamulce tarczowe dla samochodów wysokiej klasy. Plik danych zawiera wyniki pomiarów średnicy 16 tarcz hamulcowych pochodzących z każdej z 8 maszyn produkcyjnych. Docelowa średnica tarcz wynosi 322 milimetry.
- **breakfast.sav.** W klasycznym badaniu [3] 21 studentów MBA z Wharton School i ich małżonków poproszono o uszeregowanie 15 artykułów śniadaniowych według preferencji: od 1 = „najbardziej preferowany" do 15 = „najmniej preferowany". Preferencje zarejestrowano w ramach sześciu różnych scenariuszy, od „Preferencji ogólnej" do „Przekąska, tylko z napojem".
- **breakfast-overall.sav.** Ten plik danych zawiera preferencje dotyczące artykułów śniadaniowych tylko dla pierwszego scenariusza, „Preferencja ogólna".
- **broadband_1.sav.** Jest to plik danych hipotetycznych, zawierający liczbę abonentów (według regionu) krajowej stacji telewizyjnej. Plik danych zawiera miesięczne liczby abonentów z 85 regionów w okresie czterech lat.
- **broadband_2.sav.** Ten plik danych jest taki sam jak plik *broadband_1.sav*, ale zawiera dane za trzy dodatkowe miesiące.
- **car_insurance_claims.sav.** Zbiór danych przedstawiony i przeanalizowany w innej pracy [4] dotyczy roszczeń o odszkodowania motoryzacyjne. Średnią kwotę roszczenia można zamodelować jako zmienną o rozkładzie gamma, używając odwróconej funkcji łączenia w celu powiązania średniej zmiennej zależnej z liniową kombinacją wieku ubezpieczonego, typu pojazdu i jego wieku. Liczba wniesionych roszczeń może być wykorzystana jako waga skalująca.
- **car_sales.sav.** Ten plik danych zawiera hipotetyczne oszacowania sprzedaży, ceny i specyfikacje fizyczne różnych marek i modeli pojazdów. Ceny i specyfikacje fizyczne uzyskano naprzemiennie z witryny *edmunds.com* i z witryn producentów.
- **car_sales_uprepared.sav.** Jest to zmodyfikowana wersja pliku *car_sales.sav* , która nie obejmuje żadnych przetransformowanych wersji tych pól.
- **carpet.sav.** W popularnym przykładzie [5] spółka zainteresowana wprowadzeniem na rynek nowego środka do czyszczenia dywanów chce zbadać, jaki wpływ na preferencje klientów ma pięć czynników: wzornictwo opakowania, marka, cena, znak *Good Housekeeping* i gwarancja zwrotu pieniędzy. Istnieją trzy poziomy czynników dotyczące konstrukcji opakowań (różniących się położeniem aplikatora), trzy marki (*K2R*, *Glory* i *Bissell*), trzy poziomy cen i dwa poziomy (tak lub nie) dla każdego z dwóch ostatnich

[1] Van der Ham, T., J. J. Meulman, D. C. Van Strien i H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry,* 170, 363-368.

[2] Price, R. H., i D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology,* 30, 579-586.

[3] Green, P. E., i V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.

[4] McCullagh, P., i J. A. Nelder. 1989. *Generalized Linear Models*, wydanie drugie. London: Chapman & Hall.

[5] Green, P. E., i Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.

czynników. Dziesięciu klientów ocenia 22 profile zdefiniowane przez te czynniki. Zmienna *Preferencja* zawiera rangę średnich ocen każdego profilu. Niskie rangi odpowiadają wysokiej preferencji. Zmienna ta odzwierciedla ogólną miarę preferencji każdego z profili.

- **carpet_prefs.sav.** Ten plik danych bazuje na przykładzie omówionym w opisie pliku *carpet.sav*, ale zawiera rzeczywiste rangowanie uzyskane od każdego z 10 klientów. Klientów poproszono o uszeregowanie 22 profili produktów, od najbardziej do najmniej preferowanego. Zmienne od *PREF1* do *PREF22* zawierają identyfikatory powiązanych profili, zgodnie z definicją w pliku *carpet_plan.sav*.
- **catalog.sav.** Ten plik danych zawiera hipotetyczne miesięczne wielkości sprzedaży trzech produktów oferowanych przez firmę wydającą katalogi. Uwzględniono także dane pięciu możliwych predyktorów.
- **catalog_seasfac.sav.** Ten plik danych jest taki sam jak plik *catalog.sav*, ale dodano w nim zestaw czynników sezonowych obliczonych za pomocą procedury Dekompozycja sezonowa oraz towarzyszące im zmienne daty.
- **cellular.sav.** Jest to plik danych hipotetycznych dotyczący dążeń operatora telefonii komórkowej do zmniejszenia poziomu odejścia klientów. Oceny tendencji do odejścia są zastosowane do klientów porangowanych od 0 do 100. Klienci z oceną równą 50 lub wyższą mogą rozważać zmianę dostawcy usług.
- **ceramics.sav.** Jest to plik danych hipotetycznych dotyczący dążeń producenta do określenia, czy nowy stop wysokiej klasy ma większą odporność termiczną niż stop standardowy. Każda obserwacja odpowiada odrębnemu testowi jednego ze stopów i zawiera wartość temperatury, przy której łożysko uległo awarii.
- **cereal.sav.** Jest to plik danych hipotetycznych, który dotyczy badania 880 osób w zakresie ich preferencji śniadaniowych, rejestrującego także ich wiek, płeć, stan cywilny i informacje o aktywnym stylu życia (w zależności od tego, czy co najmniej dwa razy w tygodniu uprawiają sport). Każda obserwacja reprezentuje jednego respondenta.
- **clothing_defects.sav.** Jest to plik danych hipotetycznych, który dotyczy procesu kontroli jakości w fabryce odzieży. Z każdej partii produkowanej w zakładzie inspektorzy pobierają próbę odzieży i liczą, ile sztuk odzieży jest nie do przyjęcia.
- **coffee.sav.** Ten plik danych dotyczy postrzegania wizerunków sześciu marek kawy mrożonej [6]. Dla każdego z 23 atrybutów kawy mrożonej respondenci wybierali wszystkie marki, które były opisywane przez dany atrybut. Sześć marek oznaczono symbolami AA, BB, CC, DD, EE i FF, aby zachować poufność.
- **contacts.sav.** Jest to plik danych hipotetycznych, który dotyczy list kontaktowych dla grupy przedstawicieli handlowych sprzedających komputery dla firm. Każdy kontakt jest przydzielony do kategorii według działu firmy, w którym pracuje dany przedstawiciel, oraz jego pozycji w firmie. Zarejestrowano także kwotę ostatniej zawartej transakcji sprzedaży, czas, jaki upłynął od ostatniej transakcji, oraz wielkość firmy osoby kontaktowej.
- **creditpromo.sav.** Jest to plik danych hipotetycznych, który dotyczy dążeń domu towarowego do oszacowania skuteczności ostatniej promocji kart kredytowych. W tym celu wybrano losowo 500 posiadaczy kart. Połowa z nich otrzymała reklamę promującą obniżoną stopę procentową na zakupy dokonane w ciągu trzech następnych miesięcy. Druga połowa otrzymała standardową sezonową reklamę.
- **customer_dbase.sav.** Jest to plik danych hipotetycznych, dotyczący dążeń firmy do wykorzystania informacji znajdujących się w jej magazynie danych w celu złożenia specjalnych ofert klientom, w których przypadku prawdopodobieństwo odpowiedzi jest największe. Losowo wybrano podzbiór klientów z bazy klientów firmy, skierowano do nich oferty specjalne i zarejestrowano ich reakcje.
- **customer_information.sav.** Plik z hipotetycznymi danymi, zawierający informacje mailingowe klientów, takie jak nazwy i adresy.
- **customer_subset.sav.** Podzbiór 80 obserwacji z *customer_dbase.sav*.
- **debate.sav.** Jest to plik danych hipotetycznych, który dotyczy par odpowiedzi na ankietę przeprowadzaną wśród uczestników debaty politycznej przed debatą i po niej. Każda obserwacja odpowiada jednemu respondentowi.

[6] Kennedy, R., C. Riquier i B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing,* 5, 56-70.

- **debate_aggregate.sav.** Jest to plik danych hipotetycznych, w którym zagregowano odpowiedzi z pliku *debate.sav*. Każda obserwacja odpowiada klasyfikacji krzyżowej preferencji przed debatą i po niej.
- **demo.sav.** Jest to plik danych hipotetycznych, który dotyczy bazy danych klientów zakupionej do celów wysyłania comiesięcznych ofert. Zarejestrowano fakt, czy dany klient zareagował na ofertę, oraz różne informacje demograficzne.
- **demo_cs_1.sav.** Jest to plik danych hipotetycznych, który dotyczy pierwszego etapu dążeń pewnej firmy do skompilowania bazy danych informacji uzyskanych w ankiecie. Każda obserwacja odpowiada innemu miastu. Rejestrowane są region, prowincja, okręg i identyfikator miasta.
- **demo_cs_2.sav.** Jest to plik danych hipotetycznych, który dotyczy drugiego etapu dążeń pewnej firmy do skompilowania bazy danych informacji uzyskanych w ankiecie. Każda obserwacja odpowiada różnym jednostkom gospodarstw domowych z miast wybranych w pierwszym etapie; rejestrowane są region, prowincja, okręg, miasto, oddział i identyfikator jednostki. Uwzględniono także informacje na temat doboru próby z pierwszych dwóch etapów konstruowania bazy.
- **demo_cs.sav.** Jest to plik danych hipotetycznych, który zawiera wyniki ankiety zgromadzone przy użyciu złożonego modelu planu losowania. . Każda obserwacja odpowiada innej jednostce gospodarstwa domowego; rejestrowane są różne informacje demograficzne i dotyczące doboru próby.
- **diabetes_costs.sav.** Ten plik danych hipotetycznych zawiera informacje towarzystwa ubezpieczeniowego na temat osób ubezpieczonych chorych na cukrzycę. Każda obserwacja odpowiada jednemu ubezpieczonemu.
- **dietstudy.sav.** Ten plik danych hipotetycznych zawiera wyniki badania nad tzw. dietą Stillmana [7]. Każda obserwacja odpowiada odrębnemu badanemu i zawiera jego wagę w funtach przed dietą i po niej oraz poziom trójglicerydów we krwi w mg/100 ml.
- **dmdata.sav.** Jest to plik danych hipotetycznych, który zawiera informacje demograficzne i informacje dotyczące zakupów firmy zajmującej się marketingiem bezpośrednim. *dmdata2.sav* zawiera informacje o podzbiorze kontaktów, które otrzymały wiadomości testowe, a *dmdata3.sav* zawiera informacje o pozostałych kontaktach, które nie otrzymały wiadomości testowych.
- **dvdplayer.sav.** Jest to plik danych hipotetycznych dotyczący prac nad nowym odtwarzaczem płyt DVD. Używając prototypu urządzenia, dział marketingu zgromadził dane z grupy fokusowej. Każda obserwacja odpowiada jednemu badanemu użytkownikowi i zawiera pewne dane demograficzne o badanych oraz ich odpowiedzi na pytania dotyczące prototypu.
- **german_credit.sav.** Ten plik danych pochodzi ze zbioru danych „German credit", znajdującego się w Repository of Machine Learning Databases [8] na Uniwersytecie Kalifornijskim w Irvine.
- **grocery_1month.sav.** Ten plik danych hipotetyczny to plik danych *grocery_coupons.sav*, w którym tygodniowe wielkości zakupów zostały zagregowane tak, aby każda obserwacja odpowiadała jednemu klientowi. W rezultacie pewne zmienne, które zmieniały się w skali tygodnia, zniknęły, a zarejestrowana kwota wydatków jest teraz sumą wydatków w ciągu czterech tygodni badania.
- **grocery_coupons.sav.** Jest to plik danych hipotetycznych, który zawiera dane zebrane w ankiecie prowadzonej przez sieć sklepów spożywczych, zainteresowaną nawykami zakupowymi swoich klientów. Każdy klient jest badany przez cztery tygodnie, a każda obserwacja odpowiada odrębnemu tygodniowi tego klienta i zawiera obserwacje o tym, gdzie i kiedy dana osoba robi zakupy oraz ile wydała na artykuły spożywcze w danym tygodniu.
- **guttman.sav.** Bell [9] przedstawił tabelę ilustrującą możliwe grupy społeczne. Guttman [10] wykorzystał część tej tabeli, w której pięć zmiennych opisujących takie zagadnienia, jak interakcje społeczne, poczucie przynależności do grupy, fizyczna bliskość członków grupy i formalność relacji zostało skrzyżowanych z siedmioma teoretycznymi grupami społecznymi, obejmującymi tłum (np. widzów

[7] Rickman, R., N. Mitchell, J. Dingman i J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association,* 228:, 54-58.

[8] Blake, C. L. i C. J. Merz. 1998. "UCI Repository of machine learning databases." Pod adresemhttp://www.ics.uci.edu/~mlearn/MLRepository.html.

[9] Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.

[10] Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika,* 33, 469-506.

meczu piłki nożnej), publiczność (np. widzów w teatrze lub słuchaczy wykładu), publikę (np. grupę osób oglądających telewizję lub czytających gazetę), mob (podobny do tłumu, ale ze znacznie intensywniejszymi interakcjami), grupę pierwotną (intymną), grupę wtórną (dobrowolne) i nowoczesną społeczność (luźną grupę sprzymierzeńców, ukształtowaną przez bliskość fizyczną i zapotrzebowanie na specjalistyczne usługi).

- **health_funding.sav.** Jest to plik danych hipotetycznych, który zawiera dane na temat finansowania opieki zdrowotnej (kwoty na 100 członków populacji), wskaźników zachorowań (wskaźnik na 10 000 członków populacji) i wizyt w zakładach opieki zdrowotnej (wskaźnik na 10 000 członków populacji). Każda obserwacja reprezentuje inne miasto.
- **hivassay.sav.** Jest to plik danych hipotetycznych, który dotyczy dążeń laboratorium farmaceutycznego do opracowania szybkiej metody oznaczania próbek w celu wykrywania zakażenia wirusem HIV. Wyniki oznaczania to osiem ciemniejących odcieni czerwieni, gdzie ciemniejsze odcienie oznaczają większe prawdopodobieństwo zachorowania. Próba laboratoryjna została przeprowadzona na 2000 próbek krwi, z których połowa była zainfekowana wirusem HIV, druga połowa zaś wolna od wirusa.
- **hourlywagedata.sav.** Jest to plik danych hipotetycznych, który dotyczy wynagrodzeń godzinowych pielęgniarek zatrudnionych w przychodniach i szpitalach i mających różne poziomy doświadczenia.
- **insurance_claims.sav.** Jest to plik danych hipotetycznych, który dotyczy firmy ubezpieczeniowej, która chce zbudować model do flagowania podejrzanych, potencjalnie oszukańczych roszczeń. Każda obserwacja reprezentuje jedno roszczenie.
- **insure.sav.** Jest to plik danych hipotetycznych, który dotyczy firmy ubezpieczeniowej badającej czynniki ryzyka, wskazujących, czy klient będzie musiał wnieść roszczenie w ramach 10-letniej umowy ubezpieczenia na życie. Każda obserwacja w pliku danych reprezentuje parę umów, z których w jednej nastąpiło roszczenie, dopasowanych według wieku i płci.
- **judges.sav.** Jest to plik danych hipotetycznych, który dotyczy ocen przyznawanych przez wykwalifikowanych sędziów (i jednego ochotnika) 300 występom gimnastycznym. Każdy wiersz reprezentuje jeden występ; sędziowie oglądali te same występy.
- **kinship_dat.sav.** Rosenberg i Kim [11] postanowili przeanalizować 15 terminów związanych z pokrewieństwem (ciotka, brat, kuzyn, córka, ojciec, wnuczka, dziadek, babcia, wnuk, matka, bratanek, bratanica, siostra, syn, wuj). Poprosili cztery grupy studentów (dwie męskie i dwie żeńskie) o posortowanie tych terminów na podstawie podobieństw. Dwie grupy (jedną męską i jedną żeńską) poproszono o posortowanie terminów dwa razy, przy czym drugie sortowanie miało być oparte na innym kryterium niż pierwsze. W ten sposób uzyskano łącznie sześć „źródeł". Każde źródło odpowiada macierzy odległości 15 x 15, której komórki odpowiadają liczbie osób w źródle pomniejszonej o liczbę określającą, ile razy zmienne w tym źródle były łączone w podzbiory.
- **kinship_ini.sav.** Ten plik danych zawiera konfigurację początkową trójwymiarowego rozwiązania zagadnienia z pliku *kinship_dat.sav*.
- **kinship_var.sav.** Ten plik danych zawiera zmienne niezależne *gender* [płeć], *gener*(ation) [pokolenie] i *degree* (of separation) [stopień oddzielenia], które można wykorzystać do interpretacji wymiarów rozwiązania zagadnienia z pliku *kinship_dat.sav*. W szczególności można je wykorzystać do ograniczenia przestrzeni rozwiązania do liniowej kombinacji tych zmiennych.
- **marketvalues.sav.** Ten plik danych dotyczy sprzedaży domów na nowym osiedlu w Algonquin (USA) w latach 1999–2000. Dane o tej sprzedaży są ogólnodostępne.
- **nhis2000_subset.sav.** National Health Interview Survey (NHIS) to duże, oparte na całej populacji badanie stanu zdrowia ludności cywilnej w USA. Wywiady są przeprowadzane osobiście w reprezentatywnej w skali kraju próbie gospodarstw domowych. Dla każdego członka rodziny gromadzi się informacje i spostrzeżenia o zachowaniach dotyczących zdrowia i stanie zdrowia. Ten plik danych zawiera podzbiór informacji z badania przeprowadzonego w 2000 r. National Center for Health Statistics. National Health Interview Survey, 2000. Plik danych i dokumentacja do użytku publicznego. *ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/*. Dostęp uzyskano w 2003 r.
- **ozone.sav.** Dane uwzględniają 330 obserwacji sześciu zmiennych meteorologicznych, służących do przewidywania koncentracji ozonu na podstawie pozostałych zmiennych. Poprzedni badacze,

[11] Rosenberg, S., i M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research,* 10, 489-502.

m.in. [12], [13], znaleźli wśród tych zmiennych nieliniowości, które utrudniają zastosowanie standardowych podejść opartych na regresji.

- **pain_medication.sav.** Ten plik danych hipotetycznych zawiera wyniki badań klinicznych leku przeciwzapalnego służącego do leczenia chronicznego bólu stawów. Przedmiotem szczególnego zainteresowania jest czas, jaki potrzebny jest do działania leku, oraz jego porównanie z istniejącymi lekami.
- **patient_los.sav.** Ten plik danych hipotetycznych zawiera dane leczenia pacjentów przyjętych do szpitala z podejrzeniem zawału mięśnia sercowego („ataku serca"). Każda obserwacja odpowiada jednemu pacjentowi i zawiera wiele zmiennych związanych z jego pobytem w szpitalu.
- **patlos_sample.sav.** Ten plik danych hipotetycznych zawiera dane leczenia próby pacjentów, u których podczas leczenia zawału mięśnia sercowego („ataku serca") podawano leki trombolityczne (rozpuszczające skrzepy krwi). Każda obserwacja odpowiada jednemu pacjentowi i zawiera wiele zmiennych związanych z jego pobytem w szpitalu.
- **poll_cs.sav.** Jest to plik danych hipotetycznych, który dotyczy dążeń ankieterów państwowych do określenia poziomu wsparcia publicznego dla ustawy przed jej przyjéem przez zgromadzenie ustawodawcze. Obserwacje odpowiadają zarejestrowanym osobom głosującym. Każda obserwacja zawiera dane o okręgu, okręgu miejskim i osiedlu, w których dana osoba głosująca mieszka.
- **poll_cs_sample.sav.** Ten plik danych hipotetycznych zawiera próbę osób głosujących w wymienionych w pliku *poll_cs.sav*. Próbę dobrano zgodnie z wzorcem określonym w pliku *poll.csplan*, a w pliku danych zarejestrowano prawdopodobieństwo uwzględnienia i przykładowe wagi. Należy jednak pamiętać, że ponieważ plan dobierania próby wykorzystuje metodę prawdopodobieństwa proporcjonalnego do rozmiaru (PPS), istnieje także plik zawierający prawdopodobieństwa dla łącznego wyboru (*poll_jointprob.sav*). Dodatkowe zmienne odpowiadające danym demograficznym osób głosujących i ich opiniom o proponowanej ustawie zostały zebrane i dodane do pliku danych po doborze próby.
- **property_assess.sav.** Jest to plik danych hipotetycznych, które dotyczą dążeń asesora okręgu do utrzymania aktualności wycen wartości nieruchomości przy ograniczonych zasobach. Obserwacje odpowiadają nieruchomościom sprzedanym w obrębie hrabstwa w ostatnim roku. W każdej obserwacji w pliku danych zarejestrowano okręg miejski, w którym znajduje się nieruchomość, nazwisko asesora, który ostatnio odwiedzał nieruchomość, czas, jaki upłynął od tej wyceny, dokonaną wówczas wycenę wartości oraz wartość sprzedaży nieruchomości.
- **property_assess_cs.sav.** Jest to plik danych hipotetycznych, które dotyczą dążeń urzędnika stanu do utrzymania aktualności wycen wartości nieruchomości przy ograniczonych zasobach. Obserwacja odpowiadają nieruchomościom w obrębie stanu. W każdej obserwacji w pliku danych zarejestrowano okręg, okręg miejski i dzielnicę, w której znajduje się nieruchomość, czas, jaki upłynął od ostatniej wyceny, oraz dokonaną wówczas wycenę wartości.
- **property_assess_cs_sample.sav.** Ten plik danych hipotetycznych zawiera próbę nieruchomości wymienionych w pliku *property_assess_cs.sav*. Próbę dobrano zgodnie z wzorcem określonym w pliku *property_assess.csplan*, a w pliku danych zarejestrowano prawdopodobieństwo uwzględnienia i przykładowe wagi. Po dobraniu próby do pliku danych dodano nowo zgromadzoną zmienną, *Current value* (Wartość bieżąca).
- **recidivism.sav.** Jest to plik danych hipotetycznych, które dotyczą dążeń rządowej agencji organów ścigania do zrozumienia wskaźników recydywy w jej obszarze jurysdykcji. Każda obserwacja odpowiada jednej osobie, która wcześniej złamała prawo, i zawiera informacje demograficzne, pewne szczegóły pierwszego przestępstwa oraz czas, jaki upłynął do drugiego aresztowania, jeśli miało ono miejsce w ciągu dwóch lat od pierwszego aresztowania.
- **recidivism_cs_sample.sav.** Jest to plik danych hipotetycznych, które dotyczą dążeń rządowej agencji organów ścigania do zrozumienia wskaźników recydywy w jej obszarze jurysdykcji. Każda obserwacja odpowiada jednej osobie, która wcześniej złamała prawo, zwolnionej z pierwszego aresztu w czerwcu 2003 r., i zawiera jej informacje demograficzne, pewne szczegóły pierwszego przestępstwa oraz dane drugiego aresztowania, jeśli miało ono miejsce przed końcem czerwca 2006 r. Przestępców wybrano z

[12] Breiman, L., i J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association,* 80, 580-598.

[13] Hastie, T., i R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.

departamentów wyznaczonych zgodnie z planem dobierania próby określonym w pliku *recidivism_cs.csplan*; ponieważ wykorzystano w nim metodę prawdopodobieństwa proporcjonalnego do rozmiaru (PPS), istnieje także plik zawierający prawdopodobieństwa dla łącznego wyboru (*recidivism_cs_jointprob.sav*).

- **rfm_transactions.sav.** Plik z danymi hipotetycznymi, który zawiera dane dotyczące transakcji zamówień, w tym datę zamówienia, elementy zamawiane i wartość pieniężną każdej transakcji.
- **salesperformance.sav.** Jest to plik danych hipotetycznych dotyczący oceny dwóch nowych kursów dla sprzedawców. Sześćdziesięciu pracowników podzielonych na trzy grupy otrzymuje szkolenie standardowe. Dodatkowo grupa 2 uczestniczy w szkoleniu technicznym, a grupa 3 w kursie praktycznym. Na zakończenie kursu przetestowano wszystkich pracowników i zarejestrowano ich oceny. Każda obserwacja w pliku danych odpowiada jednemu kursantowi i zawiera informacje o grupie, do której dany kursant został przydzielony, oraz ocenę, jaką uzyskał na egzaminie.
- **satisf.sav.** Jest to plik danych hipotetycznych, który dotyczy badania zadowolenia klientów przeprowadzonego przez firmę handlową w 4 sklepach. Przebadano łącznie 582 klientów, a każda obserwacja reprezentuje odpowiedzi jednego klienta.
- **screws.sav.** Ten plik danych zawiera informacje o cechach śrub, wkrętów, nakrętek i gwoździ [14].
- **shampoo_ph.sav.** Jest to plik danych hipotetycznych, który dotyczy procesu kontroli jakości w fabryce produktów do pielęgnacji włosów. W regularnych odstępach czasu dokonywane są pomiary sześciu odrębnych partii produktów i rejestrowane jest ich pH. Docelowy zakres wartości to 4,5–5,5.
- **ships.sav.** Zbiór danych przedstawiony i zanalizowany w innej pracy [15], dotyczący uszkodzeń statków towarowych spowodowanych przez fale. Liczebność wypadków można zamodelować jako zmienną o rozkładzie Poissona, mając dany typ statku, okres jego budowy i serwisowania. Łączna liczba miesięcy pracy dla każdej komórki tabeli utworzonej przez klasyfikację krzyżową czynników stanowi wartość narażenia na ryzyko.
- **site.sav.** Jest to plik danych hipotetycznych, który dotyczy dążeń firmy do wybrania nowych obszarów rozwijania działalności. Firma zatrudniła dwóch konsultantów, aby odrębnie ocenili te lokalizacje. Oprócz rozszerzonego raportu, każdej lokalizacji przyznali oni sumaryczną ocenę, uznając ją za „dobrą”. „średnią” lub „kiepską” możliwość.
- **smokers.sav.** Ten plik danych pochodzi z przeprowadzonego w USA w 1998 r. krajowego badania gospodarstw domowych dotyczącego stosowania używek i stanowi prawdopodobną próbę amerykańskich gospodarstw domowych. (http://dx.doi.org/10.3886/ICPSR02934) Pierwszym krokiem analizy tego pliku danych powinno być więc ważenie danych tak, aby odzwierciedlały trendy widoczne w populacji.
- **stocks.sav** Ten plik danych hipotetycznych zawiera ceny akcji i ich ilość dla jednego roku.
- **stroke_clean.sav.** Ten plik danych hipotetycznych zawiera stan bazy danych medycznych po jej wyczyszczeniu za pomocą procedur z opcji Przygotowanie danych.
- **stroke_invalid.sav.** Ten plik danych hipotetycznych zawiera początkowy stan bazy danych medycznych, w tym kilka błędów we wprowadzaniu danych.
- **stroke_survival.** Ten plik danych hipotetycznych zawiera dane o czasie przeżycia pacjentów kończących program rehabilitacji pacjentów po udarze niedokrwiennym, stających przed różnymi wyzwaniami. Notowane są takie zdarzenia, jak udar wtórny, zawał serca, udar niedokrwienny lub krwotoczny, oraz czas ich wystąpienia. Próba jest obcięta z lewej strony, ponieważ obejmuje tylko pacjentów, którzy dożyli do końca programu rehabilitacji zaleconego po udarze.
- **stroke_valid.sav.** Ten plik danych hipotetycznych zawiera stan bazy danych medycznych po sprawdzeniu wartości za pomocą procedury Walidacja danych. Nadal zawiera obserwacje, które mogą być anomaliami.
- **survey_sample.sav.** Ten plik danych zawiera wyniki badania obejmujące dane demograficzne i różne miary poglądów. Bazuje on na podzbiorze zmiennych z ogólnych badań opinii społecznej, przeprowadzonych w 1998 roku przez NORC (amerykański ośrodek badania opinii społecznej), mimo to

[14] Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.
[15] McCullagh, P., i J. A. Nelder. 1989. *Generalized Linear Models*, wydanie drugie. London: Chapman & Hall.

niektóre wartości danych zostały zmodyfikowane i dodatkowe, fikcyjne zmienne zostały dodane do celów demonstracyjnych.

- **tcm_kpi.sav.** Jest to plik danych hipotetycznych zawierający wartości kluczowych wskaźników wydajności (KPI) przedsiębiorstwa za poszczególne tygodnie. Zawiera także dane różnych metryk podlegających kontroli — z poszczególnych tygodni w tym samym okresie, którego dotyczą dane o wskaźnikach wydajności.
- **tcm_kpi_upd.sav.** Ten plik danych jest taki sam jak plik *tcm_kpi.sav.sav*, ale zawiera dane za cztery dodatkowe tygodnie.
- **telco.sav.** Ten plik danych hipotetycznych dotyczy dążeń firmy telekomunikacyjnej do zmniejszenia poziomu odejścia w bazie klientów. Każda obserwacja odpowiada jednemu klientowi i zawiera różne informacje demograficzne oraz dane o wykorzystaniu usług.
- **telco_extra.sav.** Ten plik danych jest podobny do pliku danych *telco.sav*, ale usunięto z niego zmienne „tenure" (tytuł użytkowania) i dane wydatków przedstawione na przekształconej skali logarytmicznej, zastąpione ustandaryzowanymi zmiennymi wydatków klientów, przedstawionymi na przekształconej skali logarytmicznej.
- **telco_missing.sav.** Ten plik danych jest taki sam jak plik danych *telco.sav* , ale niektóre dane demograficzne zastąpiono w nim brakami danych.
- **testmarket.sav.** Ten plik danych hipotetycznych dotyczy planów sieci barów szybkiej obsługi co do wprowadzenia nowej pozycji w menu. Istnieją trzy możliwe kampanie promujące nowy produkt, nowa pozycja jest więc wprowadzana w oddziałach na kilku losowo wybranych rynkach. W każdym oddziale wykorzystuje się inną kampanię promocyjną i rejestruje się tygodniową wielkość sprzedaży nowego produktu. Każda obserwacja odpowiada jednemu tygodniowi w określonej lokalizacji.
- **testmarket_1month.sav.** Ten plik danych hipotetycznych to plik danych *testmarket.sav*, w którym tygodniowe wielkości sprzedaży zostały zagregowane tak, aby każda obserwacja odpowiadała jednej lokalizacji. W rezultacie niektóre zmienne, zmieniające się w skali tygodnia, zniknęły, a zarejestrowana wielkość sprzedaży jest teraz sumą sprzedaży w ciągu czterech tygodni badania.
- **tree_car.sav.** Jest to plik danych hipotetycznych zawierający dane demograficzne i informacje o cenie zakupu pojazdu.
- **tree_credit.sav.** Jest to plik danych hipotetycznych zawierający dane demograficzne i informacje o historii kredytowej.
- **tree_missing_data.sav** Jest to plik danych hipotetycznych zawierający dane demograficzne i informacje o historii kredytowej, z dużą liczbą braków danych.
- **tree_score_car.sav.** Jest to plik danych hipotetycznych zawierający dane demograficzne i informacje o cenie zakupu pojazdu.
- **tree_textdata.sav.** Prosty plik danych, zawierający tylko dwie zmienne, których pierwotnym przeznaczeniem było pokazanie domyślnego stanu zmiennych przed przypisaniem poziomu pomiaru oraz etykiet wartości.
- **tv-survey.sav.** Jest to plik danych hipotetycznych dotyczący badania przeprowadzonego przez studio telewizyjne, które rozważa przedłużenie emisji cieszącego się powodzeniem programu. 906 respondentów zapytano, czy oglądaliby ten program w różnych sytuacjach. Każdy wiersz reprezentuje jednego respondenta, a każda kolumna — jedną sytuację.
- **ulcer_recurrence.sav.** Ten plik zawiera częściowe informacje pochodzące z badania, którego celem było porównanie skuteczności dwóch terapii zapobiegających nawrotom wrzodów. Stanowi dobry przykład danych ocenzurowanych interwałowo i został przedstawiony oraz zanalizowany w innej pracy [16].
- **ulcer_recurrence_recoded.sav.** Ten plik zawiera zreorganizowane dane z pliku *ulcer_recurrence.sav*, aby umożliwić modelowanie prawdopodobieństwa zdarzeń dla każdego interwału badania, zamiast po prostu prawdopodobieństwa zdarzenia na zakończenie badania. Zagadnienie to zostało przedstawione i zanalizowane w innej pracy [17].

[16] Collett, D. 2003. *Modelling survival data in medical research*, wydanie drugie. Boca Raton: Chapman & Hall/CRC.

- **verd1985.sav.** Ten plik danych dotyczy badania [18]. Zarejestrowano w nim odpowiedzi 15 badanych na 8 zmiennych. Interesujące nas zmienne podzielono na trzy zestawy. Zestaw 1 obejmuje *wiek* i *stan cywilny*, zestaw drugi *zwierzę domowe* i *wiadomości*, a zestaw trzeci *muzykę* i *na żywo*. Zmienna *zwierzę* jest typu wielokrotnego nominalnego, a *wiek* typu porządkowego. Wszystkie pozostałe zmienne są typu jednokrotnego nominalnego.
- **virus.sav.** Jest to plik danych hipotetycznych, który dotyczy dążenia dostawcy usług internetowych (ISP) do określenia wpływu wirusów na jego sieci. Prześledzono (szacowany) udział procentowy ruchu zainfekowanych wiadomości e-mail w sieci dostawcy w pewnym okresie, od momentu wykrycia wirusa do powstrzymania zagrożenia.
- **wheeze_steubenville.sav.** Jest to podzbiór danych z długotrwałego badania wpływu zanieczyszczenia powietrza na zdrowie dzieci [19]. Dane zawierają binarne pomiary sapania dzieci w wieku 7, 8, 9 i 10 lat z miejscowości Steubenville w stanie Ohio, a także stałe informacje o tym, czy matka dziecka paliła papierosy w pierwszym roku badania.
- **workprog.sav.** Jest to plik danych hipotetycznych, który dotyczy programu prac rządowych mającego na celu ułatwienie osobom niepełnosprawnym zdobycia lepszych miejsc pracy. Prześledzono próbę potencjalnych uczestników programu; niektórzy zostali losowo wybrani do udziału w programie, inni zaś nie. Każda obserwacja reprezentuje jednego uczestnika programu.
- **worldsales.sav** Ten plik danych hipotetycznych zawiera przychód ze sprzedaży według kontynentu i produktu.

[17] Collett, D. 2003. *Modelling survival data in medical research*, wydanie drugie. Boca Raton: Chapman & Hall/CRC.

[18] Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (w jęz. niderlandzkim)*. Leiden: Department of Data Theory, University of Leiden.

[19] Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer i B. G. Ferris Jr. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases,* 129, 366-374.

Informacje

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.

IBM może nie oferować w innych krajach produktów, usług lub opcji omawianych w tej publikacji. Informacje o produktach i usługach dostępnych w danym kraju można uzyskać od lokalnego przedstawiciela IBM. Jakakolwiek wzmianka na temat produktu, programu lub usługi IBM nie oznacza, że może być zastosowany jedynie ten produkt, ten program lub ta usługa 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 tej 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 typograficzne. 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 podmiotów 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 podmiotów 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.

LICENCJA W ZAKRESIE PRAW AUTORSKICH:

Niniejsza publikacja zawiera przykładowe aplikacje w kodzie źródłowym, ilustrujące techniki programowania w różnych systemach operacyjnych. Użytkownik może kopiować, modyfikować i rozpowszechniać te programy przykładowe w dowolnej formie bez uiszczania opłat na rzecz IBM, w celu rozbudowy, użytkowania, handlowym lub w celu rozpowszechniania aplikacji zgodnych z aplikacyjnym interfejsem programowym dla tego systemu operacyjnego, dla którego napisane były programy przykładowe. Programy przykładowe nie zostały gruntownie przetestowane. IBM nie może zatem gwarantować ani sugerować niezawodności, użyteczności i funkcjonalności tych programów. Programy przykładowe są dostarczane w stanie, w jakim się znajdują ("AS IS"), bez udzielania jakichkolwiek gwarancji (rękojmię również wyłącza się). IBM nie ponosi odpowiedzialności za jakiekolwiek szkody wynikające z używania programów przykładowych.

Każda kopia programu przykładowego lub jakikolwiek jego fragment, jak też jakiekolwiek prace pochodne muszą zawierać następujące uwagi dotyczące praw autorskich:

Znaki towarowe

IBM, logo IBM oraz ibm.com są znakami towarowymi lub zastrzeżonymi znakami towarowymi International Business Machines Corp. zarejestrowanymi w wielu systemach prawnych na całym świecie. Nazwy innych produktów i usług mogą być znakami towarowymi IBM lub innych podmiotów. 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 oraz Pentium są znakami towarowymi lub zastrzeżonymi znakami towarowymi Intel Corporation lub przedsiębiorstw podporządkowanych Intel Corporation w Stanach Zjednoczonych i/lub 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 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 Oracle.

Indeks

B

braki danych
  uwzględnianie w tabelach użytkownika 72
  wpływ na obliczenia procentów 72

C

chi-kwadrat
  tabele użytkownika 54
czas
  uwzględnianie bieżącego czasu w tabelach użytkownika 17

D

data
  uwzględnianie bieżącej daty w tabelach użytkownika 17
dominanta
  tabele użytkownika 9
drukowanie tabel z warstwami 30

E

etykiety
  zmiana tekstu etykiet statystyk podsumowujących 75
etykiety zmiennych
  ukrywanie w tabelach użytkownika 5

F

formaty wyświetlania
  statystyki podsumowujące w tabelach użytkownika 11, 74

K

kategorie po przeliczeniu
  tabele użytkownika 14, 35

L

liczebność
  a N ważnych 49

M

maksimum
  tabele użytkownika 9
mediana
  tabele użytkownika 9
miejsca dziesiętne
  ustawianie liczby miejsc dziesiętnych wyświetlanych w tabelach użytkownika 6, 22, 74
minimum
minimum *(kontynuacja)*
  tabele użytkownika 9

N

N ważnych
  tabele użytkownika 9
nagłówki
  tabele użytkownika 17
narożniki
  tabele użytkownika 17

O

odchylenie standardowe
  tabele użytkownika 9
Ogółem N 72

P

pliki przykładowe
  lokalizacja 77
podsumowania
  miejsce wyświetlania 32
  podsumowania ogółem w grupie 32
  sumy brzegowe w tabelach użytkownika 22
  tabele użytkownika 11, 21, 30
  tabele zagnieżdżone 32
  warstwy 33
  wykluczone kategorie 31
podsumowania grupowe
  zmienne ilościowe 51
podsumowania ogółem w grupie 32
podsumowania ogółem w podgrupie 32
pomijanie kategorii
  tabele użytkownika 23
poziom pomiaru
  zmiana w tabelach użytkownika 1
procenty
  braki danych 72
  w tabelach użytkownika 7, 8, 20, 21
  zestawy wielokrotnych odpowiedzi 8
przedziały ufności 53
przeliczone kategorie
  formaty wyświetlania 15
  tabele użytkownika 14, 35
  ukrywanie kategorii w wyrażeniu 36
  z sum pośrednich 37
przetwarzanie podzielonych danych
  tabele użytkownika 4
puste komórki
  wyświetlana wartość w tabelach użytkownika 15, 76

R

rozstęp

rozstęp *(kontynuacja)*
tabele użytkownika 9
różne statystyki podsumowujące dla różnych zmiennych
zestawione tabele 50

S

sortowanie kategorii
tabele użytkownika 23
statystyki
statystyki podsumowania wybierane przez użytkownika 46
statystyki podsumowujące 43
zestawione tabele 45
statystyki dla proporcji kolumnowych
tabele użytkownika 59
statystyki dla średnich kolumnowych
tabele użytkownika 56
statystyki podsumowania wybierane przez użytkownika 46
statystyki podsumowujące
format wyświetlania 74
różne podsumowania dla różnych zmiennych w zestawionych tabelach 50
statystyki podsumowania wybierane przez użytkownika 46
wymiar źródłowy 43
zestawione tabele 45
zmiana tekstu etykiet 75
zmienna źródłowa 43
suma
tabele użytkownika 9
sumy pośrednie
tabele użytkownika 11, 30
ukrywanie kategorii z sumami pośrednimi 34
systemowe braki danych 71
szerokość kolumny
zmiana w tabelach użytkownika 15, 75

Ś

średnia
tabele użytkownika 9

T

tabela krzyżowa
tabele użytkownika 21
tabele
tabele użytkownika 1
tabele częstości
tabele użytkownika 15, 39
tabele częstości średnich 10, 46
tabele porównawcze 15, 39
tabele użytkownika
drukowanie tabel warstwowych 30
etykiety wartości zmiennych kategorialnych 1
formaty wyświetlania 6
formaty wyświetlania statystyk podsumowujących 11
kategorie po przeliczeniu 14, 35
nagłówki 17
narożniki 17
podsumowania 11, 21, 30
tabele użytkownika *(kontynuacja)*
podsumowania w tabelach z wykluczonymi kategoriami 23
procent w wierszu a procent w kolumnie 20
procenty 7, 8, 20, 21
procenty dla zestawów wielokrotnych odpowiedzi 8
proste tabele dla zmiennych kategorialnych 19
przeliczone kategorie 11, 14, 35
przetwarzanie podzielonych danych 4
puste komórki 15
sortowanie kategorii 23
statystyki podsumowania wybierane przez użytkownika 10
statystyki podsumowujące 7–9
sumy brzegowe 22
sumy pośrednie 11, 30
szerokość kolumny 15
tabela krzyżowa 21
tabele częstości 15, 39
tabele częstości średnich 10
tabele porównawcze 15, 39
tabele zmiennych zawierających wspólne kategorie 15, 39
testowanie istotności i wielokrotne odpowiedzi 64
testy 17, 53
tworzenie tabeli 3
tytuły 17
ukrywanie etykiet statystyk 19
ukrywanie kategorii z sumami pośrednimi 34
widok kompaktowy 27
wykluczanie braków danych w zmiennych ilościowych 15
wykluczanie kategorii 11, 23
wymiar statystyki źródłowej 21
wyświetlanie lub ukrywanie nazw i etykiet zmiennych 5
zagnieżdżanie zmiennych 26, 27
zagnieżdżanie zmiennych w warstwach 30
zamiana miejscami zmiennych w wierszach i kolumnach 28
zestawy wielokrotnych kategorii 15
zestawy wielokrotnych odpowiedzi 1, 64
zmiana etykiet statystyk podsumowujących 20
zmiana kolejności kategorii 11
zmiana liczby wyświetlanych miejsc dziesiętnych 6
zmiana poziomu pomiaru 1
zmiana wymiaru statystyk podsumowujących 10
zmienne ilościowe 1
zmienne kategorialne 1
zmienne w warstwie 28–30
zmienne zestawiające 25
zwijanie kategorii 34
testy
tabele użytkownika 17, 53
testy istotności
tabele użytkownika 17
zestawy wielokrotnych odpowiedzi 69, 70
tytuły
tabele użytkownika 17

U

ukrywanie etykiet statystyk w tabelach użytkownika 19
usuwanie kategorii
tabele użytkownika 11, 23

W

wariancja
    tabele użytkownika 9
wartości
    wyświetlanie etykiet i wartości kategorii 47
wartości i etykiety wartości 47
wykluczanie kategorii
    tabele użytkownika 11, 23
wyświetlanie wartości kategorii 47

Z

zagnieżdżanie zmiennych
    tabele użytkownika 26, 27
    zmienne ilościowe 52
zdefiniowane braki danych 71
zestawy wielokrotnych odpowiedzi
    powtórzone odpowiedzi w zestawach wielokrotnych kategorii 15
    procenty 8
    testowanie istotności 64, 69, 70
zmiana kolejności kategorii
    tabele użytkownika 11
zmiana liczby wyświetlanych miejsc dziesiętnych 22
zmienna podsumowująca statystyki źródłowej
    zmienne ilościowe 52
zmienne ilościowe
    podsumowania grupowe 51
    podsumowania pogrupowane według zmiennych jakościowych w wierszach i kolumnach 51
    statystyki podsumowujące 48
    wielokrotne statystyki podsumowujące 48
    zagnieżdżanie 52
    zestawianie 48
zmienne w warstwie
    drukowanie tabel warstwowych 30
    tabele użytkownika 28–30
    zagnieżdżanie zmiennych w warstwach 30
    zestawianie zmiennych w warstwach 29
zmienne zestawiające
    różne statystyki podsumowujące dla różnych zmiennych 50
    tabele użytkownika 25
    wiele zmiennych podsumowujących statystyki źródłowej 45
    zestawianie zmiennych w warstwach 29
    zmienne ilościowe 48
zwijanie kategorii
    tabele użytkownika 34