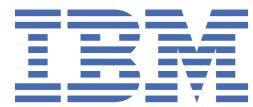


IBM SPSS Modeler 18.3 사용자 안내서



참고

이 정보와 이 정보가 지원하는 제품을 사용하기 전에, [235 페이지의 『주의사항』](#)에 있는 정보를 확인하십시오.

제품 정보

이 개정판은 새 개정판에 별도로 명시하지 않는 한, IBM® SPSS® Modeler 18, 릴리스 3, 수정사항 0 및 모든 후속 릴리스와 수정에 적용됩니다.

© Copyright International Business Machines Corporation .

목차

제 1 장 IBM SPSS Modeler 정보	1
IBM SPSS Modeler 제품.....	1
IBM SPSS Modeler	1
IBM SPSS Modeler Server	1
IBM SPSS Modeler Administration Console	1
IBM SPSS Modeler Batch	2
IBM SPSS Modeler Solution Publisher	2
IBM SPSS Collaboration and Deployment Services용 IBM SPSS Modeler Server 어댑터	2
IBM SPSS Modeler 에디션.....	2
문서.....	3
SPSS Modeler Professional 문서.....	3
SPSS Modeler Premium 문서.....	3
애플리케이션 예제.....	3
Demos 폴더.....	4
라이센스 추적.....	4
제 2 장 IBM SPSS Modeler 18.3.0의 새로운 기능.....	5
제 3 장 제품 개요.....	7
시작하기.....	7
IBM SPSS Modeler 시작	7
명령행에서 시작.....	7
IBM SPSS Modeler Server에 연결	8
Analytic Server에 연결	9
임시 디렉토리 변경.....	10
다중 IBM SPSS Modeler 세션 시작.....	10
IBM SPSS Modeler 인터페이스 살펴보기.....	11
IBM SPSS Modeler 스트림 캔버스.....	11
노드 팔레트.....	12
IBM SPSS Modeler 관리자.....	13
IBM SPSS Modeler 프로젝트.....	14
IBM SPSS Modeler 도구 모음.....	15
도구 모음 사용자 정의.....	16
IBM SPSS Modeler 창 사용자 정의.....	16
스트림 아이콘 크기 변경.....	17
IBM SPSS Modeler에서 마우스 사용	18
단축키 사용.....	18
인쇄.....	19
IBM SPSS Modeler 자동화	19
제 4 장 데이터 마이닝 이해.....	21
데이터 마이닝 개요.....	21
데이터 평가.....	22
데이터 마이닝 전략.....	23
CRISP-DM 프로세스 모델.....	23
모델 유형.....	24
데이터 마이닝 예제.....	28
제 5 장 스트림 작성.....	29
스트림 작성 개요.....	29

데이터 스트림 작성.....	29
노드에 대한 작업.....	29
스트림에 대한 작업.....	34
스트림 설명.....	44
스트림 실행.....	45
모델에 대한 작업.....	46
노드와 스트림에 주석과 주석(Annotation) 추가.....	46
데이터 스트림 저장.....	51
파일 로드.....	52
데이터 스트림 맵핑.....	53
팁 및 단축키.....	54
제 6 장 데이터에 대한 작업.....	57
차트 작성.....	57
레이아웃 및 용어.....	57
차트 유형 갤러리에서 차트 작성.....	58
차트 유형.....	58
대시보드.....	91
글로벌 시각화 기본 설정.....	91
제 7 장 출력 작업.....	93
뷰어.....	93
결과 표시 및 숨김.....	93
출력 이동, 삭제, 복사.....	94
초기 맞춤 변경.....	94
출력 항목의 맞추기 변경.....	94
뷰어 개요.....	94
뷰어에 항목 추가 및 편집.....	95
뷰어에서 찾기 및 바꾸기.....	96
다른 애플리케이션으로 출력 복사.....	97
대화형 출력.....	98
출력 내보내기.....	98
HTML 옵션.....	99
웹 보고서 옵션.....	100
Word 옵션.....	100
Excel 옵션.....	100
PowerPoint 옵션.....	101
PDF 옵션.....	101
텍스트 옵션.....	102
이미지만 옵션.....	103
그래픽 형식 옵션.....	103
뷰어 인쇄.....	104
출력과 차트 인쇄.....	104
인쇄 미리보기.....	104
페이지 속성: 머리말 및 꼬리말.....	105
페이지 속성: 옵션.....	105
출력 저장.....	106
뷰어 문서 저장.....	106
피벗 테이블.....	107
피벗 테이블.....	107
피벗 테이블 조작.....	107
레이어 작업.....	110
항목 표시 및 숨김.....	110
테이블모양.....	111
테이블 특성.....	112
셀 특성.....	114
꼬리말 및 캡션.....	114

데이터 셀 너비.....	116
열 너비 변경.....	116
피벗 테이블에서 숨겨진 경계선 표시.....	116
피벗 테이블에서 행, 열과 셀을 선택.....	116
피벗 테이블 인쇄.....	117
피벗 테이블에서 차트 작성.....	118
레거시 테이블.....	118
옵션	118
옵션.....	118
일반 옵션	118
부어 옵션.....	119
피벗 테이블 옵션.....	119
출력 옵션.....	120

제 8 장 결측값 처리..... 121

결측값의 개요.....	121
결측값 처리.....	121
결측값이 있는 레코드 처리.....	122
결측값이 있는 필드 처리.....	122
시스템 결측값이 있는 레코드 처리.....	122
결측값 대치 또는 채우기.....	124
CLEM 함수 결측값.....	125

제 9 장 CLEM 표현식 작성..... 127

CLEM 정보.....	127
CLEM 예제.....	127
값 및 데이터 유형.....	128
표현식 및 조건.....	129
스트림, 세션 및 수퍼노드 모수.....	130
문자열에 대한 작업.....	130
공백 및 결측값 처리.....	131
숫자에 대한 작업.....	131
시간 및 날짜에 대한 작업.....	131
다중 필드 요약.....	132
다중반응 데이터에 대한 작업.....	133
표현식 작성기.....	134
표현식 작성기에 액세스.....	134
표현식 작성.....	134
함수 선택.....	134
필드, 모수 및 글로벌 변수 선택.....	138
값 보기 또는 선택.....	138
CLEM 표현식 검사.....	138
찾기 및 바꾸기.....	139

제 10 장 CLEM 언어 참조..... 143

CLEM 참조 개요.....	143
CLEM 데이터 유형.....	143
정수.....	143
실수.....	143
문자.....	144
문자열.....	144
목록.....	144
필드.....	144
날짜.....	144
시간.....	145
CLEM 연산자.....	146
함수 참조.....	148

함수 설명 규칙.....	149
정보 함수.....	150
변환 함수.....	150
비교 함수.....	151
논리 함수.....	153
숫자 함수.....	153
삼각 함수.....	155
확률 함수.....	155
공간 함수.....	156
Bitwise 정수 연산.....	156
랜덤 함수.....	157
문자열 함수.....	158
SoundEx 함수.....	163
날짜 및 시간 함수.....	163
시퀀스 함수.....	167
글로벌 함수.....	172
공백 및 널값 처리 함수.....	173
특수 필드.....	173
제 11 장 리포지토리에서 IBM SPSS Modeler 사용.....	175
IBM SPSS Collaboration and Deployment Services Repository 정보.....	175
리포지토리 오브젝트 저장 및 배포.....	175
리포지토리에 연결.....	176
리포지토리의 신임 정보 입력.....	176
리포지토리 신임 정보 찾기.....	176
리포지토리 내용 찾아보기.....	176
리포지토리에 오브젝트 저장.....	177
오브젝트 특성 설정.....	177
스트림 저장.....	179
프로젝트 저장.....	179
노드 저장.....	179
출력 오브젝트 저장.....	180
모델 및 모델 팔레트 저장.....	180
리포지토리에서 오브젝트 검색.....	180
검색할 오브젝트 선택.....	181
오브젝트 버전 선택하기.....	181
리포지토리에서 오브젝트 검색.....	181
리포지토리 오브젝트 수정.....	182
폴더 생성, 이름 변경 및 삭제.....	183
리포지토리 오브젝트 잠금 및 잠금 해제.....	183
리포지토리 오브젝트 삭제.....	183
리포지토리 오브젝트의 특성 관리.....	183
폴더 특성 보기.....	184
오브젝트 특성 보기 및 편집.....	184
오브젝트 버전 레이블 관리.....	185
스트림 배포.....	185
스트림 배포 옵션.....	186
스코어링 분기.....	187
제 12 장 IBM Cloud Pak for Data에 스트림 저장.....	191
제 13 장 외부 애플리케이션으로 내보내기.....	193
외부 애플리케이션으로 내보내기 정보.....	193
스트림을 IBM SPSS Modeler Advantage에서 열기.....	193
모델을 PMML로 가져오기 및 내보내기.....	193
PMML을 지원하는 모델 유형.....	194

제 14 장 프로젝트 및 보고서.....	197
프로젝트에 대한 소개.....	197
CRISP-DM 보기.....	197
클래스 보기.....	198
프로젝트 작성.....	198
새 프로젝트 작성.....	198
프로젝트에 추가.....	198
프로젝트를 IBM SPSS Collaboration and Deployment Services Repository에 전송.....	199
프로젝트 특성 설정.....	199
프로젝트에 주석달기.....	200
오브젝트 특성.....	200
프로젝트 닫기.....	201
보고서 생성.....	201
생성된 보고서 저장 및 내보내기.....	202
제 15 장 IBM SPSS Modeler 사용자 정의	205
IBM SPSS Modeler 옵션 사용자 정의.....	205
IBM SPSS Modeler 옵션 설정.....	205
시스템 옵션.....	205
기본 디렉토리 설정.....	206
사용자 옵션 설정.....	206
노드 팔레트 사용자 정의.....	213
팔레트 관리자 사용자 정의.....	214
팔레트 탭 보기 변경.....	216
제 16 장 스트림과 노드의 성능 고려사항.....	217
노드의 순서.....	217
노드 캐시.....	218
성능: 프로세스 노드.....	218
성능: 모델링 노드.....	219
성능: CLEM 표현식.....	220
제 17 장 IBM SPSS Modeler의 내게 필요한 옵션.....	221
IBM SPSS Modeler의 내게 필요한 옵션 개요	221
내게 필요한 옵션 지원 유형.....	221
시각 장애인을 위한 내게 필요한 옵션.....	221
맹인 사용자를 위한 내게 필요한 옵션.....	222
내게 필요한 옵션의 키보드 기능.....	222
스크린 리더 사용.....	229
사용 팁.....	230
다른 소프트웨어와의 충돌.....	230
JAWS 및 Java.....	230
IBM SPSS Modeler에서 그래프 사용	231
제 18 장 유니코드 지원.....	233
IBM SPSS Modeler에서 유니코드 지원	233
주의사항	235
상표.....	236
제품 문서의 이용 약관.....	236
색인.....	239

제 1 장 IBM SPSS Modeler 정보

IBM SPSS Modeler는 비즈니스 전문 지식을 사용하여 예측 모형을 신속하게 개발하고 이를 비즈니스 운영에 배포하여 의사결정의 정확성을 향상시켜주는 데이터 마이닝 도구 세트입니다. 산업 표준 CRISP-DM 모델을 중심으로 디자인된 IBM SPSS Modeler는 데이터에서 보다 나아진 비즈니스 결과에 이르는 전체 데이터 마이닝 프로세스를 지원합니다.

IBM SPSS Modeler는 기계 학습, 인공지능 및 통계로부터 취한 다양한 모델링 방법을 제공합니다. 모델링 팔레트에서 사용할 수 있는 이러한 방법을 통해 데이터로부터 새로운 정보를 얻어서 예측 모형을 개발할 수 있습니다. 각각의 방법은 그것만의 장점이 있으며 특정한 문제점 유형에 가장 적합합니다.

SPSS Modeler는 독립형 제품으로 구매하거나 SPSS Modeler Server와 통합하여 클라이언트로 사용할 수 있습니다. 다음 절에 요약된 바와 같이 여러가지 추가 옵션도 사용할 수 있습니다. 자세한 정보는 <https://www.ibm.com/analytics/us/en/technology/spss/>의 내용을 참조하십시오.

IBM SPSS Modeler 제품

IBM SPSS Modeler 제품군 및 연관 소프트웨어는 다음으로 구성됩니다.

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console(IBM SPSS Deployment Manager와 함께 포함)
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- IBM SPSS Collaboration and Deployment Services용 IBM SPSS Modeler Server 어댑터

IBM SPSS Modeler

SPSS Modeler는 개인용 컴퓨터에 설치하여 실행되는 기능적으로 완벽한 버전의 제품입니다. 로컬 모드에서 독립형 제품으로 SPSS Modeler를 실행하거나 대형 데이터 세트에 대한 성능 향상을 위해 분산 모드에서 IBM SPSS Modeler Server와 함께 사용할 수 있습니다.

SPSS Modeler를 사용하여 프로그래밍하지 않고 신속하게 직관적으로 정확한 예측 모델을 작성할 수 있습니다. 고유한 시각적 인터페이스를 사용하면 데이터 마이닝 프로세스를 쉽게 시각화할 수 있습니다. 제품에 포함된 고급 분석 지원을 통해 데이터에서 이전에 숨겨진 패턴과 추세를 발견할 수 있습니다. 결과를 모델링하고 결과에 영향을 주는 요인을 이해하여 비즈니스 기회를 활용하고 위험을 줄일 수 있습니다.

SPSS Modeler는 두 개의 에디션(SPSS Modeler Professional과 SPSS Modeler Premium)으로 사용할 수 있습니다. 자세한 정보는 2 페이지의 [『IBM SPSS Modeler 에디션』](#)의 내용을 참조하십시오.

IBM SPSS Modeler Server

SPSS Modeler는 클라이언트/서버 아키텍처를 사용하여 자원 집약적 작업에 대한 요청을 강력한 서버 소프트웨어로 분배하여 대형 데이터 세트에 대한 성능을 향상시킵니다.

SPSS Modeler Server는 하나 이상의 IBM SPSS Modeler 설치와 함께 서버 호스트의 분산 분석 모드에서 계속해서 실행되는 별도로 라이센스가 부여된 제품입니다. 이런 방법으로 클라이언트 컴퓨터로 데이터를 다운로드하지 않고 서버에서 메모리 집약적 작업을 수행할 수 있기 때문에 SPSS Modeler Server는 대형 데이터 세트에 대한 우수한 성능을 제공합니다. 또한 IBM SPSS Modeler Server는 SQL 최적화 및 In-Database 모델링 기능에 대한 지원을 제공하여 성능 및 자동화의 이점도 추가로 제공합니다.

IBM SPSS Modeler Administration Console

Modeler Administration Console은 옵션 파일을 통해서도 구성 가능한 다수의 SPSS Modeler Server 구성 옵션을 관리하기 위한 그래픽 사용자 인터페이스입니다. 콘솔은 IBM SPSS Deployment Manager에 포함되며, SPSS Modeler Server 설치를 모니터하고 구성하는 데 사용할 수 있고, 현재 SPSS Modeler Server 고객이 무료로 사용

할 수 있습니다. 이 애플리케이션은 Windows 컴퓨터에만 설치할 수 있지만 지원되는 플랫폼에 설치된 서버를 관리할 수 있습니다.

IBM SPSS Modeler Batch

데이터 마이닝은 일반적으로 대화식 처리인 반면, 그래픽 사용자 인터페이스가 없어도 명령행에서 SPSS Modeler를 실행할 수 있습니다. 예를 들어, 사용자 개입 없이 수행할 장기 실행 또는 반복 작업이 있습니다. SPSS Modeler Batch는 정규 사용자 인터페이스에 대한 액세스 없이 SPSS Modeler의 전체 분석 기능에 대한 지원을 제공하는 특수 버전의 제품입니다. SPSS Modeler Batch를 사용하려면 SPSS Modeler Server가 필요합니다.

IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher는 외부 런타임 엔진을 통해 실행하거나 외부 애플리케이션에 포함될 수 있는 SPSS Modeler 스트림의 패키지 버전을 작성할 수 있게 하는 도구입니다. 이런 방법으로 SPSS Modeler가 설치되지 않는 환경에 사용할 수 있도록 전체 SPSS Modeler 스트림을 출판하고 배포할 수 있습니다. SPSS Modeler Solution Publisher는 별도의 라이센스가 필요한 IBM SPSS Collaboration and Deployment Services - Scoring 서비스의 일부로 분배됩니다. 이 라이센스가 있으면 출판된 스트림을 실행할 수 있게 하는 SPSS Modeler Solution Publisher Runtime을 수신합니다.

SPSS Modeler Solution Publisher에 대한 자세한 정보는 IBM SPSS Collaboration and Deployment Services 문서를 참조하십시오. IBM SPSS Collaboration and Deployment Services IBM Documentation에는 "IBM SPSS Modeler Solution Publisher" 및 "IBM SPSS Analytics Toolkit" 섹션이 포함되어 있습니다.

IBM SPSS Collaboration and Deployment Services용 IBM SPSS Modeler Server 어댑터

SPSS Modeler와 SPSS Modeler Server가 IBM SPSS Collaboration and Deployment Services 리포지토리와 상호작용할 수 있게 하는 IBM SPSS Collaboration and Deployment Services용 어댑터를 상당수 사용할 수 있습니다. 이런 방법으로 리포지토리에 배포된 SPSS Modeler 스트림을 여러 사용자가 공유하거나 씬 클라이언트 애플리케이션 IBM SPSS Modeler Advantage에서 액세스할 수 있습니다. 리포지토리를 호스팅하는 시스템에 어댑터를 설치하십시오.

IBM SPSS Modeler 에디션

SPSS Modeler는 다음 에디션으로 사용할 수 있습니다.

SPSS Modeler Professional

SPSS Modeler Professional은 CRM 시스템, 인구 통계, 구매 동작, 판매 데이터에서 추적된 동작 및 상호작용과 같은 대부분의 구조화된 데이터 유형에 대한 작업을 하는 데 필요한 모든 도구를 제공합니다.

SPSS Modeler Premium

SPSS Modeler Premium은 특수 데이터 및 비정형 텍스트 데이터에 대한 작업을 할 수 있도록 SPSS Modeler Professional을 확장하는 별도로 라이센스가 부여된 제품입니다. SPSS Modeler Premium에는 IBM SPSS Modeler Text Analytics가 포함됩니다.

IBM SPSS Modeler Text Analytics는 고급 언어 기술 및 자연어 처리(NLP)를 사용하여 다양한 비정형 텍스트 데이터를 빠르게 처리하고, 주요 개념을 추출 및 구성하며, 이러한 개념을 범주로 분류합니다. 추출된 개념과 범주는 인구 통계와 같은 기준의 구조화된 데이터와 결합할 수 있고 보다 나은 집중적인 의사결정을 내리기 위해 전체 IBM SPSS Modeler 데이터 마이닝 세트를 사용하여 모델링에 적용할 수 있습니다.

IBM SPSS Modeler Subscription

IBM SPSS Modeler Subscription에서는 일반적인 IBM SPSS Modeler 클라이언트와 동일한 모든 예측 분석 공정능력을 제공합니다. Subscription 에디션을 사용하면 제품 업데이트를 정기적으로 다운로드할 수 있습니다.

문서

문서는 SPSS Modeler의 도움말 메뉴에서 사용할 수 있습니다. 이는 제품 외부에서 항상 사용 가능한 온라인 IBM Documentation을 엽니다.

설치 지시사항을 포함하여 각 제품에 대한 전체 문서 또한 제품 다운로드의 일부로 별도의 압축 폴더에 PDF 형식으로 제공됩니다. 또는 웹(<https://www.ibm.com/support/pages/spss-modeler-183-documentation>)에서 최신 PDF 문서를 다운로드할 수 있습니다.

SPSS Modeler Professional 문서

SPSS Modeler Professional 문서 스위트(설치 지시사항은 제외)는 다음과 같습니다.

- **IBM SPSS Modeler 사용자 안내서.** SPSS Modeler 사용에 관한 일반적인 소개이며, 데이터 스트림 작성, 결측값 처리, CLEM 표현식 작성 프로젝트 및 보고서에 대한 작업, IBM SPSS Collaboration and Deployment Services 또는 IBM SPSS Modeler Advantage에 배포하기 위한 스트림 패키지 방법이 포함됩니다.
- **IBM SPSS Modeler 소스, 프로세스 및 출력 노드.** 여러 형식의 데이터를 읽고 처리하며, 출력하는 데 사용하는 모든 노드에 대한 설명입니다. 실질적으로 이는 모델링 노드 이외의 모든 노드를 의미합니다.
- **IBM SPSS Modeler 모델링 노드.** 데이터 마이닝 모델을 작성하는 데 사용하는 모든 노드에 대한 설명입니다. IBM SPSS Modeler는 기계 학습, 인공지능 및 통계로부터 취한 다양한 모델링 방법을 제공합니다.
- **IBM SPSS Modeler 애플리케이션 안내서.** 이 안내서의 예제는 특정 모델링 방법과 기법을 중점적으로 간략히 소개합니다. 이 안내서의 온라인 버전은 도움말 메뉴에서도 사용할 수 있습니다. 자세한 정보는 [3 페이지의 『애플리케이션 예제』](#) 주제를 참조하십시오.
- **IBM SPSS Modeler Python 스크립팅 및 자동화.** 노드와 스트림을 조작하는 데 사용할 수 있는 특성을 포함하여 Python 스크립팅을 통한 시스템 자동화에 대한 정보입니다.
- **IBM SPSS Modeler 배포 안내서.** IBM SPSS Deployment Manager에서 작업 처리 단계로 IBM SPSS Modeler 스트림 실행에 대한 정보입니다.
- **IBM SPSS Modeler CLEF 개발자 안내서.** CLEF는 데이터 처리 루틴 또는 모델링 알고리즘과 같은 써드파티 프로그램을 IBM SPSS Modeler의 노드로 통합하는 기능을 제공합니다.
- **IBM SPSS Modeler In-Database 마이닝 안내서.** 데이터베이스의 능력을 사용하여 성능을 향상시키고 써드파티 알고리즘을 통해 분석 기능 범위를 확장하는 방법에 대한 정보입니다.
- **IBM SPSS Modeler Server 관리 및 성능 안내서.** IBM SPSS Modeler Server 구성 및 관리 방법에 대한 정보입니다.
- **IBM SPSS Deployment Manager 사용자 안내서.** IBM SPSS Modeler Server 모니터링 및 구성용 Deployment Manager 애플리케이션에 포함된 관리 콘솔 사용자 인터페이스를 사용하는 데 대한 정보
- **IBM SPSS Modeler CRISP-DM 안내서.** SPSS Modeler에서 데이터 마이닝에 CRISP-DM 방법론을 사용하기 위한 단계별 안내서입니다.
- **IBM SPSS Modeler Batch 사용자 안내서.** 일괄처리 모드 실행 및 명령행 인수 세부사항을 포함하여 일괄처리 모드에서 IBM SPSS Modeler 사용을 위한 전체 안내서입니다. 이 안내서는 PDF 형식으로만 사용할 수 있습니다.

SPSS Modeler Premium 문서

SPSS Modeler Premium 문서 스위트(설치 지시사항은 제외)는 다음과 같습니다.

- **SPSS Modeler Text Analytics 사용자 안내서.** SPSS Modeler에서 텍스트 분석 사용에 대한 정보로, 텍스트 마이닝 노드, 대화식 워크벤치, 템플리트 및 기타 자원에 대해 설명합니다.

애플리케이션 예제

SPSS Modeler의 데이터 마이닝 도구가 광범위한 비즈니스 및 조직의 문제점을 해결하는 데 도움을 주는 가운데, 애플리케이션 예제는 특정 모델링 방법 및 기술에 대해 대상화된 간략한 소개를 제공합니다. 여기서 사용된 데이터 세트는 일부 데이터 마이너에서 관리하는 거대한 데이터 스토어보다 훨씬 작지만, 관련된 개념과 방법은 실제 애플리케이션으로 확장 가능합니다.

예제에 액세스하려면 SPSS Modeler의 도움말 메뉴에서 **애플리케이션 예제**를 클릭하십시오.

데이터 파일 및 샘플 스트림은 제품 설치 디렉토리 아래에 있는 Demos 폴더에 설치됩니다. 추가 정보는 4 페 이지의 [『Demos 폴더』](#)의 내용을 참조하십시오.

데이터베이스 모델링 예제. *IBM SPSS Modeler In-Database* 마이닝 안내서의 예제를 참조하십시오.

스크립팅 예제. *IBM SPSS Modeler* 스크립팅 및 자동화 안내서의 예제를 참조하십시오.

Demos 폴더

애플리케이션 예제에서 사용하는 데이터 파일 및 샘플 스트림은 제품 설치 디렉토리 아래의 Demos 폴더(예: C:\Program Files\IBM\SPSS\Modeler\<version>\Demos)에 설치됩니다. Windows 시작 메뉴의 IBM SPSS Modeler 프로그램 그룹에서, 또는 파일 > 스트림 열기 대화 상자의 최근 디렉토리 목록에서 Demos를 클릭해서도 이 폴더에 액세스할 수 있습니다.

라이센스 추적

SPSS Modeler를 사용할 때 라이센스 사용이 정기적으로 추적되고 로그됩니다. 로그되는 라이센스 메트릭은 AUTHORIZED_USER 및 CONCURRENT_USER이며 로그되는 메트릭의 유형은 SPSS Modeler에 대해 가진 라이센스의 유형에 의해 결정됩니다.

생성되는 로그 파일은 사용자가 라이센스 사용 보고서를 생성할 수 있는 IBM 라이센스 메트릭 도구에 의해 처리될 수 있습니다.

라이센스 로그 파일은 SPSS Modeler 클라이언트 로그 파일이 기록되는 디렉토리와 동일한 디렉토리(기본적으로 %ALLUSERSPROFILE%\IBM\SPSS\Modeler\<version>/log)에 작성됩니다.

제 2 장 IBM SPSS Modeler 18.3.0의 새로운 기능

IBM SPSS Modeler는 이 릴리스에서 다음과 같은 기능을 추가합니다.

- 모델링의 불편함은 시간이 지나면 데이터가 변경되기 때문에 모델이 구식이 되는 것입니다. 이를 일반적으로 모델 드리프트 또는 개념 드리프트라고 합니다. 모델 드리프트를 효과적으로 극복하기 위해 SPSS Modeler는 연속 자동 기계 학습을 제공합니다. IBM의 연구 결과, 생물학의 자연 선택에서 영감을 받은 연속 기계 학습을 자동 분류자 노드와 자동 숫자 노드에 제공합니다.
- 이제 IBM SPSS Modeler 클라이언트에서 직접 스트림을 IBM Cloud Pak for Data Server에 업로드할 수 있습니다. 추가 정보는 [191 페이지의 『제 12 장 IBM Cloud Pak for Data에 스트림 저장』](#)의 내용을 참조하십시오.
- IBM SPSS Modeler 클라이언트에 대해 로깅을 사용으로 설정하려면 텍스트 편집기에서 log4j2.xml 파일을 열고 다음 행에서 `level="info"`를 `level="debug"`로 변경하십시오.

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

- 음이 아닌 최소 제곱(NNLS) 수행**이라는 새로운 설정이 모수 추정값 아래 작성 옵션 탭의 GLE 노드에 추가되었습니다. 음이 아닌 최소 제곱(NNLS)은 계수가 음수가 될 수 없는 제약된 최소 제곱 문제 유형입니다. 모든 데이터 세트가 NNLS에 적합한 것은 아니므로, 예측변수와 목표값 간에 양의 상관 관계가 필요하거나 아예 필요하지 않습니다.
- 열과 유형에 대해 스캔할 행**이라는 새 설정이 Excel 소스 노드에 추가되었습니다. 이 새 설정의 기본값은 200입니다. 열 유형 및 저장 유형을 판별하기 위해 Excel 데이터의 더 많은 행을 스캔하도록 이 값을 늘려야 하는 경우 성능이 영향을 받을 수 있습니다.
- [Planning Analytics on Cloud](#)를 지원하기 위해 새로운 연결 방법(TM1 서버 연결)을 IBM Cognos TM1 소스 노드와 IBM Cognos TM1 내보내기 노드에 사용할 수 있습니다. 관리 서버는 Planning Analytics on Cloud에서 제거되었으므로, TM1 관리 서버를 통해 Planning Analytics on Cloud에 연결되는 TM1 노드가 이전 스트림에 포함되어 있는 경우 해당 스트림이 TM1 서버를 대신 가리키도록 수정할 수 있습니다.
- Db2 11.5가 이제 지원됩니다.
- Db2 Warehouse가 이제 지원됩니다.
- Db2 Big SQL 7.1.0 on Cloudera Data Platform 7.1.5가 이제 지원됩니다.
- Apache Hive 3.1.3 on Cloudera Data Platform 7.1.5가 이제 지원됩니다.
- Cloudera Impala 3.4.0 on Cloudera Data Platform 7.1.5가 이제 지원됩니다.
- Informix 14.10이 이제 지원됩니다.
- RedHat 8.3이 이제 지원됩니다.
- 새 버전의 R이 사용됩니다(4.0.4).

제 3 장 제품 개요

시작하기

데이터 마이닝 애플리케이션으로서 IBM SPSS Modeler는 대규모 데이터 세트에서 유용한 관계를 찾기 위한 전략적 접근 방식을 제공합니다. 보다 전통적인 통계 방법과는 대조적으로 시작할 때 찾아야 하는 것을 꼭 알고 있을 필요가 없습니다. 유용한 정보를 찾을 때 까지 서로 다른 모델을 맞춰보고 서로 다른 관계를 조사하면서 데이터를 탐색할 수 있습니다.

IBM SPSS Modeler 시작

애플리케이션을 시작하려면 다음을 클릭하십시오.

시작 > [모든 프로그램] > **IBM SPSS Modeler <버전>** > **IBM SPSS Modeler <버전>**

몇 초 후에 기본 창이 표시됩니다.

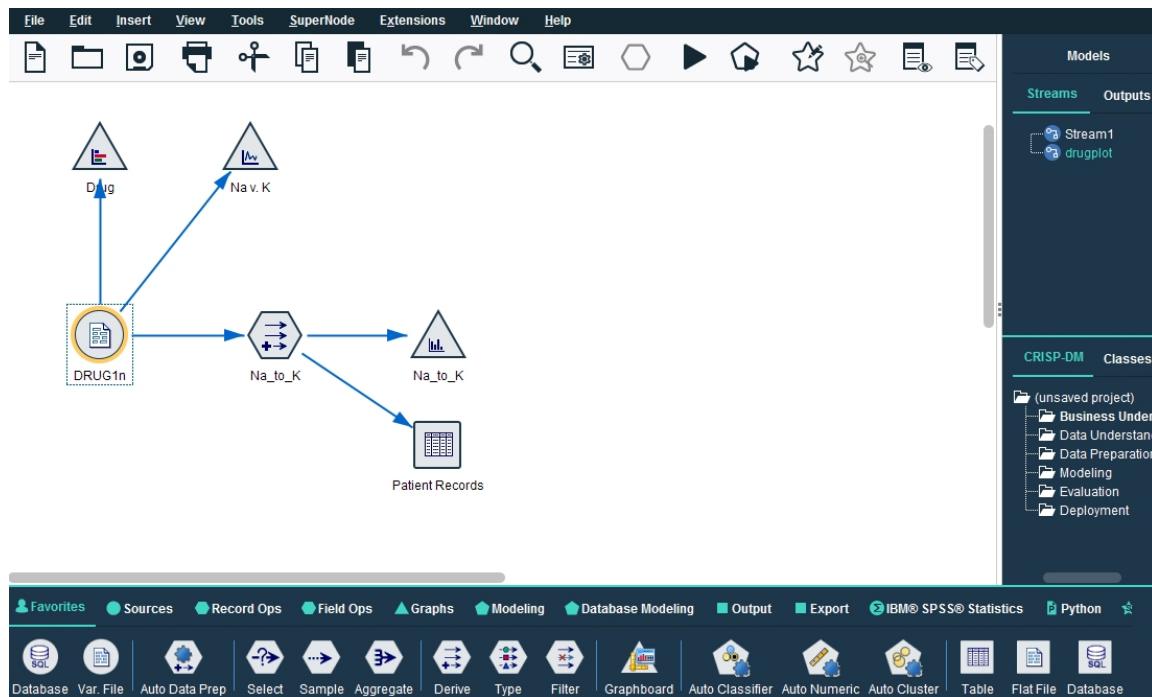


그림 1. IBM SPSS Modeler 기본 애플리케이션 창

명령행에서 시작

운영 체제의 명령행을 사용하여 다음과 같이 IBM SPSS Modeler를 시작할 수 있습니다.

1. IBM SPSS Modeler가 설치된 컴퓨터에서 DOS 또는 명령 프롬프트 창을 여십시오.
2. 대화식 모드로 IBM SPSS Modeler 인터페이스를 시작하려면 `modelerclient` 명령 뒤에 필수 인수를 입력하십시오. 예를 들어, 다음과 같습니다.

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

사용 가능한 인수(플래그)를 사용하면 필요에 따라 서버에 연결하고 스트림을 로드하며, 스크립트를 실행하거나 다른 모수를 지정할 수 있습니다.

IBM SPSS Modeler Server에 연결

IBM SPSS Modeler는 독립형 애플리케이션으로서나 IBM SPSS Modeler Server에 직접 또는 IBM SPSS Modeler Server 또는 IBM SPSS Collaboration and Deployment Services로부터 프로세스 조정자 플러그인을 통해 서버 군집에 연결된 클라이언트로서 실행할 수 있습니다. 현재 연결 상태는 IBM SPSS Modeler 창의 왼쪽 아래에 표시됩니다.

서버에 연결하려고 할 때마다 연결하려는 서버 이름을 수동으로 입력하거나 이전에 정의한 이름을 선택할 수 있습니다. 그러나 IBM SPSS Collaboration and Deployment Services가 있는 경우 서버 로그인 대화 상자에서 서버 목록이나 서버 군집을 통해 검색할 수 있습니다. 네트워크에서 실행 중인 Statistics 서비스를 통해 찾아보는 기능은 프로세스 조정자를 통해 사용 가능하게 됩니다.

서버에 연결하기

1. 도구 메뉴에서 **서버 로그인**을 클릭하십시오. 서버 로그인 대화 상자가 열립니다. 또는 IBM SPSS Modeler 창의 연결 상태 영역을 두 번 클릭하십시오.
2. 대화 상자를 사용하여 로컬 서버 컴퓨터에 연결하기 위한 옵션을 지정하거나 테이블에서 연결을 선택하십시오.
 - 연결을 추가하거나 편집하려면 **추가** 또는 **편집**을 클릭하십시오. 자세한 정보는 [8 페이지의 『IBM SPSS Modeler Server 연결 추가 및 편집』](#)의 내용을 참조하십시오.
 - 프로세스 조정자에서 서버 또는 서버 군집에 액세스하려면 **검색**을 클릭하십시오. 자세한 정보는 [9 페이지의 『IBM SPSS Collaboration and Deployment Services에서 서버 검색』](#)의 내용을 참조하십시오.

서버 테이블. 이 테이블에는 정의된 서버 연결 세트가 포함됩니다. 테이블은 기본 연결, 서버 이름, 설명 및 포트 번호를 표시합니다. 새 연결을 수동으로 추가하고 기존 연결을 선택하거나 검색할 수도 있습니다. 특정 서버를 기본 연결로 설정하려면 연결의 테이블에서 기본 열의 확인 상자를 선택하십시오.

기본 데이터 경로. 서버 컴퓨터에서 데이터에 사용된 경로를 지정하십시오. 필요한 위치로 이동하려면 생략 기호 단추(...)를 클릭하십시오.

자격 설정. 로컬 컴퓨터 사용자 이름과 비밀번호 세부사항을 사용하여 서버에 로그인하려고 시도하는 **싱글 사용자** 기능을 사용하려면 이 상자를 선택 해제 상태로 두십시오. 싱글 사용자로 가능하지 않거나 싱글 사용자를 사용 안함으로 설정하기 위해 이 상자를 선택한 경우에는(예: 관리자 계정에 로그인), 신임 정보를 입력하라는 다음 필드가 사용 가능하게 됩니다.

사용자 ID. 서버에 로그온하는 데 사용할 사용자 이름을 입력합니다.

비밀번호. 지정된 사용자 이름과 연관된 비밀번호를 입력합니다.

도메인. 서버에 로그온하는 데 사용된 도메인을 지정합니다. 도메인 이름은 서버 컴퓨터가 클라이언트 컴퓨터와 다른 Windows 도메인에 있을 때에만 필요합니다.

3. 연결을 완료하려면 **확인**을 클릭하십시오.

서버에서 연결 해제하기

1. 도구 메뉴에서 **서버 로그인**을 클릭하십시오. 서버 로그인 대화 상자가 열립니다. 또는 IBM SPSS Modeler 창의 연결 상태 영역을 두 번 클릭하십시오.
2. 대화 상자에서 로컬 서버를 선택하고 **확인**을 클릭하십시오.

IBM SPSS Modeler Server 연결 추가 및 편집

서버 로그인 대화 상자에서 수동으로 서버 연결을 편집하거나 추가할 수 있습니다. 추가를 클릭하여 서버 연결 세부사항을 입력할 수 있는 서버 추가/편집 대화 상자에 액세스할 수 있습니다. 기존 연결을 선택하고 서버 로그인 대화 상자에서 편집을 클릭하면 사용자가 변경을 수행할 수 있도록 해당 연결에 대한 세부사항과 함께 서버 추가/편집 대화 상자가 열립니다.

참고: IBM SPSS Collaboration and Deployment Services에서 추가된 서버 연결은 편집할 수 없습니다. 이름, 포트 및 기타 세부사항이 IBM SPSS Collaboration and Deployment Services에 정의되어 있기 때문입니다. 우수 사례에서는 IBM SPSS Collaboration and Deployment Services 및 SPSS Modeler 클라이언트 둘 모두와 통신하는 데에는 동일 포트를 사용해야 한다고 지시합니다. 이들은 options.cfg 파일에 max_server_port 및 min_server_port로 설정할 수 있습니다.

서버 연결 추가

1. 도구 메뉴에서 **서버 로그인**을 클릭하십시오. 서버 로그인 대화 상자가 열립니다.
 2. 이 대화 상자에서 **추가**를 클릭하십시오. 서버 로그인 서버 추가/편집 대화 상자가 열립니다.
 3. 서버 연결 세부사항을 입력하고 **확인**을 클릭하여 연결을 저장하고 서버 로그인 대화 상자로 돌아가십시오.
- **서버.** 사용 가능한 서버를 지정하거나 목록에서 하나를 선택하십시오. 서버 컴퓨터는 영숫자 이름(예: *myserver*)이나 서버 컴퓨터에 지정된 IP 주소(예: 202.123.456.78)로 식별할 수 있습니다.
 - **포트.** 서버가 청취 중인 포트 번호를 제공하십시오. 기본값이 작동하지 않으면 시스템 관리자에게 올바른 포트 번호를 문의하십시오.
 - **설명.** 이 서버 연결에 대한 선택적 설명을 입력하십시오.
 - **보안 연결 확인(SSL 사용).** SSL(**Secure Sockets Layer**) 연결을 사용해야 하는지 여부를 지정합니다. SSL은 네트워크를 통해 전송된 데이터를 보안하기 위해 주로 사용하는 프로토콜입니다. 이 기능을 사용하려면 IBM SPSS Modeler Server를 호스팅하는 서버에서 SSL이 사용 가능으로 설정되어야 합니다. 필요한 경우, 세부사항은 로컬 관리자에게 문의하십시오.

서버 연결 편집하기

1. 도구 메뉴에서 **서버 로그인**을 클릭하십시오. 서버 로그인 대화 상자가 열립니다.
2. 이 대화 상자에서 편집할 연결을 선택한 다음 **편집**을 클릭하십시오. 서버 로그인 서버 추가/편집 대화 상자가 열립니다.
3. 서버 연결 세부사항을 변경하고 **확인**을 클릭하여 변경사항을 저장하고 서버 로그인 대화 상자로 돌아가십시오.

IBM SPSS Collaboration and Deployment Services에서 서버 검색

서버 연결에 수동으로 들어가는 대신에 IBM SPSS Collaboration and Deployment Services에서 사용 가능한 프로세스 조정자를 통해 네트워크에서 사용 가능한 서버나 서버 군집을 선택할 수 있습니다. 서버 군집은 프로세스 조정자가 처리 요청에 대한 응답에 가장 적합한 서버를 판별하는 데 사용하는 서버 그룹입니다.

서버 로그인 대화 상자에서 서버를 수동으로 추가할 수 있지만 사용 가능한 서버를 검색하면 올바른 서버 이름과 포트 번호를 몰라도 서버에 연결할 수 있습니다. 이 정보는 자동으로 제공됩니다. 그러나 사용자 이름, 도메인 및 비밀번호 등과 같은 올바른 로그온 정보는 여전히 필요합니다.

참고: 프로세스 조정자 기능에 대한 액세스가 없는 경우에는 연결하려는 서버 이름을 수동으로 입력하거나 이전에 정의된 이름을 선택할 수도 있습니다. 자세한 정보는 8 페이지의 [『IBM SPSS Modeler Server 연결 추가 및 편집』](#)의 내용을 참조하십시오.

서버와 군집 검색하기

1. 도구 메뉴에서 **서버 로그인**을 클릭하십시오. 서버 로그인 대화 상자가 열립니다.
2. 이 대화 상자에서 **검색**을 클릭하여 서버 검색 대화 상자를 엽니다. 프로세스 조정자를 찾으려고 시도할 때 IBM SPSS Collaboration and Deployment Services에 로그온되어 있지 않은 경우에는 로그온하라는 메시지가 표시됩니다.
3. 목록에서 서버나 서버 군집을 선택하십시오.
4. **확인**을 클릭하여 대화 상자를 닫고 서버 로그인 대화 상자에서 테이블에 이 연결을 추가하십시오.

Analytic Server에 연결

여러 Analytic Server를 사용할 수 있으면 분석 서버 연결 대화 상자를 사용하여 IBM SPSS Modeler에서 사용할 서버를 두 개 이상 정의할 수 있습니다. 관리자가 이미 <Modeler_install_path>/config/options.cfg 파일에서 기본 Analytic Server를 설정했을 수 있습니다. 그러나 서버를 정의한 후 사용 가능한 다른 서버도 사용할 수 있습니다. 예를 들어, Analytic Server 소스 및 내보내기 노드를 사용할 때, 각 분기가 실행될 때 고유 Analytic Server를 사용하고 데이터가 IBM SPSS Modeler Server에 가져오지 않도록 스트림의 여러 다른 분기에서 다른 Analytic Server 연결을 사용할 수 있습니다. 분기에 두 개 이상의 Analytic Server 연결이 포함되어 있으면 Analytic Server에서 IBM SPSS Modeler Server로 데이터를 가져옵니다.

새 Analytic Server 연결을 작성하려면 **Tools** > **분석 서버 연결**로 이동하고 다음 대화 상자 섹션에 필수 정보를 제공하십시오.

연결

URL. `https://hostname:port/contextroot` 형식으로 Analytic Server의 URL을 입력하십시오. 여기서 `hostname`은 Analytic Server의 IP 주소 또는 호스트 이름이고, `port`는 포트 번호이며, `contextroot`는 Analytic Server의 컨텍스트 루트입니다.

테넌트 IBM SPSS Modeler Server가 멤버인 테넌트의 이름을 입력하십시오. 테넌트를 모르면 관리자에게 문의하십시오.

인증

모드 다음 인증 모드에서 선택하십시오.

- **사용자 이름 및 비밀번호**에 사용자 이름과 비밀번호를 입력해야 합니다.
- **저장된 신임 정보**에서 IBM SPSS Collaboration and Deployment Services Repository의 신임 정보를 선택해야 합니다.
- **Kerberos**의 경우, 사용자가 서비스 프린시펄 이름 및 구성 파일 경로를 입력해야 합니다. 이 정보를 모르면 관리자에게 문의하십시오.

사용자 이름 Analytic Server 사용자 이름을 입력하십시오.

영역. Analytic Server 연결에 사용할 영역을 선택하십시오.

비밀번호 Analytic Server 비밀번호를 입력하십시오.

연결 연결을 클릭하여 새 연결을 테스트하십시오.

연결

위의 정보를 지정한 다음 **연결**을 클릭하면 이 연결 테이블에 연결이 추가됩니다. 연결을 제거해야 하는 경우 연결을 선택하고 **제거**를 누르십시오.

관리자가 `options.cfg` 파일에 기본 Analytic Server 연결을 정의한 경우 **기본 연결 추가**를 클릭하여 사용 가능한 연결에도 추가할 수 있습니다. 사용자 이름과 비밀번호를 입력하도록 프롬프트됩니다.

임시 디렉토리 변경

IBM SPSS Modeler Server가 수행하는 일부 작업에는 임시 파일을 작성해야 할 수도 있습니다. 기본적으로 IBM SPSS Modeler는 시스템 임시 디렉토리를 사용하여 임시 파일을 작성합니다. 다음 단계를 사용하여 임시 디렉토리의 위치를 변경할 수 있습니다.

1. `spss`라는 새 디렉토리와 `servertemp`라는 서브디렉토리를 작성하십시오.
2. IBM SPSS Modeler 설치 디렉토리의 `/config` 디렉토리에 있는 `options.cfg`를 편집하십시오. 이 파일에서 `temp_directory` 모수를 편집하여 `temp_directory`, "C:/spss/servertemp"가 되도록 하십시오.
3. 이를 수행한 후에는 IBM SPSS Modeler Server 서비스를 다시 시작해야 합니다. Windows 제어판에서 **서비스** 탭을 클릭하여 이를 수행할 수 있습니다. 서비스를 중지한 다음 다시 시작하여 변경사항을 활성화하십시오. 시스템을 다시 시작하면 서비스도 다시 시작됩니다.

모든 임시 파일이 이 새 디렉토리에 작성됩니다.

참고: 슬래시를 사용해야 합니다.

다중 IBM SPSS Modeler 세션 시작

한 번에 둘 이상의 IBM SPSS Modeler 세션을 실행해야 하는 경우에는 IBM SPSS Modeler 및 Windows 설정을 약간 변경해야 합니다. 예를 들어, 두 개의 별도의 서버 라이센스가 있고 동일한 클라이언트 시스템에서 두 개의 서로 다른 서버에 대해 두 개의 스트림을 실행하려는 경우 이를 수행할 수도 있습니다.

다중 IBM SPSS Modeler 세션 사용

1. 다음을 클릭하십시오.

시작 > [모든] 프로그램 > **IBM SPSS Modeler**

2. IBM SPSS Modeler 단축키(아이콘이 있음)에서 마우스 오른쪽 단추를 클릭하고 **특성을 선택하십시오.**
3. **목표 텍스트** 상자에서 **-noshare**를 문자열 끝에 추가하십시오.
4. Windows 탐색기에서 다음을 선택하십시오.

도구 > 폴더 옵션...

5. 파일 유형 탭에서 IBM SPSS Modeler 스트림 옵션을 선택하고 **고급**을 클릭하십시오.
6. 파일 유형 편집 대화 상자에서 IBM SPSS Modeler로 열기를 선택하고 **편집**을 클릭하십시오.
7. 조치를 수행하는 데 사용하는 애플리케이션 텍스트 상자에서 **-stream** 인수 앞에 **-noshare**를 추가하십시오.

IBM SPSS Modeler 인터페이스 살펴보기

데이터 마이닝 프로세스의 각 지점에서 사용하기 용이한 IBM SPSS Modeler 인터페이스는 사용자의 특정 비즈니스 전문 지식을 활용합니다. 예측, 분류, 세분화, 연관 발견과 같은 모델링 알고리즘은 강력하고 정확한 모델을 보장합니다. 모델 결과를 쉽게 배포하고 데이터베이스, IBM SPSS Statistics, 다양한 기타 애플리케이션으로 읽어들일 수 있습니다.

IBM SPSS Modeler에 대한 작업은 데이터에 대해 작업하는 3단계 프로세스입니다.

- 먼저, IBM SPSS Modeler로 데이터를 읽어들입니다.
- 그 다음에는 일련의 조작을 통해 데이터를 실행합니다.
- 마지막으로, 데이터를 대상으로 보냅니다.

각 조작을 통해 소스에서 마침내 대상(모델 또는 데이터 출력 유형)으로 레코드별로 데이터가 플로우되기 때문에 이러한 작업 시퀀스를 **데이터 스트림**이라고 합니다.

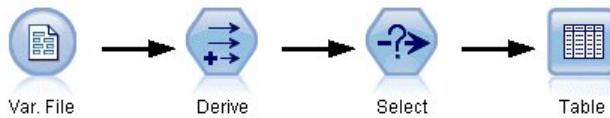


그림 2. 단순 스트림

IBM SPSS Modeler 스트림 캔버스

스트림 캔버스는 IBM SPSS Modeler 창의 가장 큰 영역이며 데이터 스트림을 작성하고 조작하는 위치입니다.

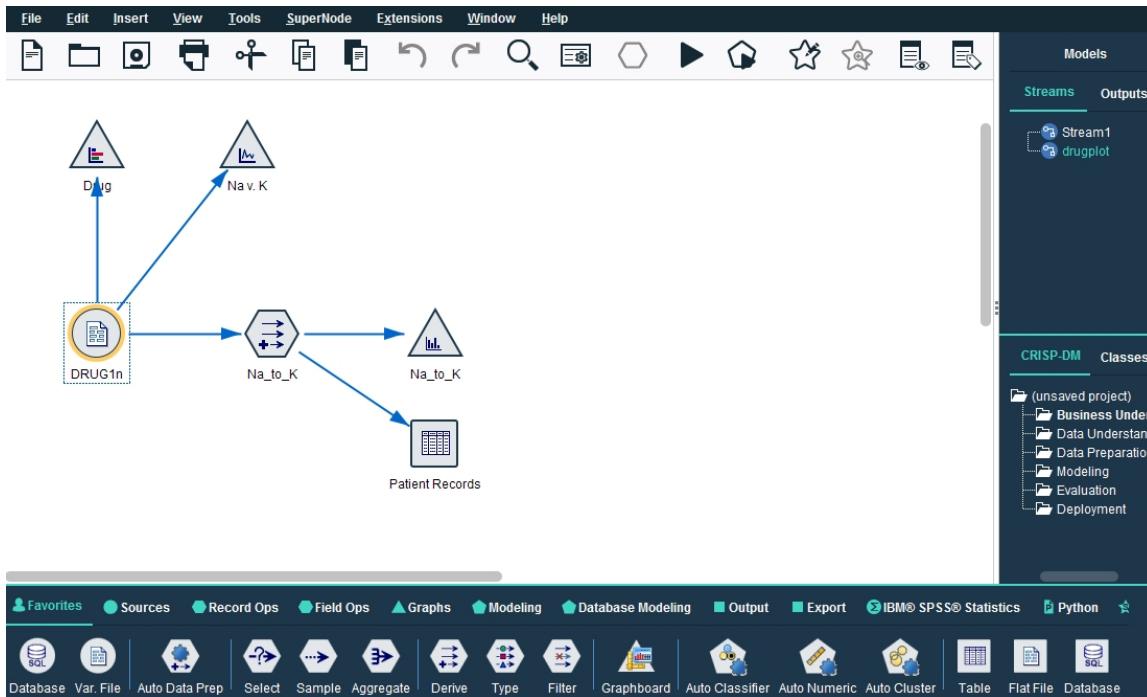


그림 3. IBM SPSS Modeler 작업공간(기본 보기)

스트림은 인터페이스의 기본 캔버스에서 비즈니스 관련 데이터 연산 다이어그램을 그려서 작성됩니다. 각 연산은 아이콘 또는 노드로 표시되며, 노드는 각 연산을 통한 데이터 플로우를 표시하는 스트림에서 함께 연결됩니다.

동일한 스트림 캔버스에서 또는 새 스트림 캔버스를 열어 IBM SPSS Modeler에서 동시에 다중 스트림에 대한 작업을 할 수 있습니다. 세션 동안 스트림은 IBM SPSS Modeler 창의 오른쪽 상단에 있는 스트림 관리자에 저장됩니다.

참고: 기본 제공 트랙패드의 클릭 및 촉감 피드백 강제 실행 설정을 사용하는 MacBook을 사용하는 경우 노드 팔레트에서 스트림 캔버스로 노드를 끌어서 놓으면 중복 노드가 캔버스에 추가됩니다. 이 문제를 방지하려면 클릭 및 촉감 피드백 강제 실행 트랙패드 시스템 환경 설정을 사용하지 않는 것이 좋습니다.

노드 팔레트

SPSS Modeler에서 대다수의 데이터와 모델링 도구는 스트림 캔버스 아래의 창 아래에 걸쳐서 노드 팔레트에서 사용할 수 있습니다.

예를 들어, **레코드 작업 팔레트** 탭에는 선택, 병합 및 붙여쓰기 등과 같이 데이터 레코드에서 작업을 수행하는데 사용할 수 있는 노드가 포함됩니다.

노드를 캔버스에 추가하려면 노드 팔레트에서 아이콘을 두 번 클릭하거나 이를 캔버스로 끄십시오. 그런 다음 이를 연결하여 데이터 플로우를 나타내는 스트림을 작성할 수 있습니다.



그림 4. 노드 팔레트의 레코드 작업 탭

각 팔레트 탭에는 다음과 같이 스트림 작업의 서로 다른 단계에 사용하는 관련 노드의 컬렉션이 포함됩니다.

- **소스 노드**는 데이터를 SPSS Modeler로 가져옵니다.
- **레코드 작업 노드**는 선택, 병합 및 붙여쓰기 등의 데이터 레코드 작업을 수행합니다.
- **필드 작업 노드**는 필터링, 새 필드 파생 및 주어진 필드의 측정 수준 판별 등의 데이터 필드 작업을 수행합니다.
- **그래프 노드**는 모델링 전후 데이터를 그래프로 표시합니다. 그래프에는 도표, 히스토그램, 웹 노드 및 평가 차트를 포함합니다.

- **모델링 노드**는 신경망, 의사결정 트리, 군집 알고리즘 및 데이터 순차규칙 등과 같이 SPSS Modeler에서 사용 가능한 모델링 알고리즘을 사용합니다.
- **데이터베이스 모델링 노드**는 Microsoft SQL Server, IBM Db2 및 Oracle과 Netezza 데이터베이스에서 사용 가능한 모델링 알고리즘을 사용합니다.
- **출력 노드**는 SPSS Modeler에서 볼 수 있는 데이터, 차트 및 모델 결과의 다양한 출력을 생성합니다.
- **내보내기 노드**는 IBM SPSS Data Collection 또는 Excel과 같은 외부 애플리케이션에서 볼 수 있는 다양한 출력을 생성합니다.
- **IBM SPSS Statistics 노드**는 IBM SPSS Statistics에서 데이터를 가져오거나 내보낼 뿐만 아니라 IBM SPSS Statistics 프로시저를 실행합니다.
- **Python 노드**는 Python 알고리즘을 실행하는 데 사용할 수 있습니다.
- **Spark 노드**를 사용하여 Spark 알고리즘을 실행할 수 있습니다.

SPSS Modeler와 익숙해짐에 따라 원하는 대로 팔레트 내용을 사용자 정의할 수 있습니다.

노드 팔레트의 왼쪽에서는 감독, 연관 또는 세분화를 선택하여 표시되는 노드를 필터링할 수 있습니다.

노드 팔레트 아래에 있는 보고서 분할창은 데이터를 데이터 스트림으로 읽어 올 때 등과 같이 다양한 작업의 진행률에 대한 피드백을 제공합니다. 마찬가지로 노드 팔레트 아래에 있는 상태 분할창은 애플리케이션이 현재 하는 작업뿐만 아니라 사용자 피드백이 필요한 시기 표시에 대한 정보를 제공합니다.

참고: 기본 제공 트래파드의 **클릭 및 촉감 피드백 강제 실행** 설정을 사용하는 MacBook을 사용하는 경우 노드 팔레트에서 스트림 캔버스로 노드를 끌어서 놓으면 중복 노드가 캔버스에 추가됩니다. 이 문제를 방지하려면 **클릭 및 촉감 피드백 강제 실행** 트래파드 시스템 환경 설정을 사용하지 않는 것이 좋습니다.

IBM SPSS Modeler 관리자

창의 오른쪽 상단은 관리자 분할창입니다. 여기에는 스트림, 출력, 모델을 관리하는 데 사용하는 세 개의 탭이 있습니다.

스트림 탭을 사용하여 세션에 작성된 스트림을 열고 이름을 변경하며, 저장 및 삭제할 수 있습니다.



그림 5. 스트림 탭

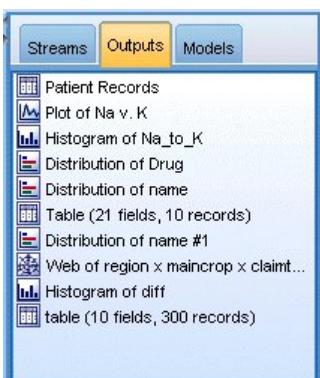


그림 6. 출력 탭

출력 탭은 IBM SPSS Modeler에서 스트림 작업을 통해 생성된 그래프, 테이블과 같은 다양한 파일을 포함합니다. 이 탭에 나열된 테이블, 그래프, 보고서를 표시하고 저장하며, 이름을 변경하고 닫을 수 있습니다.

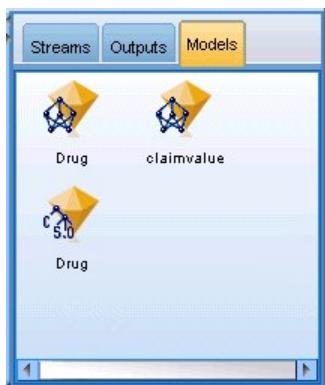


그림 7. 모델 너깃이 포함된 모델 탭

모델 탭은 관리자 탭 중 가장 강력합니다. 이 탭은 현재 세션에 대해 IBM SPSS Modeler에 생성된 모델을 포함하는 모든 모델 너깃을 포함합니다. 이러한 모델을 모델 탭에서 직접 찾아보거나 캔버스의 스트림에 추가할 수 있습니다.

IBM SPSS Modeler 프로젝트

창의 오른쪽 하단은 데이터 마이닝 프로젝트(데이터 마이닝 작업과 관련된 파일 그룹)를 작성하고 관리하는 데 사용하는 프로젝트 분할창입니다. IBM SPSS Modeler에서 사용자가 작성하는 프로젝트를 볼 수 있는 두 가지 방법이 있는데, 클래스 보기와 CRISP-DM 보기입니다.

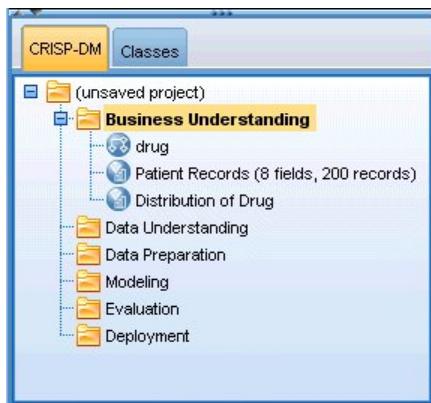


그림 8. CRISP-DM 보기

CRISP-DM 탭은 업계에서 입증된 비독점 방법론인 CRISP-DM(Cross-Industry Standard Process for Data Mining)에 따라 프로젝트를 구성하는 방법을 제공합니다. 숙련된 데이터 마이너와 초보 데이터 마이너 모두를 위해 CRISP-DM 도구를 사용하면 데이터 마이닝을 위한 노력을 쉽게 체계화하고 전달할 수 있습니다.

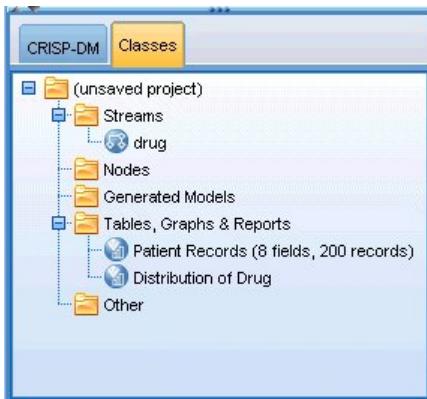


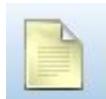
그림 9. 클래스 보기

클래스 탭은 IBM SPSS Modeler에서 작업을 범주적으로(작성하는 오브젝트 유형별로) 구성하는 방법을 제공합니다. 이 보기는 데이터, 스트림, 모델 인벤토리를 만들 때 유용합니다.

IBM SPSS Modeler 도구 모음

IBM SPSS Modeler 창의 맨 위에서 다양한 유용한 기능을 제공하는 아이콘 도구 모음을 볼 수 있습니다. 다음은 도구 모음 단추 및 해당 함수입니다.





스트림에 마크업 없음



주석 삽입



스트림 마크업 숨김(있는 경우)



숨겨진 스트림 마크업 표시



스트림을 IBM SPSS Modeler Advantage에서 열기

스트림 마크업은 스트림 주석, 모델 링크 및 스코어링 분기 표시로 구성됩니다.

모델 링크는 *IBM SPSS 모델링 노드* 안내서에 설명되어 있습니다.

도구 모음 사용자 정의

다음과 같이 도구 모음의 다양한 측면을 변경할 수 있습니다.

- 표시되는지 여부
- 아이콘에 사용 가능한 도구 텁이 있는지 여부
- 큰 아이콘 또는 작은 아이콘을 사용하는지 여부

도구 모음 표시를 켜고 끄기

1. 기본 메뉴에서 다음을 클릭하십시오.

보기 > 도구 모음 > 표시

도구 텁이나 아이콘 크기 설정 변경

1. 기본 메뉴에서 다음을 클릭하십시오.

보기 > 도구 모음 > 사용자 정의

필요에 따라 도구 텁 표시 또는 큰 단추를 클릭하십시오.

IBM SPSS Modeler 창 사용자 정의

SPSS Modeler 인터페이스의 다양한 부분 사이에 디바이더를 사용할 때 기본 설정을 충족하도록 도구의 크기를 조정하거나 닫을 수 있습니다. 예를 들어, 대용량 스트림에 대해 작업 중인 경우 각 디바이더에 있는 작은 화살표를 사용하여 노드 팔레트, 관리자 분할창 및 프로젝트 분할창을 닫을 수 있습니다. 이는 스트림 캔버스를 최소화하여 대용량 또는 다량의 스트림을 위한 충분한 작업 공간을 제공합니다.

또는 보기 메뉴에서 노드 팔레트, 관리자 또는 프로젝트를 클릭하여 이러한 항목을 켜거나 끌 수도 있습니다.

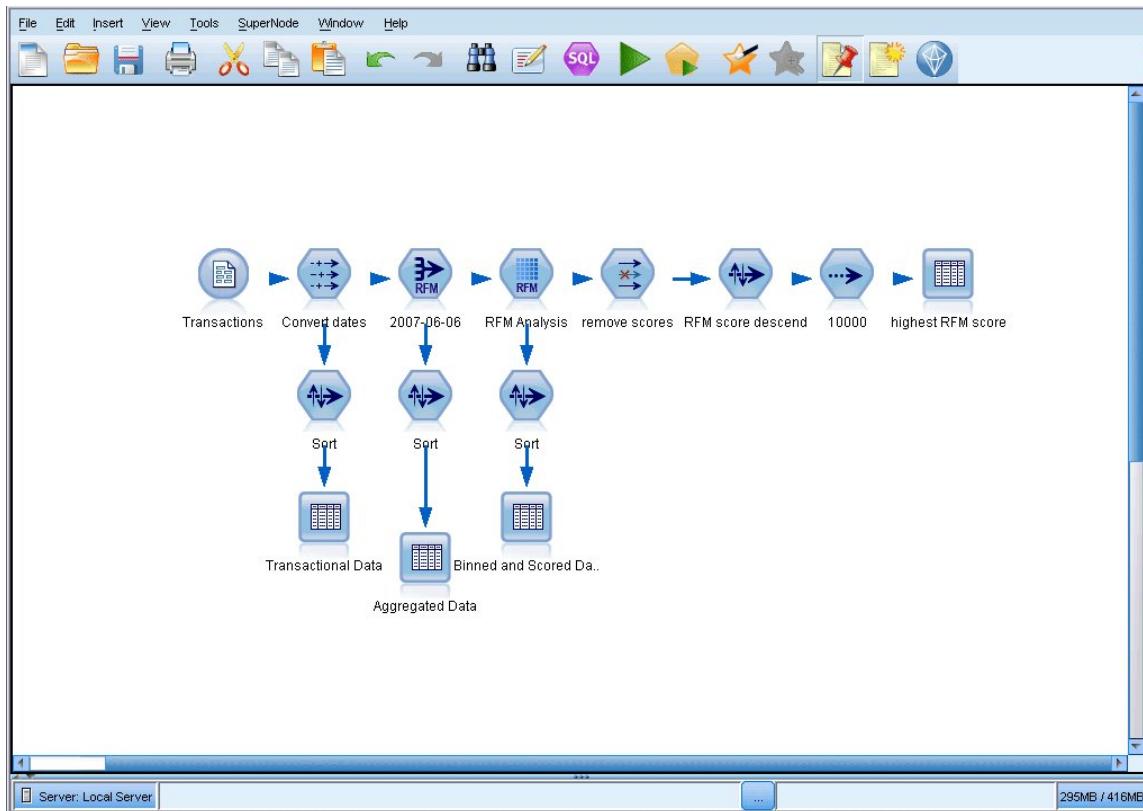


그림 10. 최대 스트림 캔버스

노드 팔레트, 관리자 및 프로젝트 분할창을 닫는 대신에 SPSS Modeler 창의 측면과 아래쪽에 있는 스크롤바를 가로 세로로 이동하여 스트림 캔버스를 스크롤 가능한 페이지로 사용할 수 있습니다.

또한 스트림 주석, 모델 링크 및 스코어링 분기 표시로 구성된 화면 마크업 표시를 제어할 수도 있습니다. 이 표시를 켜거나 끄려면 다음을 클릭하십시오.

[보기 > 스트림 마크업](#)

스트림 아이콘 크기 변경

다음 방법으로 스트림 아이콘 크기를 변경할 수 있습니다.

- 스트림 특성 설정을 통해
- 스트림의 팝업 메뉴를 통해
- 키보드 사용

전체 스트림 보기의 표준 아이콘 크기의 8% - 200% 사이의 많은 크기 중 하나로 스케일링할 수 있습니다.

전체 스트림 스케일링 방법(스트림 특성 방법)

1. 메뉴에서 다음을 선택하십시오.

[도구 > 스트림 특성 > 옵션 > 레이아웃](#)

2. 아이콘 크기 메뉴에서 원하는 크기를 선택하십시오.
3. 적용을 클릭하여 결과를 확인하십시오.
4. 확인을 클릭하여 변경사항을 저장하십시오.

전체 스트림 스케일링 방법(메뉴 방법)

1. 캔버스에서 스트림 배경을 마우스 오른쪽 단추로 클릭하십시오.

2. 아이콘 크기를 선택하고 원하는 크기를 선택하십시오.

전체 스트림 스케일링 방법(키보드 방법)

1. 기본 키보드에서 Ctrl +[-]를 눌러 크기를 축소하십시오.
2. 기본 키보드에서 Ctrl +Shift + [-]를 눌러 크기를 확대하십시오.

이 방식의 확대는 사용하는 운영 체제 및 키보드에 따라 작동하지 않을 수 있습니다.

이 기능은 복합 스트림에 대한 전체 보기로 확보할 경우 특히 유용합니다. 또한 이를 사용하여 스트림을 인쇄하는데 필요한 페이지 수를 최소화할 수 있습니다.

IBM SPSS Modeler에서 마우스 사용

IBM SPSS Modeler에서 마우스의 가장 흔한 용도는 다음과 같습니다.

- **단일 클릭.** 마우스 오른쪽이나 왼쪽 단추를 사용하여 메뉴에서 옵션을 선택하고, 팝업 메뉴를 열고, 다양한 기타 표준 제어와 옵션에 액세스할 수 있습니다. 노드를 이동하거나 끌려면 단추를 클릭한 후 계속 누르십시오.
- **두 번 클릭.** 마우스 왼쪽 단추를 두 번 클릭하여 노드를 스트림 캔버스에 놓고 기존 노드를 편집하십시오.
- **가운데 클릭.** 마우스 가운데 단추를 클릭하여 커서를 끌어서 스트림 캔버스에 있는 노드를 연결하십시오. 노드를 연결 해제하려면 마우스 가운데 단추를 두 번 클릭하십시오. 버튼이 3개인 마우스가 아닌 경우에는 마우스를 클릭하고 끄는 중에 Alt 키를 눌러서 이 기능을 시뮬레이션할 수 있습니다.

단축키 사용

IBM SPSS Modeler의 많은 비주얼 프로그래밍 작업에는 연관된 단축키가 있습니다. 예를 들어, 노드를 클릭하고 키보드에서 삭제 키를 눌러 노드를 삭제할 수 있습니다. 마찬가지로, Ctrl 키를 누른 상태에서 S 키를 눌러서 스트림을 빠르게 저장할 수 있습니다. 이와 같은 제어 명령은 Ctrl 및 다른 키 조합(예: Ctrl+S)으로 표시됩니다.

표준 Windows 작업에 사용하는 많은 단축키가 있습니다(예: 잘라내기는 Ctrl+X). 이러한 단축키는 다음 애플리케이션 특정 단축키와 함께 IBM SPSS Modeler에서 지원됩니다.

참고: 어떤 경우에는 IBM SPSS Modeler에서 사용한 이전 단축키가 표준 Windows 단축키와 충돌할 수 있습니다. 이러한 이전 단축키는 Alt 키의 추가와 함께 지원됩니다. 예를 들어, Ctrl+Alt+C는 캐시를 켜기/끄기를 토글하는 데 사용할 수 있습니다.

표 1. 지원되는 단축키	
단축키	함수
Ctrl+A	모두 선택
Ctrl+X	잘라내기
Ctrl+N	새 스트림
Ctrl+O	스트림 열기
Ctrl+P	인쇄
Ctrl+C	복사
Ctrl+V	붙여넣기
Ctrl+Z	실행 취소
Ctrl+Q	선택한 노드의 모든 노드 다운스트림 선택
Ctrl+W	모든 다운스트림 노드 선택 취소(Ctrl+Q로 토글)
Ctrl+E	선택한 노드에서 실행
Ctrl+S	현재 스트림 저장
Alt+화살표 키	선택한 노드를 캔버스 스트림에서 사용된 화살표의 방향으로 이동

표 1. 지원되는 단축키 (계속)

단축키	함수
Shift+F10	선택한 노드의 팝업 메뉴 열기

표 2. 이전 단축키의 지원되는 단축키

단축키	함수
Ctrl+Alt+D	노드 중복
Ctrl+Alt+L	노드 로드
Ctrl+Alt+R	노드 이름 변경
Ctrl+Alt+U	사용자 입력 노드 작성
Ctrl+Alt+C	캐시 켜기/끄기 토글
Ctrl+Alt+F	캐시 비우기
Ctrl+Alt+X	수퍼노드 확장
Ctrl+Alt+Z	확대/축소
Delete	노드나 연결 삭제

인쇄

다음 오브젝트를 IBM SPSS Modeler에서 인쇄할 수 있습니다.

- 스트림 다이어그램
- 그래프
- 테이블
- 보고서(보고서 노드 및 프로젝트 보고서에서)
- 스크립트(스트림 특성, 독립형 스크립트 또는 수퍼노드 스크립트 대화 상자에서)
- 모델(모델 브라우저, 현재 초점이 있는 대화 상자 탭, 트리 뷰어)
- 주석(Annotation)(출력을 위해 주석 탭 사용)

오브젝트 인쇄

- 미리보지 않고 인쇄하려면 도구 모음에서 인쇄 단추를 클릭하십시오.
- 인쇄하기 전에 페이지를 설정하려면 파일 메뉴에서 페이지 설정을 선택하십시오.
- 인쇄 전에 미리보려면 파일 메뉴에서 인쇄 미리보기를 선택하십시오.
- 프린터 선택을 위한 옵션이 있는 표준 인쇄 대화 상자를 보려면 파일 메뉴에서 인쇄를 선택하십시오.

IBM SPSS Modeler 자동화

고급 데이터 마이닝은 복잡하고 때로는 시간이 긴 프로세스일 수 있으므로 IBM SPSS Modeler에는 몇몇 코딩 유형 및 자동화 지원이 포함됩니다.

- **CLEM**(Control Language for Expression Manipulation)은 IBM SPSS Modeler 스트림과 함께 플로우하는 데 이터를 분석하고 조작하기 위한 언어입니다. 데이터 마이너는 비용과 수입 데이터에서 수익을 파생하는 것과 같은 단순한 작업 또는 웹 로그 데이터를 유용한 정보가 있는 필드와 레코드 세트로 변환하는 것과 같은 복잡한 작업을 수행하기 위해 스트림 작업에서 CLEM을 집중적으로 사용합니다.
- **스크립팅**은 사용자 인터페이스에서 프로세스를 자동화하기 위한 강력한 도구입니다. 스크립트는 사용자가 마우스나 키보드를 사용하여 수행하는 것과 동일한 유형의 조치를 수행합니다. 또는 출력을 지정하고 생성된 모델을 조작할 수도 있습니다.

제 4 장 데이터 마이닝 이해

데이터 마이닝 개요

다양한 기술을 통해 **데이터 마이닝**은 데이터 본문의 정보 너깃을 식별합니다. 데이터 마이닝은 의사결정 지원, 예측, 예측값 및 추정 등과 같은 영역에서 사용할 수 있는 방식으로 정보를 추출합니다. 데이터는 종종 볼륨이 크지만 값은 낮고 원래 상태로는 유용성이 적기도 합니다. 이는 값이 있는 데이터에 숨겨진 정보입니다.

데이터 마이닝에서 성공은 사용자 또는 전문가의 데이터 지식을 고급, 활성 분석 기술과 결합하는 것입니다. 이 기술을 통해 컴퓨터는 데이터의 기본 관계와 기능을 식별하게 됩니다. 데이터 마이닝의 프로세스는 나중에 예측, 패턴 발견 등에 사용하는 히스토리 데이터로부터 모델을 생성합니다. 이러한 모델을 작성하는 기술을 **시스템 학습** 또는 **모델링**이라고 부릅니다.

모델링 기술

IBM SPSS Modeler에는 수많은 시스템 학습과 모델링 기술이 포함됩니다. 이들은 해결하려는 문제 유형에 따라 개략적으로 그룹화할 수 있습니다.

- 예측 모델링 방법에는 의사결정 트리, 신경망 및 통계 모델이 포함됩니다.
- 군집 모델은 유사한 레코드 그룹을 식별하고 레코드에 이들이 속하는 그룹에 따라 레이블을 붙이는 데 초점을 둡니다. 군집 방법에는 코호넨, k-평균 및 TwoStep이 포함됩니다.
- 연관 규칙은 특정 결론(예: 특정 제품의 구매)을 조건 세트(몇몇 기타 제품의 구매)와 연관시킵니다.
- 선별 모델은 모델링에서 가장 관심이 있는 레코드와 필드를 찾기 위해 데이터를 선별하고 알려진 패턴과 매치하지 않을 수 있는 이상값을 식별하는 데 사용할 수 있습니다. 사용 가능한 방법에는 필드선택과 이상 항목 발견이 포함됩니다.

데이터 조작과 발견

IBM SPSS Modeler에는 또한 사용자의 전문지식을 데이터에 적용할 수 있는 여러 기능이 포함됩니다.

- **데이터 조작.** 기존 데이터 항목에서 파생된 새 데이터 항목을 작성하고 데이터를 의미있는 서브세트로 분류합니다. 다양한 소스의 데이터를 병합하고 필터링할 수 있습니다.
- **찾아보기 및 시각화.** 그래프와 통계량을 포함하여 초기 감사를 수행하기 위해 데이터 검토 노드를 사용하여 데이터의 측면을 표시합니다. 고급 시각화에는 대화식 그래프가 포함되고 이는 프로젝트 보고서에 포함하기 위해 내보낼 수 있습니다.
- **통계량.** 데이터에서 변수 사이의 추측되는 관계를 확인합니다. IBM SPSS Statistics의 통계량은 IBM SPSS Modeler 내에서도 사용할 수 있습니다.
- **가설 검정.** 데이터가 작동하는 방법의 모델을 작성하고 이러한 모델을 검증합니다.

일반적으로 이러한 기능을 사용하여 데이터에서 속성의 유망 세트를 식별합니다. 그런 다음 이러한 속성은 모델링 기법에 공급될 수 있으며 이는 기본 규칙 및 관계를 식별하려고 시도합니다.

일반 애플리케이션

데이터 마이닝 기술의 일반 애플리케이션에는 다음이 포함됩니다.

직접적인 메일. 어떤 인구통계 그룹이 응답 비율이 가장 높은지 결정합니다. 이 정보를 사용하여 이후 메일링의 응답을 극대화할 수 있습니다.

제작자 정보 스코어링. 개인의 제작자 정보 히스토리를 사용하여 제작자 정보 의사결정을 내릴 수 있습니다.

인적 자원. 지난 채용 관행을 이해하고 채용 프로세스를 간소화하기 위해 의사결정 규칙을 작성합니다.

의학 연구. 의학적 증거를 기반으로 적합한 프로시저를 제안하는 의사결정 규칙을 작성합니다.

시장 분석. 지리학, 가격 및 고객 특성 등과 같은 어떤 변수가 판매와 연관되는지를 결정합니다.

품질관리. 제품 제조의 데이터를 분석하고 제품 결함을 판별하는 변수를 식별합니다.

정책 연구. 설문조사 데이터를 사용하여 가장 중요한 변수를 선택하기 위해 의사결정 규칙을 적용하여 정책을 공식화할 수 있습니다.

건강 관리. 사용자 설문조사와 임상 데이터는 건강에 기여하는 변수를 발견하기 위해 결합할 수 있습니다.

용어

속성, 필드 및 변수 용어는 고려 중인 모든 케이스에 공통된 단일 데이터 항목을 가리킵니다. 특정 케이스를 가리키는 속성 값의 콜렉션은 **레코드, 예제** 또는 **케이스**라고 합니다.

데이터 평가

데이터 마이닝은 사용하려는 데이터가 특정 기준을 충족하지 않는 한 이익이 많지 않을 수 있습니다. 다음 섹션은 사용자가 고려해야 하는 데이터 및 해당 애플리케이션의 특정 측면을 보여줍니다.

데이터가 사용 가능하지 확인

이는 명백한 것처럼 보이지만 데이터가 사용 가능하더라도 쉽게 사용할 수 있는 양식이 아닐 수도 있습니다. IBM SPSS Modeler는 데이터베이스(ODBC를 통해) 또는 목록에서 데이터를 가져올 수 있습니다. 그러나 데이터는 직접 액세스할 수 없는 시스템에서 몇몇 다른 양식으로 되어 있을 수 있습니다. 사용하려면 먼저 적합한 양식으로 다운로드하고 덤프되어야 합니다. 이는 서로 다른 데이터베이스와 소스 사이에 흘어져 있어서 한데 모아야 할 필요가 있을 수 있습니다. 온라인에 없을 수도 있습니다. 서류상으로만 존재하는 경우에는 데이터 마이닝을 시작하려면 데이터 입력이 필요합니다.

데이터가 관련 속성을 커버하는지 여부 확인

데이터 마이닝의 오브젝트는 관련 속성을 식별하는 것이므로 이 확인을 포함하는 것은 처음에는 이상하게 보일 수 있습니다. 그러나 어떤 데이터가 사용 가능한지 확인하고 기록되지 않은 관련 요소를 식별하려고 시도하는 것은 유용할 수 있습니다. 예를 들어 아이스크림 판매를 예측하려고 시도할 때 소매 점포 또는 판매 내역에 대한 많은 정보가 있을 수 있지만 중요한 역할을 하는 날씨와 온도 정보는 가지고 있지 않을 수 있습니다. 결측 속성은 반드시 데이터 마이닝이 유용한 결과를 생성하지 않는다는 것을 의미하지는 않지만 결과로 나오는 예측의 정확도에는 제한이 있을 수 있습니다.

상황을 평가하는 가장 빠른 방법은 데이터의 포괄적 감사를 수행하는 것입니다. 계속하기 전에 데이터 검토 노드를 데이터 소스에 첨부하고 이를 실행하여 전체 보고서를 생성하는 것을 고려하십시오.

잡음 데이터 유의

데이터에는 종종 오류가 포함되고 주관적이어서 변하는 판단이 포함될 수 있습니다. 이러한 현상을 집합적으로 **잡음**이라고 부릅니다. 때로 데이터에서 잡음은 일반적입니다. 이에 대한 기본 규칙이 있을 수 있으나 전체 케이스에 해당하는 것은 아닙니다.

일반적으로 데이터에 잡음이 많을수록 정확한 결과를 얻는 것이 어렵습니다. 그러나 IBM SPSS Modeler에서 시스템 학습 방법은 잡음 데이터를 처리하고 거의 50%의 잡음을 포함하는 데이터 세트에서 성공적으로 사용하였습니다.

충분한 데이터가 있는지 확인

데이터 마이닝에서 데이터 세트의 크기가 반드시 중요한 것은 아닙니다. 가능한 결과 범위와 변수 조합과 함께 데이터 세트의 대표성이 보다 중요합니다.

일반적으로 더 많은 속성을 고려할수록 대표 범위를 제공하기 위해서는 더 많은 레코드가 필요합니다.

데이터가 대표적이고 일반적인 기본 규칙이 있는 경우에는 수천(또는 수백) 개의 데이터 샘플로도 백만 개의 레코드와 유사한 좋은 결과를 내고 더 신속하게 결과를 얻을 수도 있습니다.

데이터에 대한 전문가 찾기

대부분의 경우에 사용자는 자신의 데이터에 대해 작업할 것이므로 그 내용과 의미에 상당히 친숙합니다. 그러나 조직이 다른 부서나 클라이언트의 데이터에 대해 작업하는 경우에는 데이터에 대해 알고 있는 전문가에게 문의

를 하는 것이 바람직합니다. 이들은 관련 속성의 식별을 안내해주고, 데이터 마이닝의 결과 해석을 돋고, 정보의 참 너깃을 "거짓 금" 또는 데이터 세트의 변칙에 의해 초래된 아티팩트를 구분하는 데 도움을 줄 수 있습니다.

데이터 마이닝 전략

대부분의 비즈니스 노력과 마찬가지로 데이터 마이닝은 계획적이고 체계적인 방법으로 수행하면 훨씬 더 효과적입니다. IBM SPSS Modeler와 같은 최신 데이터 마이닝 도구를 사용하더라도 데이터 마이닝의 대부분의 작업은 프로세스를 계속 추적하기 위해 식견 있는 비즈니스 분석가가 필요합니다. 계획을 안내하려면 다음 질문에 응답 하십시오.

- 해결하려는 실질적 문제점은 무엇입니까?
- 어떤 데이터 소스가 사용 가능하고 데이터의 어떤 부분이 현재 문제점과 관련됩니까?
- 데이터 마이닝을 시작하기 전에 어떤 유형의 전처리와 데이터 정리를 수행해야 합니까?
- 어떤 데이터 마이닝 기술을 사용합니까?
- 데이터 마이닝 분석의 결과를 어떻게 평가합니까?
- 데이터 마이닝에서 얻은 정보를 어떻게 최대한 활용합니까?

일반 데이터 마이닝 프로세스는 매우 신속하게 복잡해질 수 있습니다. 복잡한 비즈니스 문제, 여러 데이터 소스, 데이터 소스 간 다양한 데이터 품질, 데이터 마이닝 기술의 배열, 데이터 마이닝 성공을 측정하는 여러 방법 등 추적할 사항이 많습니다.

계속 추적하려면 데이터 마이닝을 위해 명시적으로 정의된 프로세스 모델을 가지고 있는 것이 도움이 됩니다. 프로세스 모델은 이 섹션의 앞에 나열된 질문에 응답하는데 도움을 주고 중요한 사항이 해결되었는지 확인합니다. 이는 복잡한 데이터를 분석할 때 길을 잊지 않도록 데이터 마이닝 로드맵의 역할을 합니다.

SPSS Modeler와 함께 사용하기 위해 제안된 데이터 마이닝 프로세스는 CRISP-DM(Cross-Industry Standard Process for Data Mining)입니다. 이름에서 알 수 있는 것처럼 이 모델은 다양한 산업과 비즈니스 문제점에 적용 할 수 있는 일반 모델로 계획된 것입니다.

CRISP-DM 프로세스 모델

일반 CRISP-DM 프로세스 모델에는 데이터 마이닝에서 주요 문제를 해결하는 6개의 단계를 포함합니다. 6개의 단계는 데이터 마이닝을 큰 비즈니스 사례에 통합하도록 계획된 순환적 프로세스에서 적합합니다.

6개의 단계는 다음을 포함합니다.

- **비즈니스 이해.** 이는 데이터 마이닝의 가장 중요한 단계입니다. 비즈니스 이해에는 비즈니스 목표 결정, 상황 평가, 데이터 마이닝 목표 결정 및 프로젝트 계획 생성이 포함됩니다.
- **데이터 이해.** 데이터는 데이터 마이닝의 "원시 자료"입니다. 이 단계에서는 데이터 자원과 이러한 자원의 특성이 무엇인지에 대한 이해의 필요성을 다룹니다. 여기에는 초기 데이터 수집, 데이터 설명, 데이터 탐색 및 데이터 품질 검증이 포함됩니다. 출력 노드 팔레트에서 사용 가능한 데이터 검토 노드는 데이터 이해를 위해 꼭 필요한 도구입니다.
- **데이터 준비.** 데이터 자원을 카탈로그한 후에는 데이터의 마이닝을 준비해야 합니다. 준비에는 데이터 선택, 정리, 구성, 통합 및 형식화가 포함됩니다.
- **모델링.** 물론, 이는 데이터 마이닝의 겉으로 보이는 부분이고, 데이터에서 정보를 추출하는 데에는 정교한 분석 방법을 사용합니다. 이 단계는 모델링 기술 선택, 검정 계획 생성 및 모델 작성 및 평가가 포함됩니다.
- **평가.** 모델을 선택한 후에는 데이터 마이닝 결과가 비즈니스 목표를 달성하는 데에 어떻게 도움을 줄 수 있는지를 평가할 준비가 된 것입니다. 이 단계의 요소에는 결과 평가, 데이터 마이닝 프로세스 검토 및 다음 단계 결정이 포함됩니다.
- **배포.** 이제 모든 노력을 투자했으므로 수확을 거둘 시간입니다. 이 단계에서는 원래 비즈니스 문제를 해결하기 위해 새 지식을 일상의 비즈니스 프로세스에 통합하는 데 초점을 맞춥니다. 이 단계에는 계획 배포, 모니터링 및 유지보수, 최종 보고서 생성 및 프로젝트 검토가 포함됩니다.

이 프로세스 모델에는 몇몇 핵심 사항이 있습니다. 첫째, 이전 단락에서 요약된 순서로 단계를 통해 프로세스가 플로우되도록 하기 위한 일반적 경향이 있는 한편, 단계가 비선형 방식으로 서로에게 영향을 미치는 여러 장소도 있습니다. 예를 들어, 데이터 준비는 일반적으로 모델링보다 앞섭니다. 그러나 모델링 단계 동안 수집된 정보와 결정된 의사결정은 종종 데이터 준비 단계의 일부를 다시 생각하게 만들 수도 있으므로 새로운 모델링 문제를 제

기할 수 있습니다. 2단계는 두 단계가 모두 적절하게 해결될 때까지 서로에게 피드백합니다. 마찬가지로, 평가 단계는 원래 비즈니스 이해를 재평가하도록 유도할 수 있고 잘못된 질문에 응답하려고 시도했다고 판단할 수도 있습니다. 이 시점에서 비즈니스 이해를 개정하고 더 나은 목표를 염두에 두고 나머지 프로세스를 다시 진행할 수 있습니다.

두 번째 키 포인트는 데이터 마이닝의 반복적인 특성입니다. 간단하게 데이터 마이닝 프로젝트를 계획하고, 완료한 다음 데이터를 싸서 집으로 돌아가는 일은 거의 드물 것입니다. 고객의 요구를 해결하기 위한 데이터 마이닝은 지속적인 노력입니다. 한 번의 데이터 마이닝 주기에서 얻어진 지식은 거의 대부분 고객의 요구를 식별하고 충족시키기 위한 새로운 질문, 새로운 문제, 새로운 기회로 이어집니다. 이러한 새로운 질문, 문제 및 기회는 일반적으로 데이터를 다시 한 번 마이닝하여 해결할 수 있습니다. 이 마이닝 및 새로운 기회 식별 프로세스는 비즈니스를 생각하는 방법의 일부이고 전체 비즈니스 전략의 초석이 되어야 합니다.

이러한 소개는 CRISP-DM 프로세스 모델의 간략한 개요만을 제공합니다. 모델에 대한 전체적인 세부사항은 다음 자원을 참조하십시오.

- CRISP-DM 안내서 - 설치 디스크의 *|Documentation* 폴더에서 다른 문서와 함께 액세스할 수 있습니다.
- 시작 메뉴에서 또는 IBM SPSS Modeler의 도움말 메뉴에서 **CRISP-DM** 도움말을 클릭하여 사용 가능한 CRISP-DM 도움말 시스템.

모델 유형

IBM SPSS Modeler는 기계 학습, 인공지능 및 통계로부터 취한 다양한 모델링 방법을 제공합니다. 모델링 팔레트에서 사용할 수 있는 이러한 방법을 통해 데이터로부터 새로운 정보를 얻어서 예측 모형을 개발할 수 있습니다. 각각의 방법은 그것만의 장점이 있으며 특정한 문제점 유형에 가장 적합합니다.

IBM SPSS Modeler 애플리케이션 안내서에서는 모델링 프로세스에 대한 일반적인 소개와 함께 이러한 여러 방법의 예제를 제공합니다. 이 안내서는 온라인 자습서 및 PDF 형식으로 사용 가능합니다. 자세한 정보는 [3 페이지의 『애플리케이션 예제』](#) 주제를 참조하십시오.

모델링 방법은 다음 범주로 나뉩니다.

- 감독
- 연관
- 세분화

감독 모델

감독 모델은 하나 이상의 출력 또는 목표 필드를 예측하기 위해 하나 이상의 입력 필드 값을 사용합니다. 이러한 기술의 일부 예는 다음과 같습니다. 의사결정 트리(C&R 트리, QUEST, CHAID 및 C5.0 알고리즘), 회귀분석(1차, 로지스틱, 일반화 선형 및 Cox 회귀 알고리즘), 신경망, 지원 벡터 머신 및 베이지안 네트워크입니다.

감독 모델은 조직이 예를 들어, 고객이 구매할지 또는 떠날지 여부 또는 트랜잭션이 알려진 사기 패턴과 매치하는지 여부 등과 같이 알려진 결과를 예측하는 데 도움을 줍니다. 모델링 기법은 시스템 학습, 규칙 귀납, 하위 그룹 식별, 통계 방법 및 다중 모델 생성을 포함합니다.

감독 노드



자동 분류자 노드는 이분형 결과(예 또는 아니오, 이탈 또는 이탈 안함 등)에 대해 다수의 여러 모델을 작성하고 비교하여 주어진 분석을 위한 최상의 접근 방식을 선택할 수 있게 합니다. 많은 모델링 알고리즘이 지원되어 사용할 방법, 각각에 대한 특정 옵션, 결과 비교 기준을 선택할 수 있습니다. 이 노드는 지정된 옵션을 기반으로 모델 세트를 생성하고 사용자가 지정하는 기준에 따라 최상의 후보를 순위화합니다.



자동 수치 노드는 수많은 방법을 사용하여 연속적 수치 범위 결과의 모델을 추정하고 비교합니다. 이 노드는 자동 분류자 노드에서와 같은 방식으로 작동하므로 사용할 알고리즘을 선택하고 단일 모델링 전달에서 여러 옵션의 조합을 실험할 수 있습니다. 지원되는 알고리즘에는 신경망, C&R 트리, CHAID, 선형 회귀, 일반화 선형 회귀 및 지원 벡터 머신(SVM)이 있습니다. 모델은 상관관계, 상대 오차 또는 사용된 변수의 수를 기반으로 비교할 수 있습니다.



분류 및 회귀(C&R) 트리 노드는 추가 관측값을 예측하거나 분류할 수 있게 하는 의사결정 트리를 생성합니다. 이 방법은 재귀적 파티셔닝을 사용하여 각 단계마다 불순도를 최소화하여 학습 레코드를 세그먼트로 분할합니다. 여기서 트리의 노드는 노드의 케이스의 100%가 대상 필드의 특정 범주에 속하면 "순수"로 간주됩니다. 목표 및 입력 필드는 숫자 범위 또는 범주형(명목형, 순서형 또는 플래그)입니다. 모든 분할은 이분형입니다(오직 두 개의 부집단).



QUEST 노드는 의사결정 트리를 작성하기 위한 이분형 분류 방법을 제공하며, 대형 C&R 트리 분석에 필요한 처리 시간을 줄이는 동시에 분류 트리 방법에서 찾은 경향을 줄여 더 많은 분할을 허용하는 입력을 선호하도록 설계되었습니다. 입력 필드는 숫자 범위(연속)일 수 있지만 대상 필드는 범주형이어야 합니다. 모든 분할은 이분형입니다.



CHAID 노드는 최적 분할을 식별하기 위해 카이제곱 통계량을 사용하여 의사결정 트리를 생성합니다. C&R 트리 및 QUEST 노드와 달리, CHAID는 비이분형 트리를 생성할 수 있으며 이는 일부 분할에 둘 이상의 분기가 있음을 의미합니다. 목표 및 입력 필드는 숫자 범위(연속형) 또는 범주형입니다. Exhaustive CHAID는 가능한 모든 분할을 탐색하는 보다 철저한 작업을 수행하지만 계산하는 데 시간이 더 걸리는 변형 CHAID입니다.



C5.0 노드는 의사결정 트리 또는 규칙 세트를 작성합니다. 모델은 각 수준에서 최대 정보 이익을 제공하는 필드를 기반으로 샘플을 분할하여 작동합니다. 대상 필드는 범주형이어야 합니다. 세 개 이상의 부집단으로의 다중 분할이 허용됩니다.



의사결정 목록 노드는 전체 채우기에 상대적인 주어진 이분형 결과의 상위 또는 하위 우도를 표시하는 부집단 또는 세그먼트를 식별합니다. 예를 들어, 캠페인을 이탈할 가능성이 없거나 우호적으로 응답할 가능성이 가장 많은 고객을 찾고 있습니다. 자체 사용자 정의 세그먼트를 추가하고 대체 모델을 나란히 미리보기하여 결과를 비교함으로써 비즈니스 지식을 모델에 통합할 수 있습니다. 의사결정 목록 모델은 각 규칙에 조건과 결과가 있는 규칙 목록으로 구성됩니다. 규칙은 순서대로 적용되며 매치하는 첫 번째 규칙이 결과를 결정합니다.



선형 회귀 모형은 목표와 하나 이상의 예측변수 간의 선형 관계를 기반으로 연속형 목표를 예측합니다.



PCA/요인 노드에서는 강력한 데이터 축소 기법을 제공하여 데이터의 복잡도를 줄입니다. 비선형 주성분분석(PCA)은 구성요소가 서로 직각(수직)인 전체 필드 세트에서 변동을 캡처하는 입력 필드의 선형 조합을 찾습니다. 요인 분석은 관측된 필드 세트 내에서 상관관계 패턴을 설명하는 기본 요인을 식별하려고 시도합니다. 두 접근 방식 모두 목표는 원래 필드 세트의 정보를 효과적으로 요약하는 적은 수의 파생 필드를 찾는 것입니다.



필드선택 노드는 기준(예: 결측값의 퍼센트) 세트를 기반으로 제거용 입력 필드를 차단합니다. 그런 다음 지정된 대상에 상대적인 남아 있는 입력의 중요도에 대해 순위를 매깁니다. 예를 들어, 수백 개의 잠재 입력이 있는 데이터 세트가 있다면 환자 결과 모델링 시 어느 것이 가장 유용합니까?



판별 분석은 로지스틱 회귀분석보다 엄격한 가정을 하지만 해당 가정이 충족되면 로지스틱 회귀 분석의 귀중한 대안 또는 보조물이 될 수 있습니다.



로지스틱 회귀분석은 입력 필드 값을 기반으로 레코드를 분류하는 통계 기법입니다. 선형 회귀와 유사하지만 숫자 범위 대신 범주형 대상 필드를 사용합니다.



일반화 선형 모델은 종속변수가 요인과 선형적으로 관련되고 지정된 연결함수를 통해 공변되도록 일반 선형 모델을 확장합니다. 더욱이 모델을 사용하면 종속변수가 비정규 분포를 가질 수 있습니다. 선형 회귀, 로지스틱 회귀분석, 카운트 데이터에 대한 로그선형 모델, 간격 중도절단 생존 모델을 포함하여 상당수 통계 모델의 기능을 포함합니다.



일반화 선형 혼합 모델(GLMM)은 목표가 비정규 분포를 가질 수 있고 지정된 연결함수를 통해 요인 및 공변량과 선형적으로 관련되며 관측값을 상관시킬 수 있도록 선형 모델을 확장합니다. 일반화 선형 혼합 모델은 단순 선형 회귀에서 비정규 장기적인 데이터에 대한 복합 다중 수준 모델에 이르기까지 다양한 모델을 포함합니다.



Cox 회귀 노드를 통해 중도절단된 레코드가 있는 데서 시간 대 이벤트 데이터에 대한 생존 모델을 작성할 수 있습니다. 이 모델은 주어진 입력 변수 값에 대해 주어진 시간(t)에 흥미 있는 이벤트가 발생한 확률을 예측하는 생존함수를 생성합니다.



지원 벡터 머신(SVM) 노드를 사용하면 데이터를 과적합 없이 두 개의 그룹 중 하나로 분류할 수 있습니다. SVM은 다수의 입력 필드가 있는 데이터 세트 등과 같은 광범위한 데이터 세트와 잘 작동합니다.



베이지안 네트워크 노드를 통해 관측 및 레코드된 증거를 실세계 지식과 조합하여 발생 우도를 확립함으로써 확률 모델을 작성할 수 있습니다. 이 노드는 주로 분류에 사용하는 TAN(Tree Augmented Naïve Bayes) 및 Markov Blanket 네트워크에 초점을 맞춥니다.



SLRM(Self-Learning Response Model) 노드를 사용하면 하나의 새 케이스 또는 소수의 새 케이스를 사용하여 모든 데이터를 사용하는 모델을 다시 학습시킬 필요 없이 모델을 재평가할 수 있는 모델을 작성할 수 있습니다.



시계열 노드는 시계열 데이터에 대한 지수평활, 일변량 자기회귀 통합 이동 평균(ARIMA), 다변량 ARIMA(또는 전이 함수) 모델을 추정하고 미래 성능을 위한 예측값을 생성합니다. 이 시계열 노드는 SPSS Modeler 버전 18에서 더 이상 사용되지 않는 이전의 시계열 노드와 유사합니다. 그러나 이 새 시계열 노드는 IBM SPSS Analytic Server의 기능을 이용하여 빅 데이터를 처리해서 SPSS Modeler 버전 17에 추가된 출력 뷰어에 결과 모델을 표시하도록 설계되었습니다.



KNN(k -Nearest Neighbor) 노드는 새 케이스를 k 가 정수인 예측자 공간에서 가장 가까이에 있는 k 오브젝트의 범주 또는 값과 연관시킵니다. 유사한 케이스는 서로 가까이에 있고 유사하지 않은 케이스는 서로 멀리 떨어져 있습니다.



STP(Spatio-Temporal Prediction) 노드는 위치 데이터, 예측(예측자)을 위한 입력 필드, 시간 필드 및 대상 필드를 포함하는 데이터를 사용합니다. 각 위치에는 각 측정 시간에 각 예측변수의 값을 나타내는 데이터에 여러 행이 있습니다. 데이터가 분석된 후에는 분석에 사용된 모양 데이터 내에서 어떤 위치에서든 목표 값을 예측하는 데 사용할 수 있습니다.

연관 모델

연관 모델은 하나 이상의 엔티티(예: 이벤트, 구매 또는 속성)가 하나 이상의 다른 엔티티와 연관되어 있는 데이터에서 패턴을 발견합니다. 모델은 이러한 관계를 정의하는 규칙 세트를 구성합니다. 여기에서 데이터 내의 필드는 입력과 목표 둘 모두의 역할을 할 수 있습니다. 이러한 연관을 수동으로 찾을 수 있지만 연관 규칙 알고리즘은 이를 보다 신속하게 수행하므로 더 복잡한 패턴을 탐색할 수 있습니다. Apriori 및 Carma 모델은 이러한 알고리즘 사용의 예입니다. 연관 모델의 또 다른 유형은 순차 발견 모델이며 이는 시간 구조 데이터에서 순차 패턴을 발견합니다.

연관 모델은 다중 결과를 예측할 때 가장 유용합니다(예: 제품 X를 구매한 고객이 Y와 Z도 구매함). 연관 모델은 특정 결론(예: 구매 결정)을 조건 세트와 연관시킵니다. 다른 표준 의사결정 트리 알고리즘(C5.0 및 C&RT)에 비

해 연관 규칙 알고리즘의 장점은 어떤 속성 사이에도 연관이 있을 수 있다는 점입니다. 의사결정 트리 알고리즘은 단일 결론만 포함하는 규칙을 작성하지만, 연관 알고리즘은 각각 다른 결론을 보유할 수 있는 많은 규칙을 찾으려고 합니다.

연관 노드



Apriori 노드는 데이터에서 규칙 세트를 추출하고 정보 내용이 가장 많은 규칙을 꺼냅니다. Apriori는 규칙을 선택하는 5개의 서로 다른 방법을 제공하며 정교한 색인화 스킵을 사용하여 대형 데이터 세트를 효율적으로 처리합니다. 큰 문제점의 경우, Apriori는 일반적으로 학습 속도가 빠릅니다. 보유할 수 있는 규칙 수에 임의 제한이 없으며 최대 32개의 전제 조건을 가진 규칙을 처리할 수 있습니다. Apriori에서는 입력 및 출력 필드가 모두 범주형이어야 하지만 이런 유형의 데이터에 최적화되어 있기 때문에 우수한 성능을 제공합니다.



CARMA 모델은 입력 또는 대상 필드를 지정하지 않아도 데이터에서 규칙 세트를 추출합니다. Apriori와 대조적으로 CARMA 노드는 단지 전향 지원이 아니라 규칙 지원(전향 및 후향 둘 다에 대한 지원)을 위한 작성 설정을 제공합니다. 이는 생성된 규칙을 보다 다양한 애플리케이션에 사용하여, 예를 들어 후향이 이번 휴가철에 홍보할 항목인 제품 또는 서비스 목록을 찾을 수 있음을 의미합니다.



순차규칙 노드는 순차 또는 시간 지향 데이터에서 연관 규칙을 발견합니다. 시퀀스는 예측 가능한 순서로 발생하는 경향이 있는 항목 세트 목록입니다. 예를 들어, 면도기와 애프터 쇼이브로션을 구매하는 고객은 다음 번 구매 시에 면도용 크림을 구매할 수 있습니다. 순차규칙 노드는 시퀀스를 찾는 데 효율적인 2阶段 방법을 사용하는 CARMA 연관 규칙 알고리즘을 기반으로 합니다.



연관 규칙 노드는 Apriori 노드와 유사합니다. 그러나 Apriori와는 달리, 연관 규칙 노드는 목록 데이터를 처리할 수 있습니다. 또한, 연관 규칙 노드는 빅 데이터를 처리하고 더 빠른 병렬 처리를 사용하기 위해 IBM SPSS Analytic Server와 함께 사용할 수 있습니다.

세분화 모델

세분화 모델은 데이터를 입력 필드의 패턴이 유사한 레코드의 세그먼트 또는 군집으로 나눕니다. 이들은 입력 필드에만 관심이 있으므로 세분화 모델에는 출력이나 대상 필드와 같은 개념이 없습니다. 세분화 모델의 예는 코호넨 네트워크, K-평균 군집, 이단계 군집 및 이상 항목 발견입니다.

세분화 모델("군집 모델"로도 알려짐)은 특정 결과가 알려지지 않은 경우에 유용합니다. 예를 들어, 새로운 사기 패턴을 식별할 때나, 고객 기반에 있는 관심 그룹을 식별할 때입니다. 군집 모델은 유사한 레코드 그룹을 식별하고 레코드에 이들이 속하는 그룹에 따라 레이블을 붙이는 데 초점을 둡니다. 이는 그룹과 그룹의 특성에 대한 사전 지식 없이도 수행되고, 예측할 모델에 대한 사전에 정의된 출력이나 대상 필드가 없다는 면에서 군집 모델을 다른 모델링 기법과 구별해줍니다. 이러한 모델에는 옳고 그른 응답이 없습니다. 이를 값은 데이터에서 관심 있는 집단을 캡처하고 이러한 집단에 대한 유용한 설명을 제공하는 기능으로 판별됩니다. 군집 모델은 종종 군집이나 세그먼트를 작성하는 데 사용하고 이러한 군집이나 세그먼트는 이후의 분석에서 입력으로 사용합니다(예: 잠재 고객을 동종 하위그룹으로 분할하는 방법으로).

세분화 노드



자동 군집 노드는 유사한 특성을 가진 레코드 그룹을 식별하는 군집 모델을 추정하고 비교합니다. 이 노드는 다른 자동 모델링 노드와 동일한 방법으로 작동하여 단일 모델링 패스에서 다중 옵션 조합을 실험할 수 있습니다. 군집 모델의 유용성을 필터링하고 순위화하며 특정 필드의 중요성을 기반으로 측도를 제공하려고 시도하는 기본 측도를 사용하여 모델을 비교할 수 있습니다.



K-평균 노드는 데이터 세트를 별개의 그룹(또는 군집)으로 군집화합니다. 이 방법은 고정된 수의 군집을 정의하고 반복적으로 레코드를 군집에 지정하며, 추가 세분화가 더 이상 모델을 향상시킬 수 없을 때까지 군집 중심을 조정합니다. 결과를 예상하는 대신 k-평균은 자율 학습으로 알려진 프로세스를 사용하여 입력 필드 세트의 패턴을 찾아냅니다.



코호넨 노드는 데이터 세트를 별개의 그룹으로 군집화하는 데 사용할 수 있는 신경망 유형을 생성합니다. 네트워크가 완전히 숙달되면, 유사 레코드는 출력 맵 가까이 있지만, 다른 레코드는 멀리 떨어져 있을 것입니다. 모델 너깃에서 각 단위별로 캡처된 관측값을 살펴 강한 단위를 식별할 수 있습니다. 이것은 적당한 군집 수에 대한 감각을 제공할 것입니다.



TwoStep 노드는 2단계 군집 방법을 사용합니다. 첫 번째 단계는 원시 입력 데이터를 관리 가능한 하위 군집 세트로 압축하기 위해 데이터를 통한 단일 전달을 수행합니다. 두 번째 단계는 계층적 군집 방법을 사용하여 하위 군집을 점점 더 큰 군집으로 계속해서 병합하는 것입니다. TwoStep은 학습 데이터에 대한 최적 군집 수를 자동으로 평가하는 장점이 있습니다. 혼합 필드 유형과 대형 데이터 세트를 효율적으로 처리할 수 있습니다.



이상 항목 발견 노드는 "정상" 데이터 패턴을 따르지 않는 특이 케이스 또는 이상값을 식별합니다. 이 노드를 사용하면 이전에 알려진 패턴에 적합하지 않고, 찾고 있는 패턴을 정확하게 모르더라도 이상값을 식별할 수 있습니다.

In-Database 마이닝 모델

IBM SPSS Modeler에서 Oracle Data Miner 및 Microsoft Analysis Services를 포함하여 데이터베이스 벤더에서 제공하는 데이터 마이닝과 모델링 도구의 통합을 지원합니다. 모두 IBM SPSS Modeler 애플리케이션 내에서 시작하여 데이터베이스 내에 모델을 작성하고, 스코어링하고, 저장할 수 있습니다. 자세한 정보는 *IBM SPSS Modeler In-Database 마이닝* 안내서를 참조하십시오.

IBM SPSS Statistics 모델

컴퓨터에 IBM SPSS Statistics 사본이 설치되고 사용이 허가된 경우에는 IBM SPSS Modeler 내에서 특정 IBM SPSS Statistics 루틴에 액세스하고 실행하여 모델을 작성하고 스코어할 수 있습니다.

데이터 마이닝 예제

실제로 데이터 마이닝에 대해 배우는 최상의 방법은 예제로 시작하는 것입니다. 특정 모델링 방법과 기술을 중점적으로 간략히 소개하는 *IBM SPSS Modeler* 애플리케이션 안내서에는 다양한 애플리케이션 예제가 포함되어 있습니다. 자세한 정보는 [3 페이지의 『애플리케이션 예제』](#)의 내용을 참조하십시오.

제 5 장 스트림 작성

스트림 작성 개요

IBM SPSS Modeler를 사용하는 데이터 마이닝은 **스트림**이라고 하는 일련의 노드를 통해 데이터를 실행하는 프로세스에 초점을 맞춥니다. 이 일련의 노드는 데이터에 대해 수행될 작업을 나타내는 반면, 노드 간의 링크는 데이터 플로우의 방향을 표시합니다. 일반적으로, 데이터 스트림을 사용하여 IBM SPSS Modeler로 데이터를 읽어들이고 일련의 조작을 통해 실행한 후 테이블이나 뷰어와 같은 대상으로 보냅니다.

예를 들어, 데이터 소스를 열고 새 필드를 추가하며, 새 필드의 값을 기반으로 레코드를 선택한 후 결과를 테이블에 표시하려 한다고 가정하십시오. 이 경우, 데이터 스트림은 다음 4개의 노드로 구성됩니다.



가변파일 노드 - 데이터 소스에서 데이터를 읽도록 설정합니다.



파생 노드 - 데이터 세트에 계산된 새 필드를 추가하는 데 사용합니다.



선택 노드 - 데이터 스트림에서 레코드를 제외할 선택기준을 설정하는 데 사용합니다.



테이블 노드 - 조작 결과를 화면에 표시하는 데 사용합니다.

데이터 스트림 작성

고유한 SPSS Modeler 인터페이스를 사용하여 데이터 스트림의 디어그램에 대해 작업함으로써 데이터를 시작적으로 마이닝할 수 있습니다. 가장 기본적인 수준에서 다음 단계를 사용하여 데이터 스트림을 작성할 수 있습니다.

- 노드를 스트림 캔버스에 추가하십시오.
- 스트림을 형성하도록 노드를 연결하십시오.
- 특정 노드 또는 스트림 옵션을 지정하십시오.
- 스트림을 실행하십시오.

이 절에는 노드를 사용하여 작업함으로써 보다 복잡한 데이터 스트림을 작성하는 것에 대한 자세한 정보가 포함되어 있습니다. 또한 노드 및 스트림에 대한 옵션 및 설정에 대해 논의합니다. 프로그램 설치의 [Demos 폴더에](#) SPSS Modeler와 함께 제공되는 데이터를 사용한 스트림 작성의 단계별 예는 3 페이지의 [『애플리케이션 예제』](#)의 내용을 참조하십시오.

노드에 대한 작업

노드는 IBM SPSS Modeler에서 데이터를 탐색하는 데 도움이 됩니다. 작업공간의 다양한 노드는 서로 다른 오브젝트와 조치를 나타냅니다. IBM SPSS Modeler 창의 맨 아래에 있는 팔레트에는 스트림 작성에 사용하는 가능한 모든 노드가 포함됩니다.

몇몇 유형의 노드가 있습니다. **소스 노드**는 데이터를 스트림으로 가져오고 노드 팔레트의 소스 탭에 있습니다. **프로세스 노드**는 개별 데이터 레코드와 필드에서 조작을 수행하고 팔레트의 레코드 작업과 필드 작업 탭에서 발견할 수 있습니다. **출력 노드**는 데이터, 차트 및 모델 결과의 다양한 출력을 생성하고 노드 팔레트의 그래프, 출력 및 내보내기 탭에 포함됩니다. **모델링 노드**는 통계적 알고리즘을 사용하여 모델 너깃을 작성하고 노드 팔레트의 모

델링 탭 및 데이터베이스 모델링 탭(있는 경우)에 있습니다. 자세한 정보는 [12 페이지의 『노드 팔레트』](#)의 내용을 참조하십시오.

노드를 연결하여 스트림을 형성할 수 있습니다. 스트림은 실행되면 관계를 시각화하고 결론을 그릴 수 있게 해줍니다. 스트림은 스크립트와 같아서 이를 저장하고 다른 데이터 파일과 함께 재사용할 수 있습니다.

스트림 데이터를 처리하는 실행 가능한 노드는 **터미널 노드**로 알려져 있습니다. 모델링 또는 출력 노드는 스트림이나 스트림 분기의 끝에 있으면 터미널 노드입니다. 터미널 노드에 추가로 노드를 연결할 수 없습니다.

참고: 노드 팔레트를 사용자 정의할 수 있습니다. 자세한 정보는 [213 페이지의 『노드 팔레트 사용자 정의』](#)의 내용을 참조하십시오.

스트림에 노드 추가

노드 팔레트에서 스트림에 노드를 추가하는 몇 가지 방법이 있습니다.

- 팔레트에서 노드를 두 번 클릭하십시오. 참고: 노드를 두 번 클릭하면 현재 스트림에 노드가 자동으로 연결됩니다. 자세한 정보는 [30 페이지의 『스트림의 노드 연결』](#)의 내용을 참조하십시오.
- 팔레트에서 스트림 캔버스로 노드를 끌어서 놓으십시오.
- 팔레트에서 노드를 클릭한 후 스트림 캔버스를 클릭하십시오.
- IBM SPSS Modeler의 삽입 메뉴에서 해당 옵션을 선택하십시오.

스트림 캔버스에 노드를 추가했으면 노드를 두 번 클릭하여 해당 대화 상자를 표시하십시오. 사용 가능한 옵션은 추가하고 있는 노드 유형에 따라 다릅니다. 대화 상자 내 특정 제어에 대한 정보는 해당 도움말 단추를 클릭하십시오.

노드 제거

데이터 스트림에서 노드를 제거하려면 노드를 클릭하고 Delete 키를 누르거나 마우스 오른쪽 단추를 클릭하고 메뉴에서 **삭제**를 선택하십시오.

스트림의 노드 연결

스트림 캔버스에 추가된 노드는 연결될 때까지 데이터 스트림을 형성하지 않습니다. 노드 간의 연결은 하나의 작업에서 다른 작업으로 데이터가 흘로우될 때 데이터의 방향을 표시합니다. 노드를 연결하여 스트림을 형성하는 많은 방법이 있는데, 두 번 클릭, 마우스 가운데 단추 사용 또는 수동으로 이를 수행할 수 있습니다.

두 번 클릭으로 노드 추가 및 연결 방법

스트림을 형성하는 가장 간단한 방법은 팔레트에서 노드를 두 번 클릭하는 것입니다. 이 방법은 새 노드를 스트림 캔버스에서 선택된 노드에 자동으로 연결합니다. 예를 들어, 캔버스에 데이터베이스 노드가 포함된 경우 이 노드를 선택한 후 팔레트에서 다음 노드(예: 파생 노드)를 두 번 클릭할 수 있습니다. 이 동작은 파생 노드를 기존 데이터베이스 노드에 자동으로 연결합니다. 마지막 비터미널 노드 업스트림에 새 노드가 연결되는 지점인 터미널 노드(예: 히스토그램 또는 테이블 노드)에 도달할 때까지 이 프로세스를 반복할 수 있습니다.

마우스 가운데 단추를 사용하여 노드를 연결하는 방법

스트림 캔버스에서 마우스 가운데 단추를 사용하여 한 노드를 다른 노드로 클릭하여 끌 수 있습니다(마우스에 가운데 단추가 없으면 한 노드에서 다른 노드로 마우스를 끄는 동안 Alt 키를 눌러 이를 시뮬레이션할 수 있음).

노드를 수동으로 연결하는 방법

마우스 가운데 단추가 없고 노드를 수동으로 연결하려면 노드 팝업 메뉴를 사용하여 노드를 캔버스에 이미 있는 다른 노드에 연결할 수 있습니다.

1. 연결을 시작할 노드를 마우스 오른쪽 단추로 클릭하십시오. 이렇게 하면 노드 메뉴가 열립니다.
2. 메뉴에서 **연결**을 클릭하십시오.
3. 시작 노드와 커서 둘 다에 연결 아이콘이 표시됩니다. 캔버스에서 또 하나의 노드를 클릭하여 두 노드를 연결하십시오.

노드 연결 시 따라야 할 몇 가지 지침이 있습니다. 다음 유형의 연결을 수행하려고 시도하면 오류 메시지를 수신합니다.

- 소스 노드로 이어지는 연결

- 터미널 노드에서 이어지는 연결
- 최대 입력 연결 수보다 많은 노드
- 이미 연결된 두 개의 노드 연결
- 순환(데이터가 이미 플로우된 노드에 리턴됨)

스트림의 노드 우회

데이터 스트림의 노드를 우회하는 경우 모든 입력 및 출력 연결이 입력 노드에서 출력 노드로 직접 이끄는 연결로 대체됩니다. 노드에 입력 연결과 출력 연결이 둘 다는 없는 경우 재라우팅되지 않고 해당 연결이 모두 삭제됩니다.

예를 들어, 새 필드를 파생시키고 필드를 필터링한 후 히스토그램과 테이블에서 결과를 탐색하는 스트림이 있습니다. 필드가 필터링되기 전에 데이터에 대해서도 동일한 그래프와 테이블을 보려는 경우 새 히스토그램 및 테이블 노드를 스트림에 추가하거나 필터 노드를 우회할 수 있습니다. 필터 노드 우회 시 그래프 및 테이블로의 연결은 파생 노드에서 직접 전달됩니다. 필터 노드는 스트림에서 연결 해제됩니다.

노드 우회 방법

1. 스트림 캔버스에서 가운데 마우스 단추를 사용하여 우회할 노드를 두 번 클릭하십시오. 또는 Alt+두 번 클릭을 사용할 수 있습니다.

참고: 편집 메뉴에서 실행 취소를 클릭하거나 Ctrl+Z를 눌러 이 조치를 실행 취소할 수 있습니다.

스트림의 노드 사용 안함

스트림 내에서 단일 입력이 있는 처리 노드를 사용 안함으로 설정할 수 있는데, 그 결과 스트림 실행 중 노드를 무시합니다. 이로 인해 노드를 제거하거나 우회하지 않아도 되며, 나머지 노드에 계속 연결할 수 있음을 의미합니다. 여전히 노드 설정을 열고 편집할 수 있습니다. 그러나 노드를 다시 사용 가능하게 해야 변경사항이 적용됩니다.

예를 들어, 몇 개의 필드를 필터링한 후 줄어든 데이터 세트로 모델을 작성하는 스트림이 있습니다. 필드를 필터링하지 않고 같은 모델을 작성하여 모델 결과가 향상되는지 여부를 확인하려고 할 경우 필터 노드를 사용 안함으로 설정할 수 있습니다. 필터 노드를 사용 안함으로 설정하면 모델링 노드와의 연결은 파생 노드에서 유형 노드로 직접 전달됩니다.

노드를 사용 안함으로 설정하는 방법

1. 스트림 캔버스에서 사용 안함으로 설정할 노드를 마우스 오른쪽 단추로 클릭하십시오.
2. 도구 메뉴에서 노드 사용 안함을 클릭하십시오.

또는 편집 메뉴에서 노드 > 노드 사용 안함을 클릭할 수 있습니다. 스트림에 노드를 다시 포함시키려면 같은 방법으로 노드 사용을 클릭하십시오.

참고: 편집 메뉴에서 실행 취소를 클릭하거나 Ctrl+Z를 눌러 이 조치를 실행 취소할 수 있습니다.

기존 연결의 노드 추가

두 개의 노드를 연결하는 화살표를 끌어서 연결된 두 개의 노드 사이에 새 노드를 추가할 수 있습니다.

1. 가운데 마우스 단추를 사용하여 노드를 삽입할 위치로 연결 화살표를 클릭하여 끄십시오. 또는 클릭하여 끄는 동안 Alt 키를 눌러 가운데 마우스 단추를 시뮬레이션할 수 있습니다.
2. 포함시킬 노드로 연결을 끌고 마우스 단추를 놓으십시오.

참고: 노드에서 새 연결을 제거하고 노드를 우회하여 원래 값을 복원할 수 있습니다.

노드 간의 연결 삭제

두 노드 간의 연결을 삭제하려면 다음을 수행하십시오.

1. 연결 화살표를 마우스 오른쪽 단추로 클릭하십시오.
2. 메뉴에서 연결 삭제를 클릭하십시오.

노드 간의 연결을 모두 삭제하려면 다음 중 하나를 수행하십시오.

- 노드를 선택하고 F3을 누르십시오.
 - 기본 메뉴에서 노드를 선택하고 다음을 클릭하십시오.
- 편집 > 노드 > 연결 해제

노드 옵션 설정

노드를 작성하고 연결한 후, 노드를 사용자 정의하는 몇 가지 옵션이 있습니다. 노드를 마우스 오른쪽 단추로 클릭하고 메뉴 옵션 중 하나를 선택하십시오.

- 편집을 클릭하여 선택한 노드의 대화 상자를 여십시오.
- 연결을 클릭하여 한 노드를 다른 노드에 수동으로 연결하십시오.
- 연결 해제를 클릭하여 노드에서 모든 링크를 삭제하십시오.
- 이름 변경 및 주석달기를 클릭하여 편집 대화 상자의 주석(Annotation) 탭을 여십시오.
- 새 주석을 클릭하여 노드와 관련된 주석을 추가하십시오. 자세한 정보는 [46 페이지의 『노드와 스트림에 주석과 주석\(Annotation\) 추가』](#)의 내용을 참조하십시오.
- 노드 사용 안함을 클릭하여 처리 중 노드를 "숨기십시오". 노드를 처리를 위해 다시 표시하려면 노드 사용을 클릭하십시오. 자세한 정보는 [31 페이지의 『스트림의 노드 사용 안함』](#)의 내용을 참조하십시오.
- 잘라내기 또는 삭제를 클릭하여 스트림에서 선택한 노드를 제거하십시오. 참고: 잘라내기를 클릭하면 노드를 붙여넣을 수 있지만 삭제는 그렇게 할 수 없습니다.
- 노드 복사를 클릭하여 연결 없이 노드의 사본을 작성하십시오. 이는 신규 또는 기존 스트림에 추가할 수 있습니다.
- 노드 로드를 클릭하여 이전에 저장된 노드를 열고 해당 옵션을 현재 선택된 노드로 로드하십시오. 노드는 동일한 유형이어야 합니다.
- 노드 검색을 클릭하여 연결된 IBM SPSS Collaboration and Deployment Services Repository에서 노드를 검색하십시오.
- 노드 저장을 클릭하여 노드의 세부사항을 파일에 저장하십시오. 노드 세부사항을 동일 유형의 또 다른 노드로 로드할 수 있습니다.
- 노드 저장을 클릭하여 선택한 노드를 연결된 IBM SPSS Collaboration and Deployment Services Repository에 저장하십시오.
- 캐시를 클릭하여 선택한 노드를 캐싱하기 위한 옵션과 함께 메뉴를 펼치십시오.
- 데이터 맵핑을 클릭하여 데이터를 새 소스에 맵핑하고 필수 필드를 지정하기 위한 옵션과 함께 메뉴를 펼치십시오.
- 수퍼노드 작성을 클릭하여 현재 스트림에서 수퍼노드를 작성하기 위한 옵션과 함께 메뉴를 펼치십시오.
- 사용자 입력 노드 생성을 클릭하여 선택한 노드를 대체하십시오. 이 노드에 의해 생성된 예제에는 현재 노드와 동일한 필드가 있습니다.
- 여기부터 실행을 클릭하여 선택한 노드로부터 모든 터미널 노드 다운스트림을 실행하십시오.

노드 캐싱 옵션

실행 중인 스트림을 최적화하기 위해 비터미널 노드에서 캐시를 설정할 수 있습니다. 노드에서 캐시를 설정하면 다음 번에 데이터 스트림을 실행할 때 노드를 통과하는 데이터로 캐시가 채워집니다. 그때부터는 데이터 소스가 아닌 캐시(디스크의 임시 디렉토리에 저장됨)에서 데이터를 읽습니다.

캐싱은 정렬, 병합 또는 통합과 같이 시간이 걸리는 작업 다음으로 가장 유용합니다. 예를 들어, 데이터베이스에서 판매 데이터를 읽는 소스 노드 세트와 지역별로 판매액을 요약하는 통합 노드가 있다고 가정하십시오. 전체 데이터 세트가 아니라 통합된 데이터를 저장할 캐시를 원하기 때문에 소스 노드가 아닌 통합 노드에서 캐시를 설정할 수 있습니다.

참고: 소스 노드에서 캐싱하면(IBM SPSS Modeler로 데이터를 읽을 때 원 데이터 세트의 사본만 저장함) 대부분의 상황에서 성능이 향상되지는 않습니다.

캐싱을 사용하는 노드에는 오른쪽 상단 모서리에 작은 문서 아이콘이 표시됩니다. 노드에서 데이터가 캐싱될 때 문서 아이콘은 녹색입니다.

캐시 사용 방법

1. 스트림 캔버스에서 노드를 마우스 오른쪽 단추로 클릭하고 메뉴에서 **캐시**를 클릭하십시오.
2. 캐싱 하위 메뉴에서 **사용**을 클릭하십시오.
3. 노드를 마우스 오른쪽 단추로 클릭하고 캐싱 하위 메뉴에서 **사용 안함**을 클릭하여 캐시를 설정 해제할 수 있습니다.

데이터베이스의 노드 캐싱

데이터베이스에서 실행되는 스트림의 경우, 스트림 중간에 파일 시스템이 아닌 데이터베이스의 임시 테이블에 데이터를 캐싱할 수 있습니다. SQL 최적화와 결합되면 이로 인해 성능이 상당히 향상됩니다. 예를 들어, 다중 테이블을 병합하여 데이터 마이닝 보기를 작성하는 스트림의 출력을 캐싱하고 필요에 따라 재사용할 수 있습니다. 모든 다운스트림 노드에 대해 SQL을 자동으로 생성하면 성능이 추가로 향상될 수 있습니다.

데이터베이스 캐싱을 이용하려면 SQL 최적화와 데이터베이스 캐싱 둘 다를 사용으로 설정해야 합니다. 서버 최적화 설정은 클라이언트에서 해당 설정을 대체합니다. 자세한 정보는 [37 페이지의 『스트림의 최적화 옵션 설정』](#)의 내용을 참조하십시오.

데이터베이스 캐싱을 사용으로 설정한 상태에서는 비터미널 노드를 마우스 오른쪽 단추로 클릭하여 해당 지점에서 데이터를 캐싱하면 되며, 다음 번에 스트림이 실행될 때 데이터베이스에 자동으로 캐시가 직접 작성됩니다. 데이터베이스 캐싱 또는 SQL 최적화를 사용할 수 없으면 캐시는 파일 시스템에 대신 기록됩니다.

참고: Db2, Oracle, SQL Server 및 Teradata에서는 캐싱 목적으로 임시 테이블을 지원합니다. Netezza와 같은 기타 데이터베이스는 데이터베이스 캐싱에 정규 테이블을 사용합니다. 특정 데이터베이스에 대해 SQL 코드를 사용자 정의할 수 있습니다. 지원을 받으려면 서비스팀에 문의하십시오.

캐시 비우기 방법

노드의 흰색 문서 아이콘은 해당 캐시가 비어 있음을 표시합니다. 캐시가 가득 차면 문서 아이콘이 녹색 단색이 됩니다. 캐시 내용을 대체하려면 먼저 캐시를 비운 후 데이터 스트림을 재실행하여 다시 채우십시오.

1. 스트림 캔버스에서 노드를 마우스 오른쪽 단추로 클릭하고 메뉴에서 **캐시**를 클릭하십시오.
2. 캐싱 하위 메뉴에서 **비우기**를 클릭하십시오.

캐시 저장 방법

캐시 내용을 IBM SPSS Statistics 데이터 파일(*.sav)로 저장할 수 있습니다. 그런 다음 파일을 캐시로 재로드하거나 캐시 파일을 데이터 소스로 사용하는 노드를 설정할 수 있습니다. 또한 다른 프로젝트에서 저장한 캐시를 로드할 수 있습니다.

1. 스트림 캔버스에서 노드를 마우스 오른쪽 단추로 클릭하고 메뉴에서 **캐시**를 클릭하십시오.
2. 캐싱 하위 메뉴에서 **캐시 저장**을 클릭하십시오.
3. 캐시 저장 대화 상자에서 캐시 파일을 저장할 위치를 찾아보십시오.
4. 파일 이름 텍스트 상자에 이름을 입력하십시오.
5. 파일 유형 목록에서 ***.sav**가 선택되어 있는지 확인하고 **저장**을 클릭하십시오.

캐시 로드 방법

노드에서 제거하기 전에 캐시 파일을 저장했으면 재로드할 수 있습니다.

1. 스트림 캔버스에서 노드를 마우스 오른쪽 단추로 클릭하고 메뉴에서 **캐시**를 클릭하십시오.
2. 캐싱 하위 메뉴에서 **캐시 로드**를 클릭하십시오.
3. 캐시 로드 대화 상자에서 캐시 파일 위치를 찾아 선택한 후 **로드**를 클릭하십시오.

노드에서 데이터 미리보기

스트림을 작성할 때 예상한 방법으로 데이터가 변경 중인지 확인하려면 중요 단계마다 테이블 노드를 통해 데이터를 실행하면 됩니다. 그렇게 하지 않으려면 각 노드에서 작성되는 데이터의 샘플을 표시하는 미리보기를 생성하면 됩니다. 그러면 각 노드를 작성하는 시간이 줄어듭니다.

모델 너깃의 노드 업스트림의 경우 미리보기는 입력 필드를 표시합니다. 모델 너깃이나 너깃의 노드 다운스트림의 경우(터미널 노드는 제외) 미리보기는 입력 및 생성된 필드를 표시합니다.

표시되는 기본 행 수는 10입니다. 그러나 스트림 특성에서 이를 변경할 수 있습니다. 자세한 정보는 [35 페이지의 『스트림의 일반 옵션 설정』](#)의 내용을 참조하십시오.

생성 메뉴에서 몇몇 노드 유형을 작성할 수 있습니다.

참고: 이 노드에 의해 생성된 데이터를 미리 보면 모든 특성 변경사항이 이 노드에 적용되며 취소할 수 없습니다 (적용을 클릭하는 것과 동일한 작동입니다).

노드 잠금

다른 사용자가 스트림에서 하나 이상의 노드 설정을 수정하지 못하도록 하려면 수퍼노드라는 특수 노드 유형으로 노드를 캡슐화한 다음 비밀번호 보호를 적용하여 수퍼노드를 잠그면 됩니다.

스트림에 대한 작업

스트림 캔버스에서 소스, 프로세스 및 터미널 노드를 연결하면 스트림이 작성됩니다. 노드의 콜렉션으로서 스트림은 저장되고, 주석(Annotation)이 작성되고 프로젝트에 추가됩니다. 최적화, 날짜 및 시간 설정, 모수 및 스크립트 등과 같이 스트림에 여러 옵션을 설정할 수도 있습니다. 이러한 특성은 다음과 같은 주제에서 설명됩니다.

IBM SPSS Modeler에서, 동일한 IBM SPSS Modeler 세션에서 둘 이상의 데이터 스트림을 사용하고 수정할 수 있습니다. 기본 창의 오른쪽에는 관리자 분할창이 있고 이는 현재 열려 있는 스트림, 출력 및 모델을 탐색하는데 도움을 줍니다. 관리자 분할창이 보이지 않으면 보기 메뉴에서 관리자를 클릭한 다음 스트림 탭을 클릭하십시오.

이 탭에서 다음을 수행할 수 있습니다.

- 스트림에 액세스합니다.
- 스트림을 저장합니다.
- 스트림을 현재 프로젝트에 저장합니다.
- 스트림을 닫습니다.
- 새 스트림을 엽니다.
- 스트림을 저장하고 IBM SPSS Collaboration and Deployment Services 리포지토리(사용자의 사이트에서 사용 가능한 경우)에서 스트림을 검색합니다. 자세한 정보는 [175 페이지의 『IBM SPSS Collaboration and Deployment Services Repository 정보』](#)의 내용을 참조하십시오.

스트림 탭에서 스트림을 마우스 오른쪽 단추로 클릭하고 이러한 옵션에 액세스합니다.

스트림을 위한 옵션 설정

현재 스트림에 적용하기 위해 수많은 옵션을 지정할 수 있습니다. 이러한 옵션을 모든 스트림에 적용하기 위한 기본값으로 저장할 수도 있습니다. 옵션은 다음과 같습니다.

- **일반.** 스트림에 사용할 기호 및 텍스트 인코딩과 같은 기타 옵션. 자세한 정보는 [35 페이지의 『스트림의 일반 옵션 설정』](#)을 참조하십시오.
- **날짜/시간.** 날짜 및 시간 표현식의 형식과 관련된 옵션. 자세한 정보는 [36 페이지의 『스트림의 날짜 및 시간 옵션 설정』](#)을 참조하십시오.
- **숫자형식.** 숫자표현식의 형식을 제어하는 옵션. 자세한 정보는 [37 페이지의 『스트림의 숫자 형식 옵션 설정』](#)을 참조하십시오.
- **최적화.** 스트림 성능을 최적화하기 위한 옵션. 자세한 정보는 [37 페이지의 『스트림의 최적화 옵션 설정』](#)을 참조하십시오.
- **로깅 및 상태.** SQL 로깅 및 레코드 상태를 제어하는 옵션. 자세한 정보는 [38 페이지의 『스트림의 SQL 로깅 및 레코드 상태 옵션 설정』](#)을 참조하십시오.
- **레이아웃.** 캔버스에서 스트림의 레이아웃과 관련된 옵션. 자세한 정보는 [39 페이지의 『스트림의 레이아웃 옵션 설정』](#)을 참조하십시오.
- **Analytic Server.** SPSS Modeler와 함께 Analytic Server의 사용과 관련된 옵션. 자세한 정보는 [39 페이지의 『Analytic Server 스트림 특성』](#)을 참조하십시오.

- **지리 공간적.** 스트림에 사용하기 위한 지리 공간적 데이터 형식과 관련된 옵션. 자세한 정보는 [39 페이지의 『스트림의 지리 공간적 옵션 설정』](#)을 참조하십시오.

스트림 옵션 설정

1. 파일 메뉴에서 **스트림 특성**을 클릭하십시오. 또는 관리자 분할창의 스트림 탭에서 스트림을 선택하고 마우스 오른쪽 단추를 클릭한 다음 팝업 메뉴에서 **스트림 특성**을 클릭하십시오.

2. **옵션** 탭을 클릭하십시오.

또는 도구 메뉴에서 다음을 클릭하십시오.

스트림 특성 > 옵션

스트림의 일반 옵션 설정

일반 옵션은 현재 스트림의 다양한 측면에 적용되는 기타 옵션 세트입니다.

기본 섹션에는 다음 기본 옵션이 포함됩니다.

- **소수점 기호.** 쉼표(.) 또는 마침표(.)를 소수점 구분자로 선택하십시오.
- **그룹 기호.** 숫자 표시 형식의 경우 값을 그룹화하는 데 사용하는 기호를 선택하십시오(예: 3,000.00에서 쉼표). 옵션에는 없음, 마침표, 쉼표, 공백 및 정의된 로케일이 포함됩니다(이 경우 현재 로케일의 기본값을 사용함).
- **인코딩.** 텍스트 인코딩을 위한 스트림 기본 방법을 지정하십시오. (참고: 변수에 적용됩니다. 파일 소스 노드와 플랫 파일 내보내기 노드에만 해당합니다. 다른 노드는 이 설정을 사용하지 않습니다. 대부분의 데이터 파일에 임베드된 인코딩 정보가 있습니다.) 시스템 기본값 또는 UTF-8을 선택할 수 있습니다. 시스템 기본값은 Windows 제어판에 지정되어 있거나 분산 모드에서 실행 중인 경우 서버 컴퓨터에 지정되어 있습니다. 자세한 정보는 [233 페이지의 『IBM SPSS Modeler에서 유니코드 지원』](#)의 내용을 참조하십시오.
- **규칙 세트 평가.** 규칙 세트 모델이 평가되는 방법을 결정합니다. 기본적으로 규칙 세트는 **투표**를 사용하여 개별 규칙으로부터 예측을 조합하고 최종 예측을 결정합니다. 해당 규칙 세트가 기본적으로 처음 적중하는 규칙을 사용하도록 하려면 처음 적중을 선택하십시오. 이 옵션은 항상 알고리즘이 정의한 첫 번째 적중을 사용하는 의사결정 목록 모델에는 적용되지 않음을 유의하십시오.

데이터 미리보기에 표시할 최대 행 수. 노드에 대해 데이터 미리보기가 요청될 때 표시될 행 수를 지정합니다. 자세한 정보는 [33 페이지의 『노드에서 데이터 미리보기』](#)의 내용을 참조하십시오.

명목 필드의 최대 수. 필드의 데이터 유형이 **유형 없음** 상태가 되기 전까지의 명목(세트) 필드의 최대 멤버 수를 지정하려면 선택하십시오. 이 옵션은 대규모의 명목 필드에서 작업할 때 유용합니다. 참고: 필드의 측정 수준이 **유형 없음**으로 설정되면 해당 역할은 자동으로 **없음**으로 설정됩니다. 즉 필드를 모델링에 사용할 수 없음을 의미합니다.

코호넨, K-평균 모델링의 변수군 크기 제한. 코호넨 네트과 K-평균 모델링에 사용된 명목 필드의 최대 멤버 수를 지정하려면 선택하십시오. 기본 변수군 크기는 20이고, 20이 넘으면 필드는 무시되고 경고가 발행되어 문제의 필드에 대한 정보를 제공합니다.

호환성을 위해서 이 옵션은 또한 IBM SPSS Modeler의 14 버전에 대체된 이전 신경망 노드에도 적용됩니다. 몇몇 레거시 스트림에는 여전히 이 노드가 포함될 수 있습니다.

실행 시 소스 노드 새로 고침. 현재 스트림을 실행할 때 모든 소스 노드를 자동으로 새로 고치려면 선택하십시오. 이 동작은 소스 노드에서 **새로 고침** 단추를 클릭하는 것과 유사합니다. 단지 이 옵션은 현재 스트림의 모든 소스 노드(사용자 입력 노드는 제외)를 자동으로 새로 고친다는 점이 다릅니다.

참고: 이 옵션을 선택하면 데이터가 변경되지 않았더라도 다운스트림 노드의 캐시가 비워집니다. 도구 모음에서 **현재 스트림 실행**을 실행할 경우, 비우기는 스트림 실행 중 한 번만 발생하지만, 이는 다운스트림 캐시를 여전히 단일 실행 시 임시 저장 공간으로서 사용할 수 있음을 의미합니다. 예를 들어, 복합 파생 작업 후에 캐시 미드스트림을 설정했고 이 파생 노드의 다운스트림에 몇몇 그래프와 보고서가 첨부되었다고 가정해 보십시오. 스트림을 실행할 때 파생 노드에 있는 캐시는 첫 번째 그래프 또는 보고서에 대해서만 비워지고 다시 채워집니다. 후속 터미널 노드는 파생 노드 캐시로부터 데이터를 읽습니다. (2개 이상의 터미널 노드가 있을 때) **현재 스트림 실행**을 사용하는 대신 각각의 터미널 노드를 개별적으로 실행하도록 선택하는 경우, 캐시 비우기는 사용자가 터미널 노드를 실행할 때마다 발생한다는 점을 참고하십시오.

출력에 필드와 값 레이블 표시. 테이블, 차트 및 기타 출력에 필드와 값 레이블을 표시합니다. 레이블이 없는 경우에는 필드 이름과 데이터 값이 대신에 표시됩니다. 레이블은 기본적으로 꺼져 있습니다. 그러나 IBM SPSS

Modeler의 다른 곳에서 개별적으로 레이블을 토글할 수 있습니다. 도구 모음에서 사용 가능한 토글 단추를 사용하여 출력 레이블을 표시하도록 선택할 수도 있습니다.



그림 11. 필드와 값 레이블을 토글하는 데 사용하는 도구 모음 아이콘

실행 시간 표시. 스트림이 실행된 후 실행 시간 탭에서 스트림 노드의 개별 실행 시간을 표시합니다. 자세한 정보는 41 페이지의 『실행 시간 보기』의 내용을 참조하십시오.

자동 노드 작성 섹션에는 개별 스트림에서 노드를 자동으로 작성하기 위해 다음 옵션을 포함합니다. 이러한 옵션은 새 너깃을 생성할 때 스트림 캔버스에 모델링 너깃을 삽입하는지 여부를 제어합니다. 기본적으로 이러한 옵션은 버전 16 이상에서 작성된 스트림에만 적용됩니다. IBM SPSS Modeler 16 이상에서 버전 15 이전에서 작성된 스트림을 열고 모델링 노드를 실행하면 너깃은 이전 릴리스에서 사용된 것처럼 스트림 캔버스에 배치되지 않습니다. IBM SPSS Modeler 16 이상을 사용하여 새 스트림을 작성하고 모델링 노드를 실행하는 경우 생성된 너깃이 스트림 캔버스에 배치됩니다. 이는 예를 들어, 새 모델 출력의 모델 적용 노드 작성 옵션이 IBM SPSS Collaboration and Deployment Services 및 IBM SPSS Modeler Server 클라이언트 사용자 인터페이스가 없는 다른 환경에서 배치로 실행된 앞의 16개 스트림을 구분하기 때문입니다.

- **새 모델 출력의 모델 적용 노드 작성.** 새 모델 출력의 모델 적용 노드를 자동으로 작성합니다. 이 옵션을 선택하면 링크를 사용, 사용 안함 또는 작성하지 않음으로 설정하는지 여부와 관계없이 모델 업데이트 링크 작성에서 선택할 수도 있습니다.

새 모델 적용자 또는 소스 노드가 작성되면 드롭 다운 제어의 링크 옵션은 작성자 노드와 새 노드 간의 업데이트 링크가 작성되는지 여부를 제어하고, 작성되는 경우 어떤 모드에 있는지를 제어합니다. 링크가 작성되면 이를 사용 가능하게 만들 수 있지만 이러한 옵션은 사용자에게 완전한 제어를 제공합니다.

- **소스 작성기에서 소스 노드 작성.** 소스 작성기에서 소스 노드를 자동으로 작성합니다. 이전 옵션과 유사하게, 이 옵션을 선택하면 새로 고침 링크가 사용, 사용 안함으로 설정되는지 또는 이를 작성하지 않는지 여부와 관계없이 소스 새로 고침 링크 작성 드롭 다운에서 선택할 수도 있습니다.

기본값으로 저장. 지정된 옵션은 현재 스트림에만 적용됩니다. 이러한 옵션을 모든 스트림의 기본값으로 설정하려면 이 단추를 클릭하십시오.

스트림의 날짜 및 시간 옵션 설정

이러한 옵션은 현재 스트림에서 다양한 날짜 및 시간 표현식에 사용하기 위한 형식을 지정합니다.

다른 이름으로 날짜/시간 가져오기 - 날짜/시간 필드에 날짜/시간 저장 공간을 사용할지 또는 이를 문자열 변수로 가져올지 여부를 선택하십시오.

날짜 형식 - 날짜 저장 공간 필드에 사용하거나 문자열이 CLEM 날짜 함수를 통해 날짜로 해석될 때 사용할 날짜 형식을 선택하십시오.

시간 형식 - 시간 저장 공간 필드에 사용하거나 문자열이 CLEM 시간 함수를 통해 시간으로 해석될 때 사용할 시간 형식을 선택하십시오.

롤오버 일/분 - 시간 형식에서 음의 시간 차이가 이전 일 또는 시간을 가리키는 것으로 해석할지 여부를 선택하십시오.

날짜 기준선(1월 1일) - 단일 날짜를 처리하는 CLEM 날짜 함수가 사용할 기준선 연도(항상 1월 1일)를 선택하십시오.

두자리 시작 날짜 - 두 자릿수로 표시되는 연도에 세기 자릿수를 추가하기 위한 분리점 연도를 지정하십시오. 예를 들어, 1930을 분리점 연도로 지정하면 05/11/02는 2002년에 있다고 가정합니다. 이와 같은 설정은 30년 후의 날짜에도 20세기를 사용합니다. 따라서 05/11/73은 1973년으로 가정됩니다.

시간대 - datetime_now 표현식에서 사용할 시간대의 선택 방법을 선택하십시오.

- 서버를 선택할 경우 시간대가 다음 항목에 따라 달라집니다.

- 현재 스트림이 Analytic Server 데이터 소스를 사용할 경우 datetime_now 표현식에는 Analytic Server의 시간이 사용됩니다. 기본적으로 서버는 협정 세계시(UTC)를 사용합니다.
- 현재 스트림이 데이터베이스 소스 노드를 사용할 경우 지원되는 데이터베이스에는 SQL 푸시백이 사용되고, datetime_now 표현식에는 데이터베이스의 시간이 사용됩니다.

- 기타 모든 스트림의 경우 시간대가 SPSS Modeler Server의 시간을 사용합니다.
 - **Modeler 클라이언트**를 선택할 경우 SPSS Modeler가 설치된 시스템의 시간대 세부사항이 시간대에 적용됩니다.
 - 또는 시간대에 대해 협정 세계시(UTC) 값을 선택할 수 있습니다.
- 기본값으로 저장** - 현재 스트림에만 적용되는 옵션입니다. 이러한 옵션을 모든 스트림의 기본값으로 설정하려면 이 단추를 클릭하십시오.

스트림의 숫자 형식 옵션 설정

이러한 옵션은 현재 스트림에서 다양한 숫자 표현식에 사용하기 위한 형식을 지정합니다.

숫자 표시 형식. 표준(**#####.####**), 지수표기(**#.####E+##**) 또는 통화 표시 형식(**\$###.##**)에서 선택할 수 있습니다.

소수점이하자리수(표준, 지수표기, 통화). 숫자 표시 형식의 경우 실수를 표시하거나 인쇄할 때 사용할 소수점이하자리수를 지정합니다. 이 옵션은 각 표시 형식마다 별도로 지정됩니다.

계산 위치. 삼각법 CLEM 표현식에 측정 단위로 사용할 **라디안** 또는 **도**를 선택하십시오. 자세한 정보는 [155 페이지의 『삼각 함수』](#)의 내용을 참조하십시오.

기본값으로 저장. 지정된 옵션은 현재 스트림에만 적용됩니다. 이러한 옵션을 모든 스트림의 기본값으로 설정하려면 이 단추를 클릭하십시오.

스트림의 최적화 옵션 설정

최적화 설정을 사용하여 스트림 성능을 최적화할 수 있습니다. IBM SPSS Modeler Server(사용하는 경우)의 성능 및 최적화 설정은 클라이언트의 해당 설정을 대체합니다. 이러한 설정을 서버에서 사용할 수 없는 경우 클라이언트는 이를 사용할 수 없습니다. 그러나 서버에서 사용으로 설정된 경우에는 클라이언트가 사용 안함으로 설정하도록 선택할 수 있습니다.

참고: 데이터베이스 모델링 및 SQL 최적화를 위해서는 IBM SPSS Modeler 컴퓨터에서 IBM SPSS Modeler Server 연결이 사용으로 설정되어 있어야 합니다. 이 설정이 사용 가능하면 데이터베이스 알고리즘에 액세스하고, SQL을 IBM SPSS Modeler에서 직접 푸시백하고, IBM SPSS Modeler Server에 액세스할 수 있습니다. 현재 라이센스 상태를 검증하려면 IBM SPSS Modeler 메뉴에서 다음을 선택하십시오.

도움말 > 정보 > 추가 세부사항

연결성이 사용 가능으로 설정되면 라이센스 상태 탭에 **서버 사용 가능** 옵션이 표시됩니다.

자세한 정보는 [8 페이지의 『IBM SPSS Modeler Server에 연결』](#)을 참조하십시오.

참고: SQL 푸시백 및 최적화 지원 여부는 사용 중인 데이터베이스 유형에 따라 다릅니다. IBM SPSS Modeler에서 사용할 수 있도록 검정되고 지원되는 데이터베이스 및 ODBC 드라이버에 대한 최신 정보는 회사 지원 사이트 (<http://www.ibm.com/support>)를 참조하십시오.

스트림 재작성 사용. IBM SPSS Modeler에서 스트림 재작성을 사용하려면 이 옵션을 선택하십시오. 4가지 유형의 재작성을 사용할 수 있으며 이 중 하나 이상을 선택할 수 있습니다. 스트림 재작성은 스트림 시맨틱을 변경하지 않고 보다 효율적인 작업을 위해 막후에서 스트림의 노드를 재정렬합니다.

- **SQL 생성 최적화**. 이 옵션을 사용하면 데이터베이스에서의 실행에 SQL 생성을 사용하여 더 많은 작업을 푸시백할 수 있도록 스트림 내에서 노드를 재정렬할 수 있습니다. SQL로 렌더링할 수 없는 노드를 발견하면 최적화 프로그램이 SQL로 렌더링할 수 있고 스트림 시맨틱에 영향을 주지 않고 문제점 노드 앞에 안전하게 이동할 수 있는 다운스트림 노드가 있는지 여부를 미리 확인합니다. 데이터베이스가 IBM SPSS Modeler 보다 효율적으로 작업을 수행할 수 있을 뿐만 아니라 이러한 푸시백은 처리를 위해 IBM SPSS Modeler로 리턴되는 데이터의 크기를 줄입니다. 이는 결국 네트워크 트래픽을 줄이고 스트림 작업 속도를 높일 수 있습니다. SQL 최적화가 효과를 보려면 **SQL 생성 확인** 상자를 선택해야 합니다.
- **CLEM 표현식 최적화**. 이 옵션을 통해 최적화 프로그램은 처리 속도를 높이기 위해 스트림이 실행되기 전에 사전 처리할 수 있는 CLEM 표현식을 검색할 수 있습니다. 단순한 예로, *log(salary)*와 같은 표현식이 있으면 최적화 프로그램이 실제 급여 값을 계산하고 처리를 위해 이를 전달합니다. 이를 사용하여 SQL 푸시백과 IBM SPSS Modeler Server 성능을 모두 향상시킬 수 있습니다.

- **구문 실행 최적화.** 이러한 스트림 재작성 방법은 IBM SPSS Statistics 구문이 포함된 두 개 이상의 노드를 통합하는 작업의 효율성을 증대합니다. 구문 명령문을 각각 별도의 작업으로 실행하는 대신 단일 작업에 결합하여 최적화를 달성합니다.
- **기타 실행 최적화.** 이러한 스트림 재작성 방법은 데이터베이스에 위임할 수 없는 작업의 효율성을 증대합니다. 스트림의 데이터 양을 가능한 조기에 줄여 최적화를 달성합니다. 데이터 무결성을 유지하는 동안 스트림이 재작성되어 데이터 소스 가까이에 작업을 푸시하므로 결합과 같이 비용이 많이 드는 작업에 대한 데이터 다운스트림을 줄입니다.

병렬 처리 사용. 다중 프로세서를 가진 컴퓨터에서 실행할 때 이 옵션을 사용하면 시스템이 해당 프로세서 전반에 서 로드 균형을 맞출 수 있으므로 성능이 향상됩니다. 다중 노드를 사용하거나 개별 노드(C5.0, 병합(키별), 정렬, 구간(순위 및 분위수 방법), 통합(하나 이상의 키 필드 사용))를 사용하면 병렬 처리를 통해 이익을 얻을 수 있습니다.

SQL 생성. 실행 프로세스를 생성하는 데 SQL 코드를 사용하여 데이터베이스에 스트림 작업을 푸시백(이렇게 하면 성능이 향상됨)할 수 있도록 SQL 생성을 사용하려면 이 옵션을 선택하십시오. 성능을 추가로 개선하려면 **SQL 생성 최적화**도 선택하여 데이터베이스에 푸시백되는 작업 수를 최대화할 수 있습니다. 노드에 대한 작업이 데이터베이스에 푸시백된 경우, 스트림이 실행될 때 노드가 보라색으로 강조 표시됩니다.

- **데이터베이스 캐싱.** 데이터베이스에서 실행될 SQL을 생성하는 스트림의 경우, 스트림 중간에 파일 시스템이 아닌 데이터베이스의 임시 테이블에 데이터를 캐싱할 수 있습니다. SQL 최적화와 결합되면 이로 인해 성능이 상당히 향상됩니다. 예를 들어, 다중 테이블을 병합하여 데이터 마이닝 보기를 작성하는 스트림의 출력을 캐싱하고 필요에 따라 재사용할 수 있습니다. 데이터베이스 캐싱을 사용으로 설정한 상태에서는 비터미널 노드를 마우스 오른쪽 단추로 클릭하여 해당 지점에서 데이터를 캐싱하면 되며, 다음 번에 스트림이 실행될 때 데이터베이스에 자동으로 캐시가 직접 작성됩니다. 따라서 다운스트림 노드에 대해 SQL이 생성되며 성능이 추가로 향상됩니다. 또는 필요하면 이 옵션을 사용 안 함으로 설정할 수 있습니다(예: 정책 또는 권한으로 인해 데이터베이스에 데이터를 기록하지 못하는 경우). 데이터베이스 캐싱 또는 SQL 최적화를 사용할 수 없으면 캐시는 파일 시스템에 대신 기록됩니다. 자세한 정보는 [32 페이지의 『노드 캐싱 옵션』](#) 주제를 참조하십시오.
- **완화된 변환 사용.** 이 옵션을 사용하면 적절한 형식으로 저장된 경우 문자열에서 숫자로 또는 숫자에서 문자열로의 데이터 변환이 가능합니다. 예를 들어, 데이터가 데이터베이스에 문자열로 유지되지만 실제로는 의미 있는 숫자를 포함하는 경우 푸시백이 발생할 때 사용하도록 데이터를 변환할 수 있습니다.

참고: SQL 구현에서의 사소한 차이 때문에 데이터베이스에서 실행되는 스트림이 IBM SPSS Modeler에서 실행될 때 리턴되는 것과 약간 다른 결과를 리턴할 수 있습니다. 비슷한 이유로 이러한 차이는 데이터베이스 벤더에 따라서도 다릅니다.

기본값으로 저장. 지정된 옵션은 현재 스트림에만 적용됩니다. 이러한 옵션을 모든 스트림의 기본값으로 설정하려면 이 단추를 클릭하십시오.

스트림의 SQL 로깅 및 레코드 상태 옵션 설정

이러한 설정에는 스트림이 생성한 SQL문의 표시와 스트림이 처리한 레코드 수의 표시를 제어하는 다양한 옵션을 포함합니다.

SQL을 스트림 실행 동안에 메시지 로그에 표시. 스트림을 실행하는 동안 생성된 SQL이 메시지 로그로 전달되는지 여부를 지정합니다.

스트림 전파 동안에 메시지 로그에 SQL 생성 세부사항 표시. 스트림 미리보기 중에 생성된 SQL의 미리보기가 메시지 로그에 전달되는지 여부를 지정합니다.

SQL 표시. 로그에 표시되는 SQL에 SPSS Modeler가 생성하는 기본 SQL 함수 또는 {fn FUNC(...)} 양식의 표준 ODBC 함수가 포함되는지 여부를 지정합니다. 전자는 구현되지 않을 수도 있는 ODBC 드라이버 기능에 의존합니다.

가독성 향상을 위해 SQL 다시 형식화. 로그에 표시된 SQL이 가독성을 위해 형식화되어야 하는지 여부를 지정합니다.

레코드의 상태 표시. 레코드가 터미널 노드에 도착할 때 언제 레코드를 보고해야 하는지를 지정합니다. 상태를 N 개의 레코드마다 업데이트하는 데 사용하는 숫자를 지정합니다.

기본값으로 저장. 지정된 옵션은 현재 스트림에만 적용됩니다. 이러한 옵션을 모든 스트림의 기본값으로 설정하려면 이 단추를 클릭하십시오.

스트림의 레이아웃 옵션 설정

이러한 설정은 스트림 캔버스의 표시 및 사용과 관련된 수많은 옵션을 제공합니다.

최소 스트림 캔버스 너비. 스트림 캔버스의 최소 너비를 픽셀 단위로 지정합니다.

최소 스트림 캔버스 높이. 스트림 캔버스의 최소 높이를 픽셀 단위로 지정합니다.

스트림 스크롤링 속도. 노드를 캔버스에서 한 위치에서 다른 위치로 끌 때 스트림 캔버스 분할창이 스크롤링되는 속도를 제어하기 위한 스트림 캔버스의 스크롤링 비율을 지정합니다. 숫자가 높을수록 더 빠른 스크롤링 속도를 지정합니다.

아이콘 이름 최대값. 스트림 캔버스에서 노드 이름의 문자 수 제한을 지정합니다.

아이콘 크기. 전체 스트림 보기를 표준 아이콘 크기의 8%와 200% 사이의 크기의 숫자 중 하나로 배율 조정하려면 옵션을 선택하십시오.

눈금 셀 크기. 목록에서 눈금 셀 크기를 선택합니다. 이 숫자는 개별 눈금을 사용하여 스트림 캔버스에서 노드를 정렬하는 데 사용합니다. 기본 눈금 셀 크기는 0.25입니다.

눈금에 정렬. 아이콘을 보이지 않는 눈금 패턴으로 맞추려면 선택합니다(기본적으로 선택됨).

생성된 아이콘 배치. 캔버스에서, 모델 너깃에서 생성된 노드의 아이콘을 배치할 위치를 선택합니다. 기본값은 왼쪽 상단입니다.

기본값으로 저장. 지정된 옵션은 현재 스트림에만 적용됩니다. 이러한 옵션을 모든 스트림의 기본값으로 설정하려면 이 단추를 클릭하십시오.

Analytic Server 스트림 특성

이러한 설정은 Analytic Server와 함께 작업하기 위한 많은 옵션을 제공합니다.

Analytic Server 외부에서 처리할 최대 레코드 수

Analytic Server 데이터 소스에서 SPSS Modeler 서버로 가져올 최대 레코드 수를 지정하십시오.

노드를 Analytic Server에서 처리할 수 없을 때 알림

이 설정은 Analytic Server에 제출될 스트림에 Analytic Server에서 처리할 수 없는 노드가 포함될 때 어떤 일이 발생하는지를 판별합니다. 경고를 발행하고 스트림을 계속 처리할지 오류를 발행하고 처리를 중지할지 여부를 지정합니다.

분할 모델 저장 공간 설정

모델 크기(MB)를 초과하는 경우 Analytic Server에 참조로 분할 모델 저장

모델 너깃은 일반적으로 스트림의 일부로 저장됩니다. 여러 분할이 있는 분할 모델은 대량의 너깃을 생성할 수 있으며 스트림과 Analytic Server 간에 앞뒤로 너깃을 이동하면 성능에 영향을 미칠 수 있습니다.

솔루션으로서, 분할 모델이 지정된 크기를 초과하면 이는 Analytic Server에 저장되고, SPSS Modeler의 너깃에는 모델에 대한 참조가 포함됩니다.

실행이 완료된 후에 Analytic Server에서 참조별로 모델을 저장하기 위한 기본 폴더

Analytic Server에 분할 모델을 저장하려는 기본 경로를 지정하십시오. 경로는 유효한 Analytic Server 프로젝트 이름으로 시작해야 합니다.

승격된 모델을 저장할 폴더

"승격된" 모델을 저장하려는 기본 경로를 지정하십시오. 승격된 모델은 SPSS Modeler 세션이 종료될 때 정리되지 않습니다.

스트림의 지리 공간적 옵션 설정

모양, 좌표 또는 단일 축 값(x 또는 y, 또는 위도와 경도)인지 관계없이 모든 지리 공간적 필드에는 연관된 좌표계가 있습니다. 이 좌표계는 속성을 원점(0,0) 및 값과 연관된 단위로 설정합니다.

수많은 좌표계가 있고 유형에는 지리적 및 투영의 두 가지가 있습니다. SPSS Modeler에서 모든 공간 함수는 투영된 좌표계와 함께만 사용할 수 있습니다.

좌표계의 특성으로 인해, 두 개의 별개의 지리 공간적 데이터 소스에서 데이터를 병합하거나 붙여쓰려면 소스가 동일한 좌표계를 사용해야 합니다. 이로 인해 스트림에 사용된 지리 공간적 데이터에 좌표 설정을 지정해야 합니다.

데이터는 자동으로 다음 상황에서 선택된 스트림 좌표계를 사용하도록 다시 투영됩니다.

- 공간 함수(예: 영역, *closeto*, *within*)의 경우, 함수에 전달된 모수는 자동으로 다시 투영됩니다. 그러나 원본 행 데이터는 변경되지 않은 상태로 남아 있습니다.
- STP(Spatio-Temporal Prediction)에서 작성 또는 스코어링(너깃) 노드를 사용할 때 위치 필드는 자동으로 재 투영됩니다. 스코어링할 때 너깃에서 나오는 위치는 원본 위치입니다.
- 맵 시각화 노드를 사용할 때.

스트림 좌표계. 확인 상자를 선택하는 경우에만 사용 가능합니다. 사용 가능한 투영된 좌표계 목록을 표시하고 현재 스트림에 사용하려는 좌표계를 선택하려면 **변경**을 클릭합니다.

기본값으로 저장. 선택하는 좌표계는 현재 스트림에만 적용됩니다. 스트림을 모든 스트림의 기본값으로 선택하려면 이 단추를 클릭하십시오.

지리 공간적 좌표계 선택

SPSS Modeler에서 모든 공간 함수는 투영된 좌표계와만 함께 사용할 수 있습니다.

스트림 좌표계 선택 대화 상자에는 스트림에 사용하는 지리 공간적 데이터를 위해 선택할 수 있는 모든 투영된 좌표계 목록이 포함됩니다.

다음 정보는 각 좌표계에 나열되어 있습니다.

- **WKID** 각 좌표계에 고유한 잘 알려진 ID.
- 이름 좌표계의 이름.
- 단위 좌표계와 연관된 측정 단위.

모든 좌표계의 목록에 추가로 대화 상자에는 **필터링** 제어가 있습니다. 필요한 좌표계 이름의 모두 또는 일부를 알고 있는 경우에는 이를 대화 상자 아래쪽의 **이름** 필드에 입력하십시오. 선택할 수 있는 좌표계 목록은 입력한 텍스트를 포함하는 이름이 있는 시스템만을 표시하기 위해 자동으로 필터링됩니다.

스트림 작업 메시지 보기

모델 작성과 평가 실행, 최적화 및 경과 시간 등과 같이 스트림 작업과 관련된 메시지는 스트림 특성 대화 상자의 메시지 탭을 사용하여 쉽게 볼 수 있습니다. 오류 메시지는 또한 이 테이블에 보고됩니다.

스트림 메시지 보기

1. 파일 메뉴에서 **스트림 특성**을 클릭하십시오. 또는 관리자 분할창의 **스트림** 탭에서 스트림을 선택하고 마우스 오른쪽 단추를 클릭한 다음 팝업 메뉴에서 **스트림 특성**을 클릭하십시오.
2. **메시지** 탭을 클릭하십시오.

또는 도구 메뉴에서 다음을 클릭하십시오.

스트림 특성 > 메시지

스트림 작업에 관한 메시지에 추가로 오류 메시지도 여기에 보고됩니다. 오류로 인해 스트림 실행이 종료되면 이 대화 상자 오류 메시지가 표시된 상태로 메시지 탭으로 열립니다. 또한 오류가 있는 노드는 스트림 캔버스에서 빨간색으로 강조 표시됩니다.

SQL 최적화와 로깅 옵션이 사용자 옵션 대화 상자에서 사용 가능한 경우에는 생성된 SQL에 대한 정보 또한 표시됩니다. 자세한 정보는 [37 페이지의 『스트림의 최적화 옵션 설정』](#)의 내용을 참조하십시오.

저장 단추 드롭 다운 목록에서(메시지 탭 바로 아래 왼쪽에서) **메시지 저장**을 클릭하여 스트림에 대해 보고된 메시지를 여기에 저장할 수 있습니다.

드롭 다운에서 **모든 메시지 지우기**를 선택하여 지정된 스트림에 대한 모든 메시지를 지울 수 있습니다.

CPU 시간은 서버 프로세스가 CPU를 이용하는 시간임을 유의하십시오. 경과 시간은 실행 시작과 실행 종료 사이의 총 시간이므로 파일 전송과 출력 렌더링 등을 포함합니다. CPU 시간은 스트림이 다중 CPU(병렬 실행)를 사용할 때는 경과 시간보다 길 수 있습니다. 스트림이 데이터 소스로 사용 중인 데이터베이스로 완전히 푸시백되면 CPU 시간은 0이 됩니다.

실행 시간 보기

메시지 탭에서 IBM SPSS Modeler Server에서 실행되는 스트림의 모든 노드에 대한 개별 실행 시간을 볼 수 있는 실행 시간을 표시하도록 선택할 수 있습니다. 시간이 R 또는 Analytic Server 등과 같은 다른 영역에서 스트림 실행이 정확하지 않을 수도 있음을 유의하십시오. 또한 일부 노드의 실행 시간은 계산할 수 없습니다.

참고: 이 기능이 작동하려면 옵션 탭의 일반 설정에서 실행 시간 표시 옵션이 선택되어야 합니다.

노드 실행 시간의 테이블에서 열은 다음과 같습니다. 열 머리말을 클릭하여 항목을 오름차순이나 내림차순으로 정렬하십시오(예를 들어, 어떤 노드가 실행 시간이 가장 긴지 확인하기 위해).

터미널 노드. 노드가 속하는 분기의 ID. ID는 분기의 끝에 있는 터미널 노드의 이름입니다.

노드 레이블. 실행 시간이 참조하는 노드의 이름입니다.

노드 ID. 실행 시간이 참조하는 노드의 고유 ID입니다. 이 ID는 노드가 작성될 때 시스템에 의해 생성됩니다.

실행 시간. 이 노드를 실행하는 데 걸리는 시간(초). 스트림을 실행할 때 데이터를 준비하고 출력에서 데이터를 검색하는 데 시간이 걸리고 이 유형의 시간은 계산할 수 없으므로 실행 시간은 일반 메시지에 표시되는 시간과 다를 수 있습니다.

스트림과 세션 모수 설정

CLEM 표현식과 스크립팅에서 사용할 모수를 정의할 수 있습니다. 사실상 모수는 현재 스트림, 세션 또는 수퍼노드와 함께 저장되고 지속되는 사용자 정의 변수이며 스크립팅을 통해서는 물론 사용자 인터페이스에서도 액세스 할 수 있습니다. 예를 들어, 스트림을 저장하면 해당 스트림에 대해 설정된 모수도 저장됩니다(이는 선언된 스크립트에서만 사용할 수 있는 로컬 스크립트 변수와 모수를 구별함). 모수는 스크립트에 하드 코드화하지 않아도 되는 필드 및 값에 대한 정보를 제공하여 스크립팅에서 스크립트 동작을 제어하는데 사용하는 경우가 많습니다.

모수 범위는 모수가 설정된 위치에 따라 다릅니다.

- 스트림 모수는 스트림 스크립트 또는 스트림 특성 대화 상자에서 설정할 수 있으며 스트림의 모든 노드에 사용 할 수 있습니다. 표현식 작성기의 모수 목록에 표시됩니다.
- 세션 모수는 독립형 스크립트 또는 세션 모수 대화 상자에서 설정할 수 있습니다. 현재 세션에서 사용하는 모든 스트림(관리자 분할창의 스트림 탭에 나열된 모든 스트림)에 사용할 수 있습니다.

수퍼노드에 대해서도 모수를 설정할 수 있는데, 이 경우 해당 수퍼노드 내에 캡슐화된 노드에만 표시됩니다.

사용자 인터페이스를 통해 스트림 및 세션 모수 설정

1. 스트림 모수를 설정하려면 기본 메뉴에서 다음을 클릭하십시오.

도구 > 스트림 특성 > 모수

2. 세션 모수를 설정하려면 도구 메뉴에서 **세션 모수 설정**을 클릭하십시오.

프롬프트 표시 여부. 사용자가 런타임 시에 이 모수의 값을 입력하도록 프롬프트하려면 이 상자를 선택하십시오.

이름 모수 이름이 여기 나열됩니다. 이 필드에 이름을 입력하여 새 모수를 작성할 수 있습니다. 예를 들어, 최저 기온에 대한 모수를 작성하려면 minvalue를 입력할 수 있습니다. CLEM 표현식에서 모수를 나타내는 \$P- 접두문자를 포함시키지 마십시오. 이 이름은 CLEM 표현식 작성기에 표시하는 데도 사용합니다.

긴 이름. 작성된 각 모수에 대한 설명 이름을 나열합니다.

저장 공간. 목록에서 저장 유형을 선택하십시오. 저장 공간은 모수에 데이터 값이 저장되는 방법을 표시합니다. 예를 들어, 유지할 선행 0이 포함된 값(예: 008)에 대한 작업 시 저장 유형으로 **문자열**을 선택해야 합니다. 그렇지 않으면, 값에서 0이 제거됩니다. 사용 가능한 저장 유형은 문자열, 정수, 실수, 시간, 날짜, 시간소인입니다. 날짜 모수의 경우, 다음 단락에 표시된 대로 ISO 표준 표기법을 사용하여 값을 지정해야 합니다.

값. 각 모수의 현재 값을 나열합니다. 필요에 따라 모수를 조정하십시오. 데이터 모수의 경우, ISO 표준 표기법 (즉, YYYY-MM-DD)으로 값을 지정해야 합니다. 다른 형식으로 지정된 날짜는 허용되지 않습니다.

유형(선택사항). 외부 애플리케이션에 스트림을 배포할 계획이면 목록에서 측정 수준을 선택하십시오. 그렇지 않으면 유형 열을 있는 그대로 두는 것이 바람직합니다. 모수의 값 제한조건(예: 숫자 범위의 상한 및 하한)을 지정 하려면 목록에서 **지정**을 선택하십시오.

사용자 인터페이스를 통해서만 모수에 대해 긴 이름, 저장 공간, 유형 옵션을 설정할 수 있습니다. 이러한 옵션은 스크립트를 사용하여 설정할 수 없습니다.

선택된 모수를 사용 가능한 모수 목록 위, 아래로 추가로 이동하려면 오른쪽에 있는 화살표를 선택하십시오. 선택된 모수를 제거하려면 삭제 단추(X로 표시)를 사용하십시오.

모수 값의 런타임 프롬프트 지정

서로 다른 경우에 동일한 모수에 대해 서로 다른 값을 입력해야 하는 스트림이 있는 경우에는 하나 이상의 스트림 또는 세션 모수 값에 런타임 프롬프트를 지정할 수 있습니다.

모수. (선택적) 모수 값을 입력하거나 기본값이 있으면 그대로 두십시오.

이러한 프롬프트 끄기. 스트림을 실행할 때 이러한 프롬프트가 표시되는 것을 원치 않으면 이 상자를 선택하십시오. 모수가 정의된 스트림 특성이나 세션 특성 대화 상자에서 **프롬프트 표시 여부** 확인 상자를 선택하여 이를 다시 표시되게 할 수 있습니다. 자세한 정보는 [41 페이지의 『스트림과 세션 모수 설정』](#)의 내용을 참조하십시오.

모수 유형의 값 제한조건 지정

데이터 모델링 스트림을 읽는 외부 애플리케이션에 스트림 배포 동안에 사용 가능한 모수의 값 제한조건을 만들 수 있습니다. 이 대화 상자를 사용하면 스트림을 실행하는 외부 사용자가 사용 가능한 값을 지정할 수 있습니다. 데이터 유형에 따라서 값 제한조건은 대화 상자에서 동적으로 변화합니다. 여기에 표시된 옵션은 유형 노드의 값에 사용 가능한 옵션과 동일합니다.

유형. 현재 선택된 측정 수준을 표시합니다. IBM SPSS Modeler에서 모수를 사용하려는 방법을 반영하여 이 값을 변경할 수 있습니다.

저장 공간. 알려진 경우 저장 유형을 표시합니다. 저장 유형은 사용자가 IBM SPSS Modeler에서 작업하기 위해 선택하는 측정 수준(연속, 명목 또는 플래그)에 영향을 받지 않습니다. 기본 모수 탭에서 저장 유형을 변경할 수 있습니다.

대화 상자의 하반부는 **유형** 필드에 선택된 측정 수준에 따라서 동적으로 변경됩니다.

연속 측정 수준

하한. 모수 값의 하한을 지정합니다.

상한. 모수 값의 상한을 지정합니다.

레이블. 범위 필드의 값에 레이블을 지정할 수 있습니다. 레이블 단추를 클릭하여 값 레이블을 지정하기 위한 별도의 대화 상자를 여십시오.

명목 측정 수준

값. 이 옵션을 사용하면 명목 필드로 사용할 모수의 값을 지정할 수 있습니다. 값은 IBM SPSS Modeler 스트림에서 강요되지 않지만 외부 배포 애플리케이션의 드롭 다운 목록에 사용합니다. 화살표와 삭제 단추를 사용하면 기존 값을 수정할 뿐만 아니라 값을 다시 정렬하고 삭제할 수 있습니다.

플래그 측정 수준

참. 조건이 충족될 때 모수의 플래그 값을 지정합니다.

거짓. 조건이 충족되지 않을 때 모수의 플래그 값을 지정합니다.

레이블. 플래그 필드의 값에 레이블을 지정할 수 있습니다.

스트림 배포 옵션

스트림 특성 대화 상자의 배포 탭에서는 모델 새로 고침 또는 자동화된 작업 스케줄링 용도로 IBM SPSS Collaboration and Deployment Services 내에 스트림을 배포하기 위한 옵션을 지정할 수 있습니다. 스트림을 배포할 수 있으려면 모든 스트림에 전용 스코어링 분기가 있어야 합니다. 자세한 정보는 [175 페이지의 『리포지토리 오브젝트 저장 및 배포』](#)를 참조하십시오.

스트림의 실행 반복

스트림 특성 대화 상자에서 실행 탭을 사용하면 현재 스트림에서 반복 작업을 자동화하도록 반복 조건을 설정할 수 있습니다.

이러한 조건을 설정한 후에는 이를 스크립팅에 소개로 사용하여 스크립트 창을 스트림의 기본 스크립팅으로 사용할 수 있습니다. 그런 다음 더 나은 스크립트를 작성하기 위한 기본으로 사용하기 위해 이를 수정할 수 있습니다. 자세한 정보는 [172 페이지의 『글로벌 함수』](#)의 내용을 참조하십시오.

스트림의 루핑 설정

1. 파일 메뉴에서 **스트림 특성**을 클릭하십시오. 또는 관리자 분할창의 스트림 탭에서 스트림을 선택하고 마우스 오른쪽 단추를 클릭한 다음 팝업 메뉴에서 **스트림 특성**을 클릭하십시오.
2. **실행** 탭을 클릭하십시오.
3. **루핑/조건부 실행** 실행 모드를 선택하십시오.
4. **루핑** 탭을 클릭하십시오.

또는 도구 메뉴에서 다음을 클릭하십시오.

스트림 특성 > 실행

또 다른 대안으로서 노드를 마우스 오른쪽 단추로 클릭하고 컨텍스트 메뉴에서 다음을 클릭하십시오.

루핑/조건부 실행 > 루핑 설정 편집

반복. 이 행 번호 값을 편집할 수 없지만 테이블 오른쪽의 단추를 사용하여 반복을 추가, 삭제 또는 위아래로 이동할 수 있습니다.

테이블 헤더. 이는 루프를 설정할 때 작성되는 반복 키와 반복 변수를 반영합니다.

스트림의 글로벌 값 보기

스트림 특성 대화 상자의 글로벌 탭을 사용하여 현재 스트림에 설정된 글로벌 값을 볼 수 있습니다. 글로벌 값은 선택한 필드의 평균, 합계 또는 표준 편차 등과 같은 통계를 판별하기 위해 전역값 설정 노드를 사용하여 작성됩니다.

전역값 설정 노드가 실행되면 이러한 값은 스트림 작업에서 다양한 용도로 사용 가능합니다. 자세한 정보는 [172 페이지의 『글로벌 함수』](#)의 내용을 참조하십시오.

스트림의 글로벌 값 보기

1. 파일 메뉴에서 **스트림 특성**을 클릭하십시오. 또는 관리자 분할창의 스트림 탭에서 스트림을 선택하고 마우스 오른쪽 단추를 클릭한 다음 팝업 메뉴에서 **스트림 특성**을 클릭하십시오.
2. **전역값** 탭을 클릭하십시오.

또는 도구 메뉴에서 다음을 클릭하십시오.

스트림 특성 > 전역값

전역값 사용가능. 사용 가능한 전역값이 이 테이블에 나열되어 있습니다. 여기에서는 글로벌 값을 편집할 수 없지만 테이블 오른쪽의 모두 선택 취소 단추를 사용하여 스트림의 모든 글로벌 값을 지울 수 있습니다.

스트림에서 노드 검색

노드 이름, 범주 및 식별자 등과 같은 여러 검색 기준을 지정하여 스트림에서 노드를 검색할 수 있습니다. 이 기능은 다수의 노드를 포함하는 복합 스트림에서 특히 유용할 수 있습니다.

스트림에서 노드 검색

1. 파일 메뉴에서 **스트림 특성**을 클릭하십시오. 또는 관리자 분할창의 스트림 탭에서 스트림을 선택하고 마우스 오른쪽 단추를 클릭한 다음 팝업 메뉴에서 **스트림 특성**을 클릭하십시오.
2. **검색** 탭을 클릭하십시오.

또는 도구 메뉴에서 다음을 클릭하십시오.

스트림 특성 > 검색

둘 이상의 옵션을 지정하여 검색을 제한할 수 있습니다. 그러나 노드 ID로 검색하면(**ID 동일** 필드 사용) 다른 옵션은 제외된다는 점은 예외입니다.

노드 레이블 포함. 특정 노드를 검색하려면 이 상자를 선택하고 노드 레이블의 모두나 일부를 입력하십시오. 검색은 대소문자를 구분하지 않고 다중 단어는 단일 텍스트 조각으로 처리됩니다.

노드 범주. 특정 유형의 노드를 검색하려면 이 상자를 선택하고 목록에서 범주를 클릭하십시오. **프로세스 노드는 노드 팔레트의 레코드 작업 또는 필드 작업 탭의 노드를 의미합니다. 모델 노드 적용은 모델 너겟을 가리킵니다.**

키워드 포함. 노드 대화 상자의 주석(Annotation) 탭에서 키워드 필드에 해당 텍스트가 있는 노드를 검색하려면 이 상자를 선택하고 하나 이상의 완전한 키워드를 입력하십시오. 입력하는 키워드 텍스트는 정확히 매치해야 합니다. 대안을 검색하려면 여러 키워드를 세미콜론으로 구분하십시오. 예를 들어, `proton;neutron`을 입력하면 이러한 키워드 중 하나가 있는 모든 노드를 찾습니다. 자세한 정보는 [50 페이지의 『주석\(Annotation\)』](#)의 내용을 참조하십시오.

주석(Annotation) 포함. 노드 대화 상자의 주석(Annotation) 탭에서 기본 텍스트 영역에 이 텍스트를 포함하는 노드를 검색하려면 이 상자를 선택하고 하나 이상의 단어를 입력하십시오. 검색은 대소문자를 구분하지 않고 다중 단어는 단일 텍스트 조각으로 처리됩니다. 자세한 정보는 [50 페이지의 『주석\(Annotation\)』](#)의 내용을 참조하십시오.

호출된 필드 생성. 이 상자를 선택하고 생성된 필드의 이름을 입력하십시오(예: `$C-Drug`). 이 옵션을 사용하여 특정 필드를 생성하는 모델링 노드를 검색할 수 있습니다. 정확히 매치해야 하는 단 하나의 필드 이름을 입력하십시오.

ID 동일. 해당 ID가 있는 특정 노드를 검색하려면 이 상자를 선택하고 노드 ID를 입력하십시오. 노드 ID는 노드가 작성될 때 시스템이 지정하고 스크립팅이나 자동화를 위해 노드를 참조할 때 사용할 수 있습니다. 정확히 매치해야 하는 단 하나의 노드 ID만을 입력하십시오. 자세한 정보는 [50 페이지의 『주석\(Annotation\)』](#)의 내용을 참조하십시오.

수퍼노드에서 검색. 이 상자는 기본적으로 선택되어 있습니다. 즉 수퍼노드의 내부와 외부 둘 모두의 노드에서 검색이 수행됨을 의미합니다. 스트림의 맨 위 수준에서 수퍼노드의 외부에 있는 노드에서만 검색을 수행하려면 이 상자를 지우십시오.

찾기. 원하는 모든 옵션을 지정한 후에는 이 단추를 클릭하여 검색을 시작하십시오.

지정된 옵션과 매치하는 노드가 대화 상자의 하부에 나열됩니다. 목록에서 노드를 선택하여 이를 스트림 캔버스에서 강조 표시하십시오.

스트림 이름 변경

스트림 특성 대화 상자에서 주석(Annotation) 탭을 사용하면 스트림의 설명 주석(Annotation)을 추가하고 스트림의 사용자 정의 이름을 작성할 수 있습니다. 이러한 옵션은 프로젝트 분할창에 추가된 스트림에 대한 보고서를 생성할 때 특히 유용합니다. 자세한 정보는 [50 페이지의 『주석\(Annotation\)』](#)의 내용을 참조하십시오.

스트림 설명

작성하는 각 스트림에 대해 IBM SPSS Modeler는 스트림의 내용에 대한 정보를 포함하는 스트림 설명을 생성합니다. 스트림의 수행 작업은 보려고 하지만 IBM SPSS Modeler가 설치되지 않은 경우(예를 들어, IBM SPSS Collaboration and Deployment Services를 통해 스트림에 액세스하는 경우) 유용할 수 있습니다.

스트림 설명은 여러 섹션으로 구성되어 HTML 문서의 양식으로 표시됩니다.

일반 스트림 정보

이 섹션에는 스트림 작성 시간과 마지막 저장 시간에 대한 세부사항과 함께 스트림 이름이 포함됩니다.

설명과 주석

이 섹션에는 다음이 포함됩니다.

- 스트림 주석(Annotation)([50 페이지의 『주석\(Annotation\)』](#) 참조)
- 특정 노드에 연결되지 않은 주석
- 스트림의 모델링 및 스코어링 분기 둘 모두에서 노드에 연결된 주석

스코어링 정보

이 섹션에는 스트림의 스코어링 분기와 관련하여 다양한 머리말 아래에 정보를 포함합니다.

- 주석 스코어링 분기에서 노드에 링크된 주석만을 포함합니다.
- 입력 입력 필드를 저장 유형과 함께 나열합니다(예: 문자열, 정수, 실수 등).
- 출력 모델링 노드에 의해 생성된 추가 필드를 포함하여 출력 필드를 저장 유형과 함께 나열합니다.
- 모수 스트림의 스코어링 분기와 관련되었고 모델이 스코어링될 때마다 보거나 편집할 수 있는 모수를 나열합니다. 이러한 모수는 스트림 특성 대화 상자의 **배포 탭**에서 **스코어링 모수** 단추를 클릭할 때 식별됩니다.
- 모델 노드 모델 이름과 유형을 표시합니다(예: 신경망, C&R 트리 등). 이는 스트림 특성 대화 상자의 **배포 탭**에서 **모델 노드** 필드에 대해 선택된 모델 너깃입니다.
- 모델 세부사항 식별된 모델 너깃의 세부사항을 이전 머리말 아래에 표시합니다. 가능한 경우 모델의 예측자 중요도와 평가 차트가 포함됩니다.

모델링 정보

스트림의 모델링 분기와 관련된 정보를 포함합니다.

- 주석 모델링 분기에서 노드에 연결된 주석 또는 주석(Annotation)을 나열합니다.
- 입력 입력 필드를 모델링 분기에서의 역할과 함께 나열합니다(예를 들어, 입력, 목표, 분할 등의 필드 역할 값의 양식으로).
- 모수 스트림의 모델링 분기와 관련되었고 모델이 업데이트될 때마다 보거나 편집할 수 있는 모수를 나열합니다. 이러한 모수는 스트림 특성 대화 상자의 **배포 탭**에서 **모델 작성 모수** 단추를 클릭할 때 식별됩니다.
- 모델링 노드 모델을 생성하거나 업데이트하는 데 사용하는 모델링 노드의 이름과 유형을 표시합니다.

스트림 설명 미리보기

스트림 특성 대화 상자에서 옵션을 클릭하여 스트림 설명의 내용을 웹 브라우저에서 볼 수 있습니다. 설명의 내용은 대화 상자의 배포 탭에 지정하는 옵션에 따라 다릅니다. 자세한 정보는 [186 페이지의 『스트림 배포 옵션』](#)의 내용을 참조하십시오.

스트림 설명 보기

1. 기본 IBM SPSS Modeler 메뉴에서 다음을 클릭하십시오.

도구 > 스트림 특성 > 배포

2. 배포 유형, 전용 스코어링 노드 및 모든 스코어링 모수를 설정하십시오.

3. 배포 유형이 모델 새로 고침이면 선택적으로 다음을 선택할 수 있습니다.

- 모델링 노드 및 모든 모델 작성 모수
- 스트림의 스코어링 분기 위의 모델 너깃

4. 스트림 설명 미리보기 단추를 클릭하십시오.

스트림 설명 내보내기

스트림 설명의 내용을 HTML 파일로 내보낼 수 있습니다.

스트림 설명 내보내기

1. 기본 메뉴에서 다음을 클릭하십시오.

파일 > 스트림 설명 내보내기

2. HTML 파일의 이름을 입력하고 저장을 클릭하십시오.

스트림 실행

필수 스트림 옵션을 지정했고 필수 노드를 연결했으면 스트림의 노드를 통해 데이터를 실행하여 스트림을 실행할 수 있습니다. IBM SPSS Modeler에서 스트림을 실행하는 몇 가지 방법이 있습니다. 다음을 수행할 수 있습니다.

- 도구 메뉴에서 **실행**을 클릭하십시오.
- 도구 모음에서 **실행...** 단추 중 하나를 클릭하십시오. 이러한 단추를 사용하여 전체 스트림을 실행하거나 선택된 터미널 노드만 실행할 수 있습니다. 자세한 정보는 15 페이지의 [『IBM SPSS Modeler 도구 모음』](#)의 내용을 참조하십시오.
- 터미널 노드를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 **실행**을 클릭하여 단일 데이터 스트림을 실행하십시오.
- 비터미널 노드를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 **여기에서 실행**을 클릭하여 데이터 스트림의 일부를 실행하십시오. 이렇게 하면 선택된 노드 뒤의 해당 작업만 수행됩니다.

진행 중인 스트림 실행을 정지하려면 도구 모음에서 빨간색 중지 단추를 클릭하거나, 도구 메뉴에서 **실행 중지**를 선택할 수 있습니다. 중지 단추를 한 번 클릭하면 Modeler가 Modeler Server 실행을 중지하도록 지시합니다. 어떤 경우에는 실행이 즉시 중지되지만, 다른 경우에는 전체 실행을 중지하기 전에 현재 단계를 완료해야 합니다. 따라서 시간이 달라집니다. 단추를 두 번 클릭하면 서버 연결이 삭제되고 새 연결이 설정됩니다. 대부분의 경우 모든 서버 프로세스가 닫힙니다(다소 시간이 걸릴 수 있음). 그러나 어떤 경우에는 일부 서버 프로세스가 중지되지 않습니다.

스트림을 실행하는 데 3초가 넘으면, 실행 피드백 대화 상자가 표시되어 진행률을 나타냅니다.

일부 노드에는 스트림 실행에 대한 추가 정보를 제공하는 추가 표시가 있습니다. 이러한 표시를 보려면 대화 상자에서 해당 행을 선택하십시오. 첫째 행이 자동으로 선택됩니다.

모델에 대한 작업

스트림이 모델링 노드(즉, 노드 팔레트의 모델링 또는 데이터베이스 모델링 탭의 노드)를 포함하는 경우 스트림이 실행될 때 모델 너깃이 작성됩니다. 모델 너깃은 모델 컨테이너 즉, 소스 데이터에 대해 예측을 생성할 수 있게 하고 예측 분석의 심장에 놓여 있는 규칙 세트, 수식 또는 등식입니다.



그림 12. 모델 너깃

모델링 노드를 실행할 때 해당 모델 너깃이 스트림 캔버스에 배치되는데, 여기서 황금색 마름모꼴 아이콘으로 표시되기 때문에 이름이 "너깃"입니다. 너깃을 열고 내용을 찾아 모델에 대한 세부사항을 볼 수 있습니다. 예측을 보려면 하나 이상의 터미널 노드에 접속하여 예측을 읽을 수 있는 양식으로 제공하는 출력을 실행하십시오.

일반적인 모델링 스트림은 2개의 분기로 구성됩니다. **모델링 분기**는 모델링 노드를 포함하며, 모델링 노드 앞에 소스 및 처리 노드가 있습니다. **스코어링 분기**는 모델링 노드를 실행할 때 작성되며, 예측을 보는 데 사용하는 모델 너깃과 터미널 노드를 포함합니다.

자세한 정보는 [IBM SPSS Modeler 모델링 노드 안내서](#)를 참조하십시오.

노드와 스트림에 주석과 주석(Annotation) 추가

사용자의 조직에서 다른 사람에게 스트림을 설명해야 할 수도 있습니다. 이를 위해서 스트림, 노드 및 모델 너깃에 설명 주석을 첨부할 수 있습니다.

그러면 다른 사람은 화면에서 이러한 주석을 보거나 주석을 포함하는 스트림의 이미지를 출력할 수 있습니다.

스트림이나 수퍼노드의 모든 주석을 나열하고, 목록에서 주석의 순서를 변경하고, 주석 텍스트를 편집하고, 주석의 전경 또는 배경 색상을 변경할 수 있습니다. 자세한 정보는 49 페이지의 [『스트림 주석 나열』](#)의 내용을 참조하십시오.

스트림 특성 대화 상자, 노드 대화 상자 또는 모델 너깃 창의 주석(Annotation) 탭을 사용하여 노트를 텍스트 주석(Annotation)의 양식으로 스트림, 노드 및 너깃에 추가할 수도 있습니다. 이러한 노트는 주석(Annotation) 탭이 열려 있는 경우에만 표시되지만 스트림 주석(Annotation)은 화면 주석으로도 표시 가능합니다. 자세한 정보는 50 페이지의 [『주석\(Annotation\)』](#)의 내용을 참조하십시오.

주석

주석은 텍스트를 원하는 양만큼 입력할 수 있고 원하는 수만큼의 주석을 추가할 수 있는 텍스트 상자의 양식으로 되어 있습니다. 주석은 독립적이거나(스트림 오브젝트에 첨부되지 않음) 스트림에서 하나 이상의 노드나 모델 너깃에 연결될 수 있습니다. 독립적인 주석은 일반적으로 스트림의 전체 목적을 설명하는 데 사용합니다. 연결된 주석은 이들이 첨부된 노드나 너깃을 설명합니다. 노드와 너깃에는 둘 이상의 주석이 첨부될 수 있으며 스트림에는 독립적인 주석을 원하는 수만큼 가질 수 있습니다.

참고: 스트림 주석(Annotation)을 화면 주석으로 표시할 수도 있지만 이들은 노드나 너깃에는 첨부할 수 없습니다. 자세한 정보는 [50 페이지의 『주석\(Annotation\)을 주석으로 변환』](#)의 내용을 참조하십시오.

텍스트 상자의 모양은 다음 테이블에 표시된 대로 주석의 현재 모드(또는 주석으로 표시된 주석(Annotation))를 나타내기 위해 변경됩니다.

표 3. 주석과 주석(Annotation) 텍스트 상자 모드				
주석 텍스트 상자	주석(Annotation) 텍스트 상자	모드	표시	획득 방법..
		편집	주석이 편집을 위해 열려 있습니다.	새 주석 또는 주석(Annotation)을 작성하거나 기존의 것을 두 번 클릭합니다.
		마지막 선택	주석은 이동하거나, 크기를 조정하거나 삭제할 수 있습니다.	기존 주석이나 주석(Annotation)을 편집하거나 한 번 클릭한 후에 스트림 배경을 클릭.
		보기	편집이 완료되었습니다.	편집 후에 또 다른 노드, 주석 또는 주석(Annotation) 클릭.

새로운 독립된 주석을 작성할 때 처음에는 스트림 캔버스의 왼쪽 상단 구석에 표시됩니다.

노드나 너깃에 주석을 첨부하는 경우에는 주석은 처음에는 주석이 첨부된 스트림 오브젝트 위에 표시됩니다.

텍스트 상자는 텍스트를 입력할 수 있다는 점을 표시하기 위해 흰색으로 표시되어 있습니다. 텍스트를 입력한 후에는 텍스트 상자 밖을 클릭하십시오. 텍스트 입력이 완료되었음을 보여주기 위해 주석 배경이 노란색으로 표시됩니다. 주석이 선택된 상태로 남아 있어서 이를 이동, 크기 조정 또는 삭제할 수 있습니다.

다시 클릭하면 경계는 편집이 완료되었음을 보여주기 위해 실선으로 변경됩니다.

주석을 두 번 클릭하면 텍스트 상자가 편집 모드로 변경됩니다. 배경이 흰색으로 변경되고 주석 텍스트를 편집할 수 있습니다.

수퍼노드에 주석을 첨부할 수도 있습니다.

주석 관련 작업

주석에 대해 다수의 조작을 수행할 수 있습니다. 다음을 수행할 수 있습니다.

- 독립된 주석 추가
- 노드 또는 너깃에 주석 첨부
- 주석 편집
- 주석 크기 조정
- 주석 이동
- 주석 연결 해제
- 주석 삭제
- 스트림에 대한 주석 모두 표시 또는 숨기기

독립된 주석 추가 방법

1. 스트림에서 아무것도 선택되어 있지 않은지 확인하십시오.

2. 다음 중 하나를 수행하십시오.

- 기본 메뉴에서 다음을 클릭하십시오.

삽입 > 새 주석

- 스트림 배경을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 새 주석을 클릭하십시오.

- 도구 모음에서 새 주석 단추를 클릭하십시오.

3. 주석 텍스트를 입력하거나 클립보드에서 텍스트를 붙여넣으십시오.

4. 스트림에서 노드를 클릭하여 주석을 저장하십시오.

노드 또는 너깃에 주석 첨부 방법

1. 스트림 캔버스에서 하나 이상의 노드 또는 너깃을 선택하십시오.

2. 다음 중 하나를 수행하십시오.

- 기본 메뉴에서 다음을 클릭하십시오.

삽입 > 새 주석

- 스트림 배경을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 새 주석을 클릭하십시오.

- 도구 모음에서 새 주석 단추를 클릭하십시오.

3. 주석 텍스트를 입력하십시오.

4. 스트림에서 다른 노드를 클릭하여 주석을 저장하십시오.

또는 다음을 수행할 수 있습니다.

5. 독립된 주석을 삽입하십시오(이전 절 참조).

6. 다음 중 하나를 수행하십시오.

- 주석을 선택하고 F2를 누른 후 노드 또는 너깃을 선택하십시오.

- 노드 또는 너깃을 선택하고 F2를 누른 후 주석을 선택하십시오.

- (3단추 마우스만) 마우스 포인터를 주석 위로 이동하고 가운데 단추를 누른 상태에서 노드나 너깃 위로 마우스 포인터를 끈 후 마우스 단추를 놓으십시오.

추가 노드 또는 너깃에 주석 첨부 방법

주석이 노드나 너깃에 이미 첨부되어 있거나 현재 스트림 수준에 있는데 추가 노드나 너깃에 첨부하려면 다음 중 하나를 수행하십시오.

- 주석을 선택하고 F2를 누른 후 노드 또는 너깃을 선택하십시오.

- 노드 또는 너깃을 선택하고 F2를 누른 후 주석을 선택하십시오.

- (3단추 마우스만) 마우스 포인터를 주석 위로 이동하고 가운데 단추를 누른 상태에서 노드나 너깃 위로 마우스 포인터를 끈 후 마우스 단추를 놓으십시오.

기존 주석 편집 방법

1. 다음 중 하나를 수행하십시오.

- 주석 텍스트 상자를 두 번 클릭하십시오.

- 텍스트 상자를 선택하고 Enter를 누르십시오.

- 텍스트 상자를 마우스 오른쪽 단추로 클릭하여 메뉴를 표시한 후 편집을 클릭하십시오.

2. 주석 텍스트를 편집하십시오. 편집 시 표준 Windows 단축키를 사용할 수 있습니다(예: 복사의 경우 Ctrl+C). 편집 중에 기타 옵션이 주석 팝업 메뉴에 나열됩니다.

3. 텍스트 상자 바깥쪽을 한 번 클릭하여 크기 조정 제어를 표시한 후 다시 주석을 완료하십시오.

주석 텍스트 상자 크기 조정 방법

1. 크기 조정 제어를 표시할 주석을 선택하십시오.

2. 제어를 클릭하여 끌어서 상자 크기를 조정하십시오.

3. 텍스트 상자 바깥쪽을 클릭하여 변경사항을 저장하십시오.

기존 주석 이동 방법

첨부된 오브젝트(있는 경우)는 이동하지 않고 주석을 이동하려면 다음 중 하나를 수행하십시오.

- 마우스 포인터를 주석 위로 이동하고 마우스 왼쪽 단추를 누른 상태에서 주석을 새 위치로 끄십시오.
- 주석을 선택하고, Alt 키를 누른 상태에서 화살표 키를 사용하여 주석을 이동하십시오.

주석이 첨부된 노드나 너깃과 함께 주석을 이동하려면 다음을 수행하십시오.

1. 이동할 오브젝트를 모두 선택하십시오.

2. 다음 중 하나를 수행하십시오.

- 마우스 포인터를 오브젝트 중 하나로 이동하고 마우스 왼쪽 단추를 누른 상태에서 오브젝트를 새 위치로 끄십시오.

• 오브젝트 중 하나를 선택하고, Alt 키를 누른 상태에서 화살표 키를 사용하여 오브젝트를 이동하십시오.

노드 또는 너깃에서 주석 연결 해제 방법

1. 연결을 해제할 하나 이상의 주석을 선택하십시오.

2. 다음 중 하나를 수행하십시오.

- F3을 누르십시오.

• 선택된 주석을 마우스 오른쪽 단추로 클릭하고 해당 메뉴에서 연결 해제를 클릭하십시오.

주석 삭제 방법

1. 삭제할 하나 이상의 주석을 선택하십시오.

2. 다음 중 하나를 수행하십시오.

- Delete 키를 누르십시오.

• 선택된 주석을 마우스 오른쪽 단추로 클릭하고 해당 메뉴에서 삭제를 클릭하십시오.

주석이 노드나 너깃에 첨부된 경우 연결 행도 삭제됩니다.

주석이 원래 독립된 주석으로 변환된 스트림이나 수퍼노드였으면 캔버스에서 주석이 삭제되지만 텍스트는 스트림 또는 수퍼노드의 주석(Annotation) 탭에 유지됩니다.

스트림에 대한 주석 표시 또는 숨기기 방법

1. 다음 중 하나를 수행하십시오.

- 기본 메뉴에서 다음을 클릭하십시오.

보기 > 주석

- 도구 모음에서 주석 표시/숨기기 단추를 클릭하십시오.

스트림 주석 나열

특정 스트림 또는 수퍼노드에 대해 작성된 모든 주석 목록을 볼 수 있습니다.

이 목록에서 다음을 수행할 수 있습니다.

- 주석 순서 변경
- 주석 텍스트 편집
- 주석의 전경 또는 배경색 변경

주석 나열

스트림에 대해 작성된 주석을 나열하려면 다음 중 하나를 수행하십시오.

- 기본 메뉴에서 다음을 클릭하십시오.

도구 > 스트림 특성 > 주석

- 관리자 분할창에서 스트림을 마우스 오른쪽 단추로 클릭하고 스트림 특성, 주석을 차례로 클릭하십시오.

- 캔버스에서 스트림 배경을 마우스 오른쪽 단추로 클릭하고 **스트림 특성**, 주석을 차례로 클릭하십시오.

텍스트. 주석 텍스트입니다. 필드를 편집 가능한 텍스트 상자로 변경하려면 텍스트를 두 번 클릭하십시오.

링크. 주석을 첨부할 노드 이름입니다. 이 필드가 비어 있으면 스트림에 주석이 적용됩니다.

위치지정 단추. 선택된 주석을 목록에서 위, 아래로 이동합니다.

주석 색상. 주석의 전경 또는 배경색을 변경하려면 주석을 선택하고 **사용자 정의 색상** 확인 상자를 선택한 후 **배경** 또는 **전경** 목록(또는 둘 다)에서 색상을 선택하십시오. **적용**을 클릭한 후 스트림 배경을 클릭하여 변경 효과를 확인하십시오. **확인**을 클릭하여 변경사항을 저장하십시오.

주석(Annotation)을 주석으로 변환

스트림이나 수퍼노드에 작성된 주석(Annotation)을 주석으로 변환할 수 있습니다.

스트림의 경우, 주석(Annotation)은 스트림 캔버스에서 독립된 주석(즉, 노드에 첨부되지 않음)으로 변환됩니다.

수퍼노드 주석(Annotation)이 주석으로 변환될 때 주석은 스트림 캔버스에서 수퍼노드에 첨부되지 않지만 수퍼노드를 확대하면 표시됩니다.

스트림 주석(Annotation)을 주석으로 변환하는 방법

1. 도구 메뉴에서 **스트림 특성을** 클릭하십시오. (또는 관리자 분할창에서 스트림을 마우스 오른쪽 단추로 클릭하고 **스트림 특성을** 클릭하십시오.)
2. **주석(Annotation)** 탭을 클릭하십시오.
3. **주석(Annotation)**을 주석으로 표시 확인 상자를 선택하십시오.
4. **확인**을 클릭하십시오.

수퍼노드 주석(Annotation)을 주석으로 변환하는 방법

1. 캔버스에서 수퍼노드 아이콘을 두 번 클릭하십시오.
2. **주석(Annotation)** 탭을 클릭하십시오.
3. **주석(Annotation)**을 주석으로 표시 확인 상자를 선택하십시오.
4. **확인**을 클릭하십시오.

주석(Annotation)

노드, 스트림 및 모델은 여러 가지 방법으로 주석(Annotation)을 달 수 있습니다. 설명 주석(Annotation)을 추가하고 사용자 정의 이름을 추가할 수 있습니다. 이러한 옵션은 프로젝트 분할창에 추가된 스트림에 대한 보고서를 생성할 때 특히 유용합니다. 노드와 모델 너깃의 경우 스트림 캔버스에서 유사한 노드 간에 구별할 수 있도록 도구 텁 텍스트를 추가할 수도 있습니다.

주석 추가

노드나 모델 너깃을 편집하면 다양한 주석(Annotation) 옵션을 설정하는 데 사용하는 주석(Annotation) 탭을 포함하는 탭이 있는 대화 상자가 열립니다. 주석(Annotation) 탭을 직접 열 수도 있습니다.

1. 노드나 너깃에 주석(Annotation)을 달려면 스트림 캔버스에서 노드나 너깃을 마우스 오른쪽 단추로 클릭하고 **이름 변경 및 주석달기**를 클릭하십시오. 편집 대화 상자는 주석(Annotation) 탭이 표시된 상태로 열립니다.
2. 스트림에 주석(Annotation)을 달려면 도구 메뉴에서 **스트림 특성을** 클릭하십시오. (또는 관리자 분할창에서 **스트림**을 마우스 오른쪽 단추로 클릭하고 **스트림 특성을** 클릭하십시오.) 주석(Annotation) 탭을 클릭하십시오.

이름 사용자 정의를 선택하여 자동 생성된 이름을 조정하거나 스트림 캔버스에 표시되는 노드의 고유 이름을 작성하십시오.

도구 텁 텍스트. (노드 및 모델 너깃에만 해당) 스트림 캔버스에서 도구 텁으로 사용하는 텍스트를 입력합니다. 이는 특히 대량의 유사한 노드에 대해 작업할 때 유용합니다.

키워드 프로젝트 보고서에 사용하고, 스트림에서 노드를 검색할 때 또는 리포지토리에 저장된 오브젝트를 추적할 때 사용할 키워드를 지정하십시오(175 페이지의 『IBM SPSS Collaboration and Deployment Services Repository 정보』 참조). 여러 개의 키워드는 세미콜론으로 구분할 수 있습니다(예: `income`; `crop type`;

claim value). 각 키워드의 시작부와 끝부분에 있는 공백은 잘립니다(예: income ; crop type은 income;crop type과 같은 결과임). (그러나, 키워드 내의 공백은 잘리지 않습니다. 예를 들어, 하나의 공백이 있는 crop type과 두 개의 공백이 있는 crop type은 서로 다릅니다.)

기본 텍스트 영역은 노드의 작업이나 노드에서 결정된 의사결정과 관계없이 긴 주석(Annotation)을 입력하는데 사용할 수 있습니다. 예를 들어, 스트림을 공유하고 재사용할 때는 필터 노드를 사용하여 여러 공백이 있는 필드를 삭제하는 등 의사결정에 대해 기록을 남겨두는 것이 유용합니다. 노드에 주석을 작성하면 이 정보가 노드와 함께 저장됩니다. 이러한 주석(Annotation)을 프로젝트 분할창에서 작성된 프로젝트 보고서에 포함하도록 선택할 수도 있습니다. 자세한 정보는 197 페이지의 『프로젝트에 대한 소개』의 내용을 참조하십시오.

주석(Annotation)을 주석으로 표시 (스트림과 수퍼노드 주석(Annotation)에만 해당) 주석(Annotation)을 스트림 캔버스에 표시되는 독립된 주석으로 변환하려면 이 상자를 선택하십시오. 자세한 정보는 46 페이지의 『노드와 스트림에 주석과 주석(Annotation) 추가』의 내용을 참조하십시오.

ID 스크립팅이나 자동화를 위해 노드를 참조하는 데 사용할 수 있는 고유한 ID를 표시합니다. 이 값은 노드가 작성될 때 자동으로 생성되고 변경되지 않습니다. 문자 "0"과의 혼동을 피하기 위해 0값은 노드 ID에는 사용하지 않음을 유의하십시오. 오른쪽의 복사 단추를 사용하여 ID를 복사하여 스크립트 또는 다른 필요한 곳에 붙여넣으십시오.

데이터 스트림 저장

스트림을 작성했으면 나중에 사용하기 위해 저장할 수 있습니다.

스트림 저장

1. 파일 메뉴에서 **스트림 저장 또는 다른 이름으로 스트림 저장**을 클릭하십시오.
2. 저장 대화 상자에서 스트림 파일을 저장할 폴더를 찾아보십시오.
3. 파일 이름 텍스트 상자에 스트림 이름을 입력하십시오.
4. 저장된 스트림을 현재 프로젝트에 추가하려면 **프로젝트에 추가**를 선택하십시오.

저장을 클릭하면 지정된 디렉토리에 *.str 확장자로 스트림이 저장됩니다.

자동 백업 파일. 스트림이 저장될 때마다 이전에 저장된 파일 버전이 백업으로 자동으로 보존되는데, 파일 이름에 하이픈이 추가됩니다(예: mystream.str-). 백업된 버전을 복원하려면 하이픈을 삭제하고 파일을 다시 열면 됩니다.

상태 저장

스트림에 추가로, 현재 표시된 스트림 다이어그램과 작성한 모든 모델 너깃(관리자 분할창의 모델에 나열됨)을 포함하는 상태를 저장할 수 있습니다.

상태 저장

1. 파일 메뉴에서 다음을 클릭하십시오.
상태 > 상태 저장 또는 다른 이름으로 상태 저장
2. 저장 대화 상자에서 상태 파일을 저장하려는 폴더로 이동하십시오.

저장을 클릭하면 확장자가 *.cst인 상태가 지정된 디렉토리에 저장됩니다.

노드 저장

스트림 캔버스에서 노드를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 **노드 저장**을 클릭하여 개별 노드를 저장할 수도 있습니다. 파일 확장자 *.nod를 사용하십시오.

여러 스트림 오브젝트 저장

스트림, 프로젝트 또는 모델 너깃과 같이 저장되지 않은 여러 오브젝트가 있는 상태로 IBM SPSS Modeler를 종료하면 소프트웨어를 완전히 닫기 전에 저장하라는 프롬프트가 표시됩니다. 항목을 저장하기로 선택하면 대화 상자가 각 오브젝트를 저장하기 위한 옵션을 표시합니다.

1. 단순히 저장하려는 오브젝트의 확인 상자를 선택하십시오.

2. 확인을 클릭하여 각 오브젝트를 필수 위치에 저장하십시오.

그런 다음 각 오브젝트의 표준 저장 대화 상자가 표시됩니다. 저장을 완료하고 나면 애플리케이션이 닫힙니다.

출력 저장

IBM SPSS Modeler 출력 노드에서 생성된 테이블, 그래프 및 보고서는 출력 오브젝트(*.cou) 형식으로 저장할 수 있습니다.

1. 저장하려는 출력을 볼 때 출력 창 메뉴에서 다음을 클릭하십시오.

파일 > 저장

2. 출력 파일의 이름과 위치를 지정하십시오.

3. 또는 저장 대화 상자에서 파일을 프로젝트에 추가를 선택하여 파일을 현재 프로젝트에 포함시키십시오. 자세한 정보는 197 페이지의 [『프로젝트에 대한 소개』](#)의 내용을 참조하십시오.

또는 관리자 분할창에 나열된 출력 오브젝트를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 저장을 선택하십시오.

정보 암호화 및 복호화

스트림, 노드, 프로젝트, 출력 파일 또는 모델 너깃을 저장할 때 암호화하여 무단 사용을 방지할 수 있습니다. 이를 수행하려면 저장 시 추가 옵션을 선택하고 저장되는 항목에 비밀번호를 추가하십시오. 이 암호화는 사용자가 저장하는 항목에 대해 설정할 수 있으며 항목에 추가 보안을 추가합니다. IBM SPSS Modeler와 IBM SPSS Modeler Server 간에 파일을 전달하는 경우 사용하는 SSL 암호화와 같지 않습니다.

암호화된 항목을 열려고 하면 비밀번호를 입력하도록 프롬프트가 표시됩니다. 올바른 비밀번호를 입력한 후에 항목이 복호화되어 평소처럼 열립니다.

항목 암호화 방법

1. 저장 대화 상자에서 암호화할 항목에 대해 옵션을 클릭하십시오. 암호화 옵션 대화 상자가 열립니다.

2. 이 파일 암호화를 선택하십시오.

3. 선택적으로, 추가 보안을 위해 비밀번호 마스크를 선택하십시오. 이렇게 하면 입력하는 것이 일련의 점으로 표시됩니다.

4. 비밀번호를 입력하십시오. 경고: 비밀번호를 잊어버리면 파일 또는 모델을 열 수 없습니다.

5. 비밀번호 마스크를 선택했으면 비밀번호를 다시 입력하여 올바로 입력했는지 확인하십시오.

6. 확인을 클릭하여 저장 대화 상자로 돌아가십시오.

참고: 암호화 보호 항목 사본을 저장하는 경우, 새 항목은 암호화 옵션 대화 상자에서 설정을 변경하지 않는 한 원래 비밀번호를 사용하여 암호화된 형식으로 자동으로 저장됩니다.

파일 로드

IBM SPSS Modeler에 저장된 다수의 오브젝트를 재로드할 수 있습니다.

- 스트림(.str)
- 상태(.cst)
- 모델(.gm)
- 모델 팔레트(.gen)
- 노드(.nod)
- 출력(.cou)
- 프로젝트(.cpj)

새 파일 열기

파일 메뉴에서 직접 스트림을 로드할 수 있습니다.

- 파일 메뉴에서 **스트림 열기**를 클릭하십시오.

파일 메뉴에서 사용 가능한 하위 메뉴 항목을 사용하여 기타 모든 파일 유형을 열 수 있습니다. 예를 들어, 모델을 로드하려면 파일 메뉴에서 다음을 클릭하십시오.

모델 > 모델 열기 또는 모델 팔레트 로드

최근에 사용한 파일 열기

최근에 사용한 파일을 빨리 로드하기 위해 파일 메뉴 맨 아래에 있는 옵션을 사용할 수 있습니다.

최근에 사용한 파일 목록을 펼치려면 **최근 스트림**, **최근 프로젝트** 또는 **최근 상태**를 선택하십시오.

데이터 스트림 맵핑

맵핑 도구를 사용하면 새 데이터 소스를 기존 스트림에 연결할 수 있습니다. 맵핑 도구는 연결을 설정할 뿐만 아니라 새 소스의 필드가 기존 스트림의 필드를 대체하는 방법을 지정하는 데에도 도움을 줍니다. 새 데이터 소스의 전체 데이터를 다시 작성하는 대신에, 간단하게 기존 스트림에 연결할 수 있습니다.

데이터 맵핑 도구를 사용하면 두 개의 스트림 단편을 함께 결합하고 모든 본질적 필드 이름이 적절하게 매치하는지 확인할 수 있습니다. 기본적으로 데이터 맵핑 결과로 단순히 새 필터 노드가 작성되고 이는 이름을 바꿔서 적합한 필드와 매치시킵니다.

데이터를 맵핑하는 데에는 두 개의 동등한 방법이 있습니다.

대체 노드 선택. 이 방법은 대체할 노드로 시작합니다. 먼저 대체할 노드를 마우스 오른쪽 단추로 클릭한 다음 팝업 메뉴에서 **데이터 맵핑 > 대체 노드 선택** 옵션을 사용하여 이를 대체할 노드를 선택하십시오.

맵핑 대상. 이 방법은 스트림에 도입될 노드로 시작합니다. 먼저 도입할 노드를 마우스 오른쪽 단추로 클릭한 다음 팝업 메뉴에서 **데이터 맵핑 > 맵핑 대상** 옵션을 사용하여 결합할 노드를 선택하십시오. 이 방법은 특히 터미널 노드에 맵핑할 때 유용합니다. 참고: 병합 또는 붙여쓰기 노드에는 맵핑할 수 없습니다. 대신 간단하게 스트림을 정상적인 방법으로 병합 노드에 연결해야 합니다.

데이터 맵핑은 스트림 작성에 엄격하게 통합됩니다. 이미 연결이 있는 노드에 연결하려고 시도하면 해당 노드에 대한 연결 또는 맵핑을 대체하는 옵션이 제공됩니다.

데이터를 템플리트에 맵핑

템플리트 스트림의 데이터 소스를 새 소스 노드로 대체하여 사용자 고유의 데이터를 IBM SPSS Modeler로 가져오려면 데이터 맵핑 팝업 메뉴에서 **대체 노드 선택** 옵션을 사용해야 합니다. 이 옵션은 병합, 통합 및 모든 터미널 노드를 제외한 모든 노드에서 사용 가능합니다. 데이터 맵핑 도구를 사용하여 이 조치를 수행하면 기존 스트림 작업과 새 데이터 소스가 제대로 매치하는지 확인하는 데 도움이 됩니다. 다음 단계는 데이터 맵핑 프로세스의 개요를 제공합니다.

1단계: 원본 소스 노드에서 본질적 필드를 지정하십시오. 스트림 작업이 제대로 실행되려면 본질적 필드가 지정되어야 합니다. 자세한 정보는 54 페이지의 **『본질적 필드 지정』**의 내용을 참조하십시오.

2단계: 새 데이터 소스를 스트림 캔버스에 추가하십시오. 소스 노드 중 하나를 사용하여 새 대체 데이터를 가져오십시오.

3단계: 템플리트 소스 노드를 대체하십시오. 템플리트 소스 노드의 팝업 메뉴에서 데이터 맵핑 옵션을 사용하여 **대체 노드 선택**을 클릭한 다음 대체 데이터의 소스 노드를 선택하십시오.

4단계: 맵핑된 필드를 확인하십시오. 대화 상자가 열리면 소프트웨어가 대체 데이터 소스에서 스트림으로 제대로 맵핑되는지 확인하십시오. 맵핑되지 않은 모든 본질적 필드가 빨간색으로 표시됩니다. 이러한 필드는 스트림 작업에 사용하고 다운스트림 작업이 제대로 작동하기 위해서는 새 데이터 소스의 유사한 필드로 대체되어야 합니다. 자세한 정보는 54 페이지의 **『맵핑된 필드 탐색』**의 내용을 참조하십시오.

대화 상자를 사용하여 모든 본질적 필드가 제대로 맵핑되었는지 확인한 후에는 이전 데이터 소스가 연결이 해제되고 새 데이터 소스가 맵이라는 필터 노드를 사용하여 스트림에 연결됩니다. 이 필터 노드는 스트림에서 필드의 실제 맵핑을 지시합니다. 맵핑 해제 필터 노드는 또한 스트림 캔버스에 포함됩니다. 맵핑 해제 필터 노드는 필드 이름 맵핑을 스트림에 추가하여 이를 역전시키는 데 사용할 수 있습니다. 이는 맵핑된 필드를 실행 취소하지만 필드와 오버레이를 다시 선택하기 위해 모든 다운스트림 터미널 노드를 편집해야 함을 유의하십시오.

스트림 사이에 맵핑

노드 연결과 유사하게, 이 데이터 맵핑 방법에서는 본질적 필드를 미리 설정할 필요가 없습니다. 이 방법을 사용하면 데이터 맵핑 팝업 메뉴에서 **맵핑 대상**을 사용하여 한 스트림에서 다른 스트림으로 간단하게 연결할 수 있습니다. 이 데이터 맵핑 유형은 터미널 노드에 맵핑하고 스트림 간에 복사하고 붙여넣는 데 유용합니다. 참고: **맵핑 대상** 옵션을 사용하면 병합, 붙여쓰기 및 모든 소스 노드 유형으로 맵핑할 수 없습니다.

스트림 간에 데이터 맵핑

1. 새 스트림에 연결하기 위해 사용하려는 노드를 마우스 오른쪽 단추로 클릭하십시오.
2. 메뉴에서 다음을 클릭하십시오.

데이터 맵핑 > 맵핑 대상

3. 목표 스트림에서 커서를 사용하여 대상 노드를 선택하십시오.
4. 대화 상자가 열리면 해당 필드가 적절하게 매치하는지 확인하고 **확인**을 클릭하십시오.

본질적 필드 지정

기존 스트림에 맵핑할 때 본질적 필드는 일반적으로 스트림 작성자가 지정합니다. 이러한 본질적 필드는 특정 필드가 다운스트림 작업에 사용하는지 여부를 표시합니다. 예를 들어, 기존 스트림은 *Churn*이라는 필드를 사용하는 모델을 작성할 수 있습니다. 이 스트림에서 *Churn*은 이 필드 없이는 모델을 작성할 수 없으므로 본질적 필드입니다. 마찬가지로, 파생 노드처럼 조작 노드에 사용하는 필드는 새 필드를 파생하는 데 필요합니다. 이러한 필드를 본질적인 것으로 명시적으로 설정하면 새 소스 노드에서 적합한 필드가 여기로 맵핑됩니다. 필수 필드가 맵핑되지 않으면 오류 메시지를 받게 됩니다. 특정 조작 또는 출력 노드가 불필요하다고 결정하는 경우에는 스트림에서 노드를 삭제하고 본질적 필드 목록에서 적합한 필드를 제거할 수 있습니다.

본질적 필드 설정

1. 대체될 템플리트 스트림의 소스 노드를 마우스 오른쪽 단추로 클릭하십시오.
2. 메뉴에서 다음을 클릭하십시오.

데이터 맵핑 > 본질적 필드 지정

3. 필드 선택기를 사용하여 목록에서 필드를 추가하거나 제거할 수 있습니다. 필드 선택기를 열려면 필드 목록 오른쪽의 아이콘을 클릭하십시오.

맵핑된 필드 탐색

하나의 데이터 스트림 또는 데이터 소스가 또 다른 데이터 스트림이나 데이터 소스로 맵핑되는 위치를 선택하면 맵핑을 위한 필드를 선택하거나 시스템 기본값 맵핑이 올바른지 확인하기 위한 대화 상자가 표시됩니다. 스트림이나 데이터 소스에 기본 필드가 설정되었고 이들이 매치하지 않은 경우에는 이러한 필드는 빨간색으로 표시됩니다. 데이터 소스에서 맵핑되지 않은 필드는 변경되지 않은 필터 노드를 통과하지만 비필수 필드 또한 맵핑할 수 있음을 유의하십시오.

원래값. 템플리트의 모든 필드 또는 기존 스트림의 현재 추가적 다운스트림인 모든 필드를 나열합니다. 새 데이터 소스의 필드는 이러한 필드로 맵핑됩니다.

맵핑됨. 템플리트 필드에 맵핑하기 위해 선택된 필드를 나열합니다. 이들은 이름을 스트림 작업에 사용된 원래 필드와 매치하도록 변경해야 할 수도 있는 필드입니다. 필드가 사용 가능한 필드 목록을 활성화하기 위해 테이블 셀을 클릭하십시오.

어떤 필드를 맵핑할지 확실하지 않으면 맵핑 전에 소스 데이터를 자세히 검사하는 것이 유용할 수 있습니다. 예를 들어, 소스 노드에서 유형 템을 사용하여 소스 데이터의 요약을 검토할 수 있습니다.

팁 및 단축키

다음 단축 아이콘과 팁에 익숙해지면 빠르고 쉽게 작업할 수 있습니다.

- **두 번 클릭하여 스트림을 빠르게 작성.** 팔레트에서 노드를 두 번 클릭하기만 하면 현재 스트림에 추가하고 연결할 수 있습니다.
- **다운스트림 노드를 선택하기 위해 키 조합 사용.** Ctrl+Q 및 Ctrl+W를 눌러 모든 노드 다운스트림 선택을 토글할 수 있습니다.

- **단축키를 사용하여 노드에 연결 및 연결 해제.** 캔버스에서 노드가 선택되었을 때 연결을 시작하려면 F2를 누르고, 필수 노드로 이동하려면 템을 누르고, 연결을 완료하려면 Shift+스페이스바를 누르십시오. 선택한 노드에 대한 모든 입력과 출력을 연결 해제하려면 F3을 누르십시오.
- **노드 팔레트 템을 즐겨찾는 노드로 사용자 정의.** 도구 메뉴에서 팔레트 관리를 클릭하여 노드 팔레트에 표시된 노드를 추가, 제거 또는 이동하기 위한 대화 상자를 여십시오.
- **노드 이름 변경 및 도구 팁 추가.** 각 노드 대화 상자에는 캔버스에서 노드에 사용자 정의 이름을 지정할 수 있는 주석(Annotation) 템뿐만 아니라 스트림을 구성하는데 도움을 주는 도구 팁이 포함됩니다. 진행률을 추적하고, 프로세스 세부사항을 저장하고, 비즈니스 의사결정을 필수 또는 달성으로 표시하기 위해 긴 주석(Annotation)을 포함할 수도 있습니다.
- **CLEM 표현식에 자동으로 값 삽입.** 다양한 대화 상자(예: 파생 및 채움 노드의 대화 상자)에서 액세스 가능한 표현식 작성기를 사용하여 필드 값을 CLEM 표현식에 자동으로 삽입할 수 있습니다. 표현식 작성기의 값 단추를 클릭하여 기존 필드 값에서 선택하십시오.



그림 13. 값 단추

- **신속하게 파일 찾아보기.** 열기 대화 상자에서 파일을 찾아볼 때, 파일 목록(찾아보기 필드 옆의 대화 상자 맨 위에 있는 노란색 다이아몬드 단추 클릭)을 사용하여 이전에 사용된 디렉토리뿐만 아니라 IBM SPSS Modeler 기본 디렉토리에 액세스하십시오. 전진 및 이전 단추를 사용하여 액세스된 디렉토리를 통해 스크롤하십시오.
 - **출력 창 클러터 최소화.** 모든 출력 창 오른쪽 상단 구석에 있는 빨간색 X 단추를 사용하여 출력을 빠르게 닫고 삭제할 수 있습니다. 그러면 관리자 분할창의 출력 탭에 유망하거나 흥미로운 결과만을 유지할 수 있습니다.
- 키보드 단축키의 전체 범위를 소프트웨어에 사용 가능합니다. 자세한 정보는 222 페이지의 『내게 필요한 옵션의 키보드 기능』의 내용을 참조하십시오.

다음을 수행할 수 있습니다.

- 마우스를 사용하여 스트림 캔버스에서 노드 그룹을 끌고 선택합니다.
- 하나의 스트림에서 다른 스트림으로 노드를 복사하고 붙여넣을 수 있습니다.
- 모든 대화 상자와 출력 창에서 도움말에 액세스합니다.
- CRISP-DM(Cross-Industry Standard Process for Data Mining)에 대한 도움말을 얻습니다. (도움말 메뉴에서 **CRISP-DM** 도움말을 클릭하십시오.)

제 6 장 데이터에 대한 작업

데이터 노드를 마우스 오른쪽 단추로 클릭한 후 **데이터 보기**를 선택하여 스트림 데이터를 탐색할 수 있습니다. 열리는 창에서 다음 작업을 수행할 수 있습니다.

- **차트** 탭에서는 고급 데이터 시각화를 작성하여 다른 퍼스펙티브에서 데이터를 탐색하고 데이터 내에서 패턴, 연결 및 관계를 식별할 수 있습니다.
- **스프레드시트** 탭에서는 데이터의 읽기 전용 보기와 표 형식으로 표시할 수 있습니다.
- **데이터 검토** 탭에는 각 열에 대한 빈도와 통계가 표시됩니다.
- **대시보드** 탭에서 **레이아웃 디자인 시작**을 클릭하고 여러 개의 차트를 단일 페이지에서 보기 위한 레이아웃을 작성할 수 있습니다. 레이아웃을 템플리트로 저장하고 저장된 차트를 레이아웃의 위치로 끌어다 놓을 수 있습니다.
- **기본 설정** 탭에서는 언어 및 모양과 느낌 등의 사용자 인터페이스 기본 설정을 지정할 수 있습니다.

이 기능은 기본적으로 포트 28900을 사용합니다. 다른 포트를 사용해야 하는 경우 `options.cfg` 파일에서 `data_view_port_number` 구성 설정의 값을 변경하십시오.

참고: 이 **데이터 보기** 기능은 현재 SPSS Modeler의 Subscription 버전에서 사용할 수 없습니다.

차트 작성

차트 탭에서는 미리 정의된 갤러리 차트 또는 개별 부분(예: 축, 막대)을 통해 차트를 작성할 수 있습니다. 차트를 작성하려면 제공된 차트 유형 옵션에서 갤러리 차트 유형 또는 기본 요소를 선택하십시오.

차트를 작성할 때 해당 차트의 미리보기가 캔버스에 표시됩니다. 미리보기에는 실제 데이터를 대표하는 실제 변수 레이블과 측정 수준이 사용됩니다.

초보자라면 차트를 작성할 때 갤러리를 사용하는 것이 좋습니다. 갤러리 사용에 대한 자세한 정보는 [58 페이지](#)의 [『차트 유형 갤러리에서 차트 작성』](#)의 내용을 참조하십시오.

차트 작성기 시작

- 작업하려는 데이터 노드를 마우스 오른쪽 단추로 클릭한 후 **데이터 보기**를 클릭하십시오.

레이아웃 및 용어

시작 화면

차트 작성기를 실행하면 파트 유형을 선택하거나 활성 데이터 세트에서 열을 선택하는 옵션이 표시됩니다. 시각화할 열을 추가하면 차트 유형 옵션이 업데이트되어 선택한 열에 권장되는 차트 유형이 표시됩니다.

캔버스

캔버스는 차트를 작성하는 데 사용되는 차트 작성기 대화 상자의 영역입니다.

차트 유형

사용 가능한 차트 유형이 나열됩니다. 그래픽 요소는 데이터를 나타내는 차트 내 항목입니다. 이러한 항목으로는 막대, 점, 선 등이 있습니다.

세부사항 분할창

세부사항 분할창은 기본 차트 작성 블록을 제공합니다.

차트 설정

차트를 작성하는 데 사용되는 변수, 분포 방법, 제목 및 부제목 필드 등을 선택하는 옵션을 제공합니다. 선택한 차트 유형에 따라 세부사항 분할창에 제공된 옵션이 달라집니다. 차트 유형별 사용 가능한 옵션에 대한 자세한 정보는 [58 페이지](#)의 [『차트 유형』](#)의 내용을 참조하십시오.

동작

차트 구성 파일 다운로드, 차트를 이미지 파일로 다운로드, 차트 재설정 및 글로벌 기본 설정 지정에 대한 옵션을 제공합니다.

차트 유형 갤러리에서 차트 작성

차트를 작성하는 가장 쉬운 방법은 차트 유형 갤러리를 사용하는 것입니다. 다음은 갤러리를 사용하여 차트를 작성하는 일반적인 단계입니다.

1. 차트 유형 섹션에서 차트 범주를 선택하십시오. 선택한 차트 유형의 미리보기 버전이 차트 캔버스에 표시됩니다.
2. 캔버스에 이미 차트가 표시된 경우 새 차트가 차트의 축 세트와 그래픽 요소를 대체합니다.
 - a. 선택한 차트 유형에 따라 사용 가능한 변수는 세부사항 분할창에서 서로 다른 여러 머리말(예: 막대형 차트의 경우 범주, 선형 차트의 경우 **X축** 및 **Y축**) 아래에 표시됩니다. 선택한 차트 유형에 적합한 변수를 선택하십시오.

차트 유형

갤러리에는 가장 일반적으로 사용되는 차트의 컬렉션이 포함되어 있습니다. 다음과 같은 차트가 있습니다.

3차원 차트

단순. 자세한 정보는 [59 페이지의 『3차원 차트』](#)의 내용을 참조하십시오.

막대형 차트

단순, 수직누적, 수평배열. 자세한 정보는 [60 페이지의 『막대형 차트』](#)의 내용을 참조하십시오.

상자도표

단순 및 수평배열. 자세한 정보는 [62 페이지의 『상자도표』](#)의 내용을 참조하십시오.

거품형

단순. 자세한 정보는 [63 페이지의 『거품형 차트』](#)의 내용을 참조하십시오.

원통형 차트

단순. 자세한 정보는 [63 페이지의 『원통형 차트』](#)의 내용을 참조하십시오.

원 채우기

단순. 자세한 정보는 [64 페이지의 『원 채우기 차트』](#)의 내용을 참조하십시오.

사용자 정의 차트

자세한 정보는 [65 페이지의 『사용자 정의 차트』](#)의 내용을 참조하십시오.

이중 Y축 차트

단순. 자세한 정보는 [65 페이지의 『이중 Y축 차트』](#)의 내용을 참조하십시오.

오차 막대 차트

단순. 자세한 정보는 [66 페이지의 『오차 막대 차트』](#)의 내용을 참조하십시오.

평가 차트

단순. 자세한 정보는 [67 페이지의 『평가 차트』](#)의 내용을 참조하십시오.

히트 맵

단순. 자세한 정보는 [69 페이지의 『히트 맵 차트』](#)의 내용을 참조하십시오.

히스토그램

단순, 수직누적, 빈도 다각형. 자세한 정보는 [70 페이지의 『히스토그램 차트』](#)의 내용을 참조하십시오.

선형 차트

단순. 자세한 정보는 [71 페이지의 『선형 차트』](#)의 내용을 참조하십시오.

맵 차트

단순. 자세한 정보는 [72 페이지의 『맵 차트』](#)의 내용을 참조하십시오.

수학 곡선 차트

단순. 자세한 정보는 [75 페이지의 『수학 곡선 차트』](#)의 내용을 참조하십시오.

다중 차트

단순. 자세한 정보는 [76 페이지의 『다중 차트』](#)의 내용을 참조하십시오.

다중 계열 차트

단순. 자세한 정보는 [77 페이지의 『다중 계열 차트』](#)의 내용을 참조하십시오.

평형 차트

단순. 자세한 정보는 [78 페이지의 『평형 차트』](#)의 내용을 참조하십시오.

원형 차트

단순. 자세한 정보는 [78 페이지의 『원형 차트』](#)의 내용을 참조하십시오.

인구 피라미드 차트

단순. 자세한 정보는 [79 페이지의 『인구 피라미드 차트』](#)의 내용을 참조하십시오.

Q-Q 도표

단순. 자세한 정보는 [80 페이지의 『Q-Q 도표』](#)의 내용을 참조하십시오.

방사형 차트

단순. 자세한 정보는 [81 페이지의 『방사형 차트』](#)의 내용을 참조하십시오.

관계 차트

단순. 자세한 정보는 [82 페이지의 『관계 차트』](#)의 내용을 참조하십시오.

산점도 및 점도표

1차원, 단순, 그룹, 오버레이 및 3차원 산점도(요약 점도표, 1차원 점도표, 하락-선 차트). 자세한 정보는 [83 페이지의 『산점도 및 점도표』](#)의 내용을 참조하십시오.

산점도 행렬 차트

단순. 자세한 정보는 [84 페이지의 『산점도 행렬 차트』](#)의 내용을 참조하십시오.

햇살 차트

단순. 자세한 정보는 [85 페이지의 『햇살 차트』](#)의 내용을 참조하십시오.

시간 도표

단순. 자세한 정보는 [86 페이지의 『시간 도표』](#)의 내용을 참조하십시오.

테마 리버

단순. 자세한 정보는 [87 페이지의 『테마 리버 차트』](#)의 내용을 참조하십시오.

트리 차트

단순. 자세한 정보는 [88 페이지의 『트리 차트』](#)의 내용을 참조하십시오.

트리 맵 차트

단순 및 햇살. 자세한 정보는 [89 페이지의 『트리 맵 차트』](#)의 내용을 참조하십시오.

t-SNE 차트

단순. 자세한 정보는 [90 페이지의 『t-SNE 차트』](#)의 내용을 참조하십시오.

단어 클라우드

단순. 자세한 정보는 [90 페이지의 『단어 클라우드 차트』](#)의 내용을 참조하십시오.

3차원 차트

3차원 차트는 일반적으로 다중 변수 함수를 나타내는데 사용되며, X축 및 Y축 변수의 함수인 Z축 변수를 포함합니다.

단순 3차원 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 **3차원** 아이콘을 클릭하십시오.

3D

캔버스가 업데이트되어 3차원 차트 템플리트가 표시됩니다.

2. 드롭 다운 목록에서 차트 유형을 선택하십시오.
3. 드롭 다운 목록에서 **X축** 변수를 선택하십시오.
4. 드롭 다운 목록에서 **Y축** 변수를 선택하십시오.
5. 드롭 다운 목록에서 **Z축** 변수를 선택하십시오.

추가 기능

유형

데이터를 표시하는데 사용할 수 있는 차트 유형이 나열됩니다.

X축

차트의 x축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

Y축

차트의 y축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

Z축

차트의 z축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

도구 팁 정보

데이터 점 위로 마우스를 이동할 때 표시되는 도구 팁 정보를 생성하는데 사용할 수 있는 변수가 나열됩니다.

색상 맵

사용 가능한 색상 맵 변수가 나열됩니다. 이 변수는 색상 진행(지정된 열의 값 범위를 기준으로 함)을 사용하여 플롯 점에 자신을 표시합니다. 색상 맵은 코로플레스 맵이라고도 합니다.

크기 맵

사용 가능한 크기 맵 변수가 나열됩니다. 이 변수는 서로 다른 크기를 사용하여 플롯 점에 자신을 표시합니다.

Z 비율

x축과 y축에 상대적인 z축 데이터 값의 척도를 설정합니다.

회전

차트 회전을 사용 및 사용 안함으로 설정합니다.

데이터 점 도구 팁

데이터 점 도구 팁이 표시되는 위치(데이터 점의 오른쪽, 차트의 오른쪽 상단, 숨기기)를 제어합니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

막대형 차트

막대형 차트는 범주형 변수 요약 시 유용합니다. 예를 들어 막대형 차트를 사용하여 설문조사에 참여한 남자 및 여자 인원 수를 표시하거나, 남자의 평균 급여 및 여자의 평균 급여를 표시할 수 있습니다.

단순 막대형 차트 작성

1. 차트 작성기의 차트 유형 세션에서 **막대** 아이콘을 클릭하십시오.

Bar
Bar

캔버스가 업데이트되어 막대형 차트 템플리트가 표시됩니다.

2. 범주형(명목 또는 순서) 변수를 범주 변수로 선택하십시오. 척도변수를 사용할 수는 있지만 소수의 특별한 케이스에서만 결과가 유용합니다. 막대형 차트는 제한된 수의 고별된 값이 있는 경우에 가장 적절하게 표시됩니다. 척도를 범주 축으로 사용하여 막대형 차트를 작성할 경우 각 막대가 정확한 값에서 그려지고 막대가 다른 연속 값과 겹칠 수 없기 때문에 막대가 매우 얇습니다.
3. **요약** 목록에서 통계를 선택하십시오. 통계 결과에 따라 막대 높이가 결정됩니다. 원하는 통계가 **요약** 목록에 나타나지 않을 경우 변수가 필요할 수 있습니다. 값 목록에서 변수를 선택하고 통계가 지금 사용 가능한지 확인하십시오. 기타 차트 유형 제한사항이 있을 수 있습니다. 예를 들어 오차 막대 차트는 특정 통계에 대해서만 계산할 수 있습니다.

추가 기능

Category

차트의 x축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

순서 기준

변수 내의 범주에 대한 정렬 옵션을 선택하십시오.

범주 이름

변수의 범주를 정렬하는 데 범주 레이블을 사용합니다. 이는 차트에 표시되는 레이블로, 일반적으로 눈금 또는 범례 레이블입니다.

범주 값

변수의 범주를 정렬하는 데 데이터 세트에 저장된 값을 사용합니다. 범주 값은 데이터 세트에서 범주를 식별하는 데 사용됩니다. 일반적으로 이 값은 레이블과는 다르며 구체적인 설명을 제공하지 않아도 됩니다. 예를 들어, 값은 1 같은 숫자가 될 수 있는 반면, 레이블은 여성과 같이 범주에 대한 텍스트 설명입니다.

범주 순서

변수 범주의 정렬 순서를 선택하십시오.

읽는 대로

변수 범주가 데이터 세트에 나타나는 대로 표시됩니다.

오름차순

변수 범주가 오름차순으로 정렬됩니다.

내림차순

변수 범주가 내림차순으로 정렬됩니다.

요약

그래픽 요소에 대한 통계 요약함수를 선택하십시오. 통계 결과에 따라 y 축에서 그래픽 요소의 위치가 결정됩니다. 2차원 차트의 경우 축의 각 x 값에 대해 통계가 계산됩니다. 3차원 차트의 경우 x 축 및 z 축의 교차 값에 대해 통계가 계산됩니다.

두 가지 유형의 통계 요약함수가 있습니다. **값** 변수를 지정해야 하는지 여부가 판별되기 때문에 구분하는 것이 중요합니다.

- **값** 변수가 필요하지 않은 함수: 변수가 필요하지 않은 함수입니다. 모든 개수 및 퍼센트 통계가 이 범주에 속합니다. 이 통계는 정의된 **값** 변수가 없을 때 사용할 수 있습니다.
- **값** 변수가 필요한 함수: 이 함수는 **값** 변수가 필요한 함수입니다. 예를 들어 평균 함수의 경우 평균을 계산할 때 변수가 필요합니다. 이 통계는 정의된 **값** 변수가 있을 때 사용할 수 있습니다.

값

이 필드는 **값** 변수가 필요한 **요약** 함수를 선택한 경우에 표시됩니다. **값**으로 사용할 변수를 선택하십시오.

분할 기준

분할 기준 변수에 범주별 셀이 있는 차트 테이블을 작성하는 범주형 변수를 선택하십시오. 그룹화와 마찬가지로, 분할 기준 변수는 기본적으로 변수 범주 각각에 대한 정보를 표시하여 차트에 차원을 더 추가합니다.

분할 유형

분할 기준 변수를 선택한 경우 생성되는 범주 막대를 수직누적 또는 수평배열로 표시하도록 선택할 수 있습니다. 군집 및 누적이 차트 내에서 차원으로 추가됩니다. 군집은 한 막대를 여러 막대로 분할하고, 누적은 각 막대에 여러 세그먼트를 작성합니다. 누적의 경우 정확한 통계를 선택해야 합니다. **값**을 함께 추가하면(누적) 적절한 결과가 제공되어야 합니다. 예를 들어 평균값을 더하거나 누적하는 것은 일반적으로 의미가 없습니다.

막대 유형

제공된 옵션에서 막대형 차트 유형을 선택하십시오.

- X축
- Y축
- X축 반전
- Y축 반전
- 극좌표 각도 축
- 극반경 축
- 극도표-레인보우

레이블 위치

드롭 다운 메뉴에서 차트의 레이블 위치를 선택하십시오.

- 없음
- 위쪽
- 왼쪽

- 오른쪽
- 아래쪽
- 내부
- insideLeft
- insideRight
- insideTop
- insideBottom
- insideTopLeft
- insideBottomLeft
- insideTopRight
- insideBottomRight

참조선 표시

이 전환 컨트롤은 차트에 참조선 표시를 사용/사용 안함으로 설정합니다. 사용 가능한 옵션은 **최소값**, **최대값** 및 **평균**이며, 차트의 최소값, 최대값 및 평균값에 각각 참조선을 표시합니다.

참조선 값 입력

참조선 표시가 사용으로 설정된 경우 이 설정을 사용하여 특정 참조선 값을 지정할 수 있습니다. 추가 참조선 값을 지정하려면 다른 열 추가를 클릭하십시오.

전치

사용으로 설정할 경우 차트의 x 축과 y 축이 전치됩니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

상자도표

상자도표는 5개의 통계(최소값, 첫 번째 사분위수, 중위수, 세 번째 사분위수 및 최대값)를 표시합니다. 척도변수의 분포를 표시하고 이상값을 정확히 표시하는 데 유용합니다.

단순 상자도표 작성

1. 차트 작성기의 **차트 유형** 섹션에서 **상자도표** 아이콘을 클릭하십시오.



캔버스가 업데이트되어 상자도표 차트 템플리트가 표시됩니다.

2. 하나 이상의 척도변수를 **열** 변수로 선택하십시오.

참고: 점도표에 대한 통계는 상자도표입니다. 이는 변경할 수 없습니다.

추가 기능

열

차트의 x 축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

열을 더 추가하려면 다른 열 추가를 클릭하십시오.

분할 기준

분할 기준 변수에 범주별 셀이 있는 차트 테이블을 작성하는 범주형 변수를 선택하십시오. 그룹화와 마찬가지로, 분할 기준 변수는 기본적으로 변수 범주 각각에 대한 정보를 표시하여 차트에 차원을 더 추가합니다.

범주 순서

변수 범주의 정렬 순서를 선택하십시오.

읽는 대로

변수 범주가 데이터 세트에 나타나는 대로 표시됩니다.

오름차순

변수 범주가 오름차순으로 정렬됩니다.

내림차순

변수 범주가 내림차순으로 정렬됩니다.

데이터 표준화

사용으로 설정할 경우 이 설정은 다중 데이터 세트 또는 다중 열의 데이터를 비교할 수 있도록 데이터를 정규 분포로 변환합니다. 이 설정은 빈도 및 변환 통계에 대한 100% 수직누적을 퍼센트로 작성합니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

거품형 차트

거품형 차트에서는 그룹의 범주를 계층 구조가 아닌 원으로 채워 표시합니다. 각 원(거품)의 크기는 해당 값에 비례합니다. 거품형 차트는 데이터의 관계를 비교하는 데 유용합니다.

단순 거품형 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 거품형 아이콘을 클릭하십시오.

Bubble



캔버스가 업데이트되어 거품형 차트 템플리트가 표시됩니다.

2. 드롭 다운 목록에서 열 변수를 선택하십시오.

참고: 열 변수를 더 포함시키려면 다른 열 추가를 클릭하십시오.

추가 기능

그룹 색상

색상 그룹화를 켜거나 끕니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

원통형 차트

원통형 차트는 재무 차트의 한 스타일로, 유가증권, 파생상품 또는 통화의 가격 변화를 설명하는 데 사용됩니다. 각 원통형 요소는 일반적으로 1일을 나타냅니다. 따라서 한 달 차트에는 거래일 20일이 20개의 원통형 요소로 표시될 수 있습니다. 원통형 차트는 주식 및 통화 가격 패턴 분석에 가장 자주 사용되며 상자도표와 유사합니다.

원통형 차트를 작성하는 데 사용되는 데이터 세트에는 표시하려는 각 기간에 대한 고가, 저가 및 종가가 포함되어 있어야 합니다.

단순 원통형 차트 작성

1. 도표 작성기의 도표 유형 섹션에서 원통형 차트 아이콘을 클릭하십시오.



캔버스가 업데이트되어 원통형 차트 템플릿이 표시됩니다.

2. 변수를 **X축** 변수로 선택하십시오.
3. 변수를 **고가** 변수로 선택하십시오.
4. 변수를 **저가** 변수로 선택하십시오.

추가 기능

X축

차트의 x축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

고가

차트의 고가 값에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

상한 필드 요약

선택한 상한 변수에 대한 통계 요약함수를 선택하십시오.

저가

차트의 저가 값에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

시가

차트의 시가 값에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

종가

차트의 종가 값에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

볼륨

차트의 볼륨 막대에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

범주 순서

변수 범주의 정렬 순서를 선택하십시오.

읽는 대로

변수 범주가 데이터 세트에 나타나는 대로 표시됩니다.

오름차순

변수 범주가 오름차순으로 정렬됩니다.

내림차순

변수 범주가 내림차순으로 정렬됩니다.

원통형

차트 데이터를 원통형 또는 선형으로 표시하도록 토글합니다.

이동 평균

이동 평균을 차트에 표시하기 위한 옵션을 제공합니다. 사용 가능한 옵션은 5, 10, 20, 30, 60 및 120입니다.

색상 전환

이 전환 컨트롤은 증가 색상과 감소 색상을 전환합니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

원 채우기 차트

원 채우기 차트는 계층 구조 데이터를 중첩 영역 세트로 표시하므로, 대량의 계층 구조 데이터를 시각화할 수 있습니다. 이 도표는 나무 지도와 비슷하지만 직사각형 대신 원을 사용합니다. 원 채우기 차트에서는 포함(중첩)을 사용하여 계층 구조 데이터를 표시합니다.

단순 원 채우기 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 시간 도표 아이콘을 클릭하십시오.

Circle packing



캔버스가 업데이트되어 원 채우기 차트 템플릿이 표시됩니다.

2. 드롭 다운 목록에서 열 변수를 선택하십시오.

참고: 열 변수를 더 포함시키려면 다른 열 추가를 클릭하십시오.

추가 기능

그룹 색상

색상 그룹화를 켜거나 끕니다.

요약

통계 요약함수를 선택하십시오. 각 범주를 요약하는 데 사용되는 방법입니다.

두 가지 유형의 통계 요약함수가 있습니다. 값 변수를 지정해야 하는지 여부가 판별되기 때문에 구분하는 것이 중요합니다.

- **값 변수가 필요하지 않은 함수:** 변수가 필요하지 않은 함수입니다. 모든 개수 및 퍼센트 통계가 이 범주에 속합니다. 이 통계는 정의된 값 변수가 없을 때 사용할 수 있습니다.
- **값 변수가 필요한 함수:** 이 함수는 값 변수가 필요한 함수입니다. 예를 들어 평균 함수의 경우 평균을 계산 할 때 변수가 필요합니다. 이 통계는 정의된 값 변수가 있을 때 사용할 수 있습니다.

값

이 필드는 값 변수가 필요한 **요약** 함수를 선택한 경우에 표시됩니다. 값으로 사용할 변수를 선택하십시오.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

사용자 정의 차트

사용자 정의 차트 옵션으로 JSON 코드를 붙여넣거나 편집하여 원하는 차트를 작성할 수 있습니다.

사용자 정의 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 사용자 정의 아이콘을 클릭하십시오.



2. 차트 지정 사항이 포함된 JSON 코드를 세부사항 분할창의 제공된 필드에 붙여넣으십시오.

3. 차트 생성을 클릭하십시오.

이중 Y축 차트

이중 Y축 차트를 사용하여 서로 다른 도메인을 사용하는 두 개의 Y축 변수를 요약하거나 차트로 작성할 수 있습니다. 예를 들어 케이스 수는 특정 축에, 평균 급여는 다른 축에 도표로 작성할 수 있습니다. 또한 이 차트는 이전에 설명한 다양한 차트 유형이 이중 Y축 차트에 포함되도록 다양한 그래픽 요소가 혼합된 것일 수 있습니다. 개수를 선으로 표시하고 각 범주의 평균을 막대로 표시할 수 있습니다.

단순 이중 Y축 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 **이중 Y축** 아이콘을 클릭하십시오.



캔버스가 업데이트되어 이중 Y축 차트 템플리트가 표시됩니다.

2. 변수를 **X축** 변수로 선택하십시오.

3. 첫 번째 **Y축** 변수에 대한 변수를 선택한 후 변수를 나타내는 차트 유형(막대, 선, 산점도)을 선택하십시오.

4. 두 번째 **Y축** 변수에 대한 변수를 선택한 후 변수를 나타내는 차트 유형(막대, 선, 산점도)을 선택하십시오.

참고: 위로 및 아래로 화살표 컨트롤을 사용하여 y축 순서를 변경할 수 있습니다.

추가 기능

X축

차트의 x축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

Y축

차트의 이중 y축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

요약

사용으로 설정할 경우 각 범주의 요약 방법을 선택할 수 있는 옵션이 표시됩니다.

왼쪽 Y축 요약

차트의 왼쪽에 표시될 Y축에 대한 요약 방법을 설정합니다. 옵션은 합계, 평균, 최대값, 최소값입니다.

오른쪽 Y축 요약

차트의 오른쪽에 표시될 Y축에 대한 요약 방법을 설정합니다. 옵션은 합계, 평균, 최대값, 최소값입니다.

데이터 표준화

사용으로 설정할 경우 데이터가 정규 분포로 변환됩니다. 이 경우 여러 데이터 세트 또는 열의 데이터를 쉽게 비교할 수 있습니다.

두 번째 축선

사용으로 설정할 경우 차트의 두 번째 축선이 표시됩니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

오차 막대 차트

오차 막대 차트는 데이터의 변동을 나타내며 보고된 측정치로 오차(불확실성)를 표시합니다. 오차 막대로 차이가 통계적으로 유의적인지 판별할 수 있습니다. 오차 막대는 제공된 함수의 적합도를 제시할 수도 있습니다.

단순 오차 막대 차트 작성

1. 도표 작성기의 도표 유형 섹션에서 **오차 막대** 아이콘을 클릭하십시오.



캔버스가 업데이트되어 오차 막대 차트 템플리트가 표시됩니다.

2. 척도변수를 범주 변수로 선택하십시오. 이 변수의 데이터는 x축에 표시됩니다.

3. 변수를 **Y축** 변수로 선택하십시오. 이 변수의 데이터는 y축에 표시됩니다.

추가 기능

Category

차트의 x축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

Y축

차트의 y축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

범주 순서

변수 범주의 정렬 순서를 선택하십시오.

읽는 대로

변수 범주가 데이터 세트에 나타나는 대로 표시됩니다.

오름차순

변수 범주가 오름차순으로 정렬됩니다.

내림차순

변수 범주가 내림차순으로 정렬됩니다.

분할 기준

분할 기준 변수에 범주별 셀이 있는 차트 테이블을 작성하는 범주형 변수를 선택하십시오. 그룹화와 마찬가지로, 분할 기준 변수는 기본적으로 변수 범주 각각에 대한 정보를 표시하여 차트에 차원을 더 추가합니다.

참조선

사용으로 설정할 경우 차트에 참조선이 표시됩니다. 참조선은 선택한 통계 방법과 상관됩니다.

오차 막대

사용으로 설정할 경우 오차 범위를 나타내는 선이 차트에 표시됩니다.

측도

오차 막대로 표시할 측도 유형을 선택하십시오.

신뢰구간

선택한 변수에 대한 신뢰구간을 설정합니다. 기본값은 0.95(95%)이며, 표시 값 필드에 적용된 것과 같습니다.

표준 오차

선택한 변수의 표준 오차를 측정합니다.

표준 편차

선택한 변수의 표준 편차를 측정합니다.

신뢰수준

이 값은 선택한 측도에 대한 신뢰구간을 나타냅니다. 기본값은 0.95(95%)입니다.

통계 방법

중심 경향 설명 방법을 선택하십시오.

평균

비율을 합산한 후 그 결과를 총 비율 수로 나눈 결과입니다.

중앙값

이 값보다 작은 비율의 수와 이 값보다 큰 비율의 수가 같은 값입니다.

표시 모드

통계 방법 선택사항의 표시 방법(막대, 선, 원)을 선택하십시오.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

평가 차트

평가 차트는 히스토그램 또는 콜렉션 그래프와 비슷합니다. 평가 차트는 특정 결과 예측 시 모델이 작동하는 방식을 보여줍니다. 평가 차트는 예측의 신뢰도 및 예측값을 기반으로 레코드를 정렬하고 레코드를 동일한 크기의 그룹(분위수)으로 분할한 후 각 분위수에 대한 기준의 값을 내림차순으로 도표로 작성하여 작동합니다. 다중 모델이 도표에 선구분 변수로 표시됩니다.

결과는 특정 값 또는 값 범위를 **적중**으로 정의하여 처리됩니다. 적중은 일반적으로 관심 있는 이벤트(예: 특정 의료 진단) 또는 일부 정렬(예: 고객에 대한 판매)의 성공을 표시합니다.

플래그

출력 필드는 직설적이어서 적중은 참 값에 해당합니다.

명목

명목 출력 필드의 경우 세트의 첫 번째 값이 적중을 정의합니다.

연속

연속 출력 필드의 경우 적중은 필드 범위의 중심점보다 큰 값과 동일합니다.

각각의 점이 해당 분위수에 대한 값에 더 높은 모든 분위수를 더한 값과 동일하도록 평가 차트는 누적일 수도 있습니다. 누적 차트가 일반적으로 모델의 전체 성능을 더 잘 전달하지만 비누적 차트가 모델에 대한 특정 문제점 영역 표시에서 뛰어날 수도 있습니다.

단순 평가 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 **평가** 아이콘을 클릭하십시오.

Evaluation



캔버스가 업데이트되어 평가 차트 템플리트가 표시됩니다.

2. **목표 필드**, **예측 필드** 및 **신뢰도 필드** 변수를 설정하십시오. 목표 필드는 둘 이상의 값을 가진 명목 필드 또는 인스턴스화된 플래그일 수 있습니다. 예측 필드는 예측값으로 사용되는 변수를 정의합니다. 신뢰도 필드는 예측의 신뢰도를 설정하는 데 사용되는 변수를 정의합니다.

참고: 예측 필드 변수 유형은 목표 필드에 대해 선택된 변수 유형과 일치해야 합니다.

3. 사용자 정의 적중을 표시하는 데 사용하는 사용자 정의 조건을 지정하십시오. 이 옵션은 값의 순서 및 대상 필드의 유형에서 관심 있는 결과를 추론할 때보다 관심 있는 결과를 정의하는 경우에 유용합니다.

적중 조건에 대해 CLEM 표현식을 지정해야 합니다. 예를 들어, @TARGET = "YES"는 목표 필드에 대한 Yes 값이 평가에서 적중으로 계수됨을 나타내는 유효한 조건입니다. 지정된 조건은 모든 대상 필드에 사용됩니다.

추가 기능

누적 도표

사용으로 설정된 경우 누적 차트를 작성합니다. 누적 차트에 각 분위수 및 더 높은 분위수에 대해 값이 표시됩니다.

표시 모드

이 설정은 미리보기 모드 및 출력에 표시되는 차트를 제어합니다.

클래식 모드

선택할 경우 모델 분류 조정, 분리점, 행렬 막대, ROC, 이득, ROI 및 이익 차트가 미리보기 모드와 출력에 표시됩니다.

단일 모드

선택할 경우 모델 분류 조정 차트가 미리보기 모드와 출력에 표시되는 차트에만 있습니다.

전체 모드

선택할 경우 모델 분류 조정, 분리점, 행렬 막대, ROC, 이득, ROI, 이익, GINI, 리프트 및 반응 차트가 미리보기 모드와 출력에 표시됩니다.

평가 차트

분리점

분리점 차트는 지정된 분리점 값의 선택된 변수에 대한 예측 값과 실제 값을 보여줍니다.

행렬 막대

행렬 막대 차트는 다중 변수 간에 선형 상관관계가 있는지 판별하는 좋은 방법입니다.

ROC

ROC(Receiver Operating Characteristic)는 개체를 분류하는 두 개의 범주가 하나의 변수에 있는 분류 방법의 성능을 평가합니다.

이득

이득은 각 분위수에서 발생하는 적중 총계의 비율로 정의됩니다. 이득은 (분위수의 적중 수 / 적중 수 총 계) $\times 100\%$ 로 계산됩니다.

ROI

ROI(Return On Investment)는 수입 및 비용 정의를 포함한다는 점에서 이익과 비슷합니다. ROI는 분위 수에 대한 비용과 이익을 비교합니다. ROI는 (분위수에 대한 이익 / 분위수에 대한 비용) $\times 100\%$ 로 계산 됩니다.

이익

이익은 각 레코드의 수입에서 레코드의 비용을 뺀 값입니다. 분위수의 이익은 단순히 분위수의 전체 레코 드 이익 합계입니다. 수입은 적중에만 적용되는 것으로 가정되지만 비용은 모든 레코드에 적용됩니다. 이 익 및 비용은 고정이거나 데이터의 필드에 의해 정의될 수 있습니다. 이익은 (분위수의 레코드에 대한 수 입의 합계 - 분위수의 레코드에 대한 비용의 합계)로 계산됩니다.

Kolmogorov-Smirnov

한 변수에 대해 관측된 누적 분포 함수를 지정된 이론적 분포와 비교합니다. 이때 이론적 분포는 정규 분 포, 균일 분포, 지수 분포, 포아송 분포일 수 있습니다.

GINI

GINI는 통계적 산포도를 측정하고 수입 또는 재산 분포를 나타냅니다. 불균등을 측정하는 데 가장 일반 적으로 사용됩니다.

리프트

리프트는 적중인 각 분위수의 레코드 백분율을 훈련 데이터의 전체 적중 백분율과 비교합니다. 리프트는 (분위수의 적중 수 / 분위수의 레코드 수) / (적중 총계 / 레코드 총계)로 계산됩니다.

응답

반응은 단순히 적중인 분위수의 레코드 백분율입니다. 반응은 (분위수의 적중 수 / 분위수의 레코드 수) $\times 100\%$ 로 계산됩니다.

평가 차트 설정

다음 설정은 이익과 ROI 차트에만 적용됩니다.

비용

각 레코드와 연관된 고정 비용을 지정합니다.

수입

적중을 나타내는 각 레코드와 연관된 고정수입을 지정합니다.

가중치

데이터의 레코드가 둘 이상의 단위를 표시하는 경우 빈도 가중치를 사용하여 결과를 조정할 수 있습니다. 각 레코드와 연관된 고정 가중치를 지정하십시오.

히트 맵 차트

히트 맵 차트는 행렬에 포함된 개별 값이 색상으로 표시되는 데이터를 제공합니다.

단순 히트 맵 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 히트 맵 아이콘을 클릭하십시오.

Heat map



캔버스가 업데이트되어 히트 맵 차트 템플릿이 표시됩니다.

2. 변수를 열 변수로 선택하십시오. 각 변수 범주가 개별 차트 열로 표시됩니다.
3. 변수를 행 변수로 선택하십시오. 각 변수 범주가 개별 차트 행으로 표시됩니다.

추가 기능

Column

차트 열에 사용할 수 있는 데이터 세트 변수가 나열됩니다. 각 변수 범주가 개별 차트 열로 표시됩니다.

행

차트 행에 사용할 수 있는 데이터 세트 변수가 나열됩니다. 각 변수 범주가 개별 차트 행으로 표시됩니다.

범주 순서

변수 범주의 정렬 순서를 선택하십시오.

읽는 대로

변수 범주가 데이터 세트에 나타나는 대로 표시됩니다.

오름차순

변수 범주가 오름차순으로 정렬됩니다.

내림차순

변수 범주가 내림차순으로 정렬됩니다.

요약

그래픽 요소에 대한 통계 요약함수를 선택하십시오. 통계 결과에 따라 y 축에서 그래픽 요소의 위치가 결정됩니다. 2차원 차트의 경우 축의 각 x 값에 대해 통계가 계산됩니다. 3차원 차트의 경우 x 축 및 z 축의 교차 값에 대해 통계가 계산됩니다.

두 가지 유형의 통계 요약함수가 있습니다. **값** 변수를 지정해야 하는지 여부가 판별되기 때문에 구분하는 것이 중요합니다.

- 값 변수가 필요하지 않은 함수:** 변수가 필요하지 않은 함수입니다. 모든 개수 및 퍼센트 통계가 이 범주에 속합니다. 이 통계는 정의된 **값** 변수가 없을 때 사용할 수 있습니다.
- 값 변수가 필요한 함수:** 이 함수는 **값** 변수가 필요한 함수입니다. 예를 들어 평균 함수의 경우 평균을 계산할 때 변수가 필요합니다. 이 통계는 정의된 **값** 변수가 있을 때 사용할 수 있습니다.

값

이 필드는 **값** 변수가 필요한 **요약** 함수를 선택한 경우에 표시됩니다. **값**으로 사용할 변수를 선택하십시오.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

히스토그램 차트

히스토그램은 막대형 차트와 모양이 유사하지만, 범주를 비교하거나 시간 경과에 따른 추세를 찾을 수 있습니다. 각 막대는 데이터가 단일 범주에서 어떻게 분포되어 있는지를 나타냅니다. 각 막대는 특정 데이터 점에 대해 연속 범위의 데이터 또는 빈도 수를 나타냅니다.

히스토그램은 단일 척도변수의 분포를 표시하는 데 유용합니다. 데이터는 개수 또는 퍼센트 통계를 사용하여 구간화되고 요약됩니다. 빈도 다각형은 히스토그램의 변동이며 막대 그래픽 요소 대신 영역 그래픽 요소가 사용되는 것을 제외하고 일반적인 히스토그램과 같습니다.

다른 히스토그램의 변동은 인구 피라미드입니다. 이 이름은 가장 일반적인 용도(인구 데이터 요약)에서 파생되었습니다. 인구 데이터와 함께 사용하는 경우 성별로 분할되어 나이 데이터에 대해 두 개의 백 투 백, 수평 히스토그램을 제공합니다. 젊은이 인구가 있는 국가의 경우 결과 그래프의 형태가 피라미드와 유사합니다.

히스토그램 차트 작성

- 차트 작성기의 차트 유형 섹션에서 히스토그램 아이콘을 클릭하십시오.



캔버스가 업데이트되어 히스토그램 차트 템플리트가 표시됩니다.

- 척도변수를 **X축** 변수로 선택하십시오.

참고: 히스토그램에 대한 통계는 히스토그램 또는 히스토그램 퍼센트입니다. 이러한 통계는 데이터를 구간화하고 각 구간에 대한 개수를 계산합니다.

추가 기능

X축

차트의 x축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

분할 기준

분할 기준 변수에 범주별 셀이 있는 차트 테이블을 작성하는 범주형 변수를 선택하십시오. 그룹화와 마찬가지로, 분할 기준 변수는 기본적으로 변수 범주 각각에 대한 정보를 표시하여 차트에 차원을 더 추가합니다.

KDE 곡선 표시

사용으로 설정할 경우, 커널 밀도 추정값 곡선이 차트에 표시됩니다.

분포 곡선 표시

사용으로 설정할 경우 분포 적합 곡선이 차트에 표시됩니다.

분포

드롭 다운 목록은 다음 분포 옵션을 제공합니다.

분포 자동 적합

기본 설정입니다.

베타

지정한 형태 모수를 사용하여 베타 분포의 값을 반환합니다.

지수

지수 분포에서 값을 리턴합니다.

감마

지정한 형태와 척도 모수를 사용하여 감마 분포의 값을 반환합니다.

로그정규

지정한 모수를 사용하여 로그-정규 분포의 값을 반환합니다.

정규

지정한 평균과 표준 편차를 사용하여 정규 분포의 값을 반환합니다.

삼각

지정한 모수를 사용하여 삼각 분포의 값을 반환합니다.

균일

최소값과 최대값 간 균일 분포의 값을 반환합니다.

와이블

지정한 모수를 사용하여 Weibull 분포의 값을 반환합니다.

구간 너비

데이터를 그룹으로 분할하는 데 사용되는 간격의 크기를 슬라이더로 제어할 수 있습니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

선형 차트

선형 차트는 일련의 데이터 점을 그래프에 도표화한 후 이를 선으로 연결합니다. 선형 차트는 미묘한 차이가 있는 추세선을 표시하거나 서로 교차하는 데이터 선이 있는 경우에 특히 유용합니다. 선형 차트를 사용하여 범주형 변수를 요약할 수 있으며 이 경우 막대형 차트와 비슷합니다([60 페이지의 『막대형 차트』](#) 참조). 선형 차트는 시계열 데이터에도 유용합니다.

단순 시계열 선형 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 선 아이콘을 클릭하십시오.



캔버스가 업데이트되어 선형 차트 템플리트가 표시됩니다.

2. 날짜변수를 **X축** 변수로 선택하십시오.

3. 척도변수를 **Y축** 변수로 선택하십시오. 이 변수는 값이 시간 경과에 따라 기록된 변수입니다.

추가 기능

X축

차트의 x축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

Y축

차트의 y축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

분할 기준

분할 기준 변수에 범주별 셀이 있는 차트 테이블을 작성하는 범주형 변수를 선택하십시오. 그룹화와 마찬가지로, 분할 기준 변수는 기본적으로 변수 범주 각각에 대한 정보를 표시하여 차트에 차원을 더 추가합니다.

영역

사용으로 설정할 경우 선 아래의 영역이 다른 색상으로 채워집니다.

평활

사용으로 설정할 경우 차트에 평활 곡선이 표시됩니다.

데이터 점 표시

사용으로 설정할 경우 차트에 데이터 점이 표시됩니다.

재정렬

이 전환 컨트롤은 X축 및 Y축 값을 기준으로 데이터를 다시 정렬합니다.

참조선 표시

사용으로 설정할 경우, X축 및 Y축 값을 기준으로 하는 차트에 참조선이 표시됩니다.

X축에 참조선 값 입력

참조선 표시가 사용으로 설정된 경우 이 설정을 사용하여 X축에 대한 특정 참조선 값을 지정할 수 있습니다. 추가 참조선 값을 지정하려면 다른 열 추가를 클릭하십시오.

Y축에 참조선 값 입력

참조선 표시가 사용으로 설정된 경우 이 설정을 사용하여 Y축에 대한 특정 참조선 값을 지정할 수 있습니다. 추가 참조선 값을 지정하려면 다른 열 추가를 클릭하십시오.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

맵 차트

맵 차트는 일반적으로 값을 비교하고 지역 전체의 범주를 표시하는 데 사용되며, 데이터에 지리 정보(국가, 지역, 시/도, County, 우편번호 등)가 포함되어 있는 경우에 특히 유용합니다.

단순 맵 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 맵 아이콘을 클릭하십시오.



캔버스가 업데이트되어 맵 차트 템플리트가 표시됩니다.

2. 맵 서비스 드롭 다운 목록에서 맵 이미지를 제공하는 데 사용할 서비스를 선택하십시오. 이 목록은 특정 세계 지역을 포함하는 옵션을 제공합니다.
3. 드롭 다운 메뉴에서 차트 유형을 선택하십시오. 선택한 차트 유형에 따라 다음 옵션이 사용 가능합니다.

경도

드롭 다운 목록에서 경도 값으로 사용할 변수를 선택하십시오.

위도

드롭 다운 목록에서 위도 값으로 사용할 변수를 선택하십시오.

그룹

드롭 다운 메뉴에서 데이터 점 위치를 그룹화하는 변수를 선택하십시오.

범주

시각화할 열 변수를 선택하십시오.

요약

통계 요약함수를 선택하십시오. 각 범주를 요약하는 데 사용되는 방법입니다.

두 가지 유형의 통계 요약함수가 있습니다. 값 변수를 지정해야 하는지 여부가 판별되기 때문에 구분하는 것이 중요합니다.

- **값 변수가 필요하지 않은 함수:** 변수가 필요하지 않은 함수입니다. 모든 개수 및 퍼센트 통계가 이 범주에 속합니다. 이 통계는 정의된 값 변수가 없을 때 사용할 수 있습니다.
- **값 변수가 필요한 함수:** 이 함수는 값 변수가 필요한 함수입니다. 예를 들어 평균 함수의 경우 평균을 계산할 때 변수가 필요합니다. 이 통계는 정의된 값 변수가 있을 때 사용할 수 있습니다.

값

이 필드는 값 변수가 필요한 **요약** 함수를 선택한 경우에 표시됩니다. 값으로 사용할 변수를 선택하십시오.

추가 기능

맵 서비스

맵 이미지를 제공하는 데 사용할 수 있는 서비스가 나열됩니다. [74 페이지의 『맵 서비스 옵션』](#)의 내용을 참조하십시오.

유형

데이터를 표시하는 데 사용할 수 있는 차트 유형이 나열됩니다.

경도

경도 값으로 사용할 수 있는 변수가 나열됩니다.

위도

위도 값으로 사용할 수 있는 변수가 나열됩니다.

그룹

데이터 점 위치를 그룹화하는 데 사용할 수 있는 변수를 나열합니다.

범주

열 변수를 나열합니다.

요약

통계 요약함수를 선택하십시오. 각 범주를 요약하는 데 사용되는 방법입니다.

두 가지 유형의 통계 요약함수가 있습니다. 값 변수를 지정해야 하는지 여부가 판별되기 때문에 구분하는 것이 중요합니다.

- **값 변수가 필요하지 않은 함수:** 변수가 필요하지 않은 함수입니다. 모든 개수 및 퍼센트 통계가 이 범주에 속합니다. 이 통계는 정의된 값 변수가 없을 때 사용할 수 있습니다.
- **값 변수가 필요한 함수:** 이 함수는 값 변수가 필요한 함수입니다. 예를 들어 평균 함수의 경우 평균을 계산할 때 변수가 필요합니다. 이 통계는 정의된 값 변수가 있을 때 사용할 수 있습니다.

값

이 필드는 값 변수가 필요한 **요약** 함수를 선택한 경우에 표시됩니다. 값으로 사용할 변수를 선택하십시오.

도구 팁 정보

데이터 점 위로 마우스를 이동할 때 표시되는 도구 팁 정보를 생성하는 데 사용할 수 있는 변수가 나열됩니다.

크기 맵

사용 가능한 크기 맵 변수가 나열됩니다. 이 변수는 서로 다른 크기를 사용하여 플롯 점에 자신을 표시합니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

맵 서비스 옵션

맵 차트의 세부사항 분할창에서 맵 이미지를 제공하는 데 사용할 맵 서비스를 선택할 수 있습니다.

맵 서비스 옵션 구성

1. IBM SPSS Modeler Server를 중지하십시오.
2. geojson 형식의 로컬 맵 파일을 적용하려면 <Modeler_installation_directory>/dataview/conf/public/mapfiles 디렉토리에 저장하십시오. 디렉토리가 없으면 작성하십시오.
3. 텍스트 편집기에서 <Modeler_installation_directory>/dataview/conf/application.conf 파일을 열고 map.resources 섹션에 다음 옵션을 추가하십시오. 이 예에서는 World - Local Map과 World - Online Map이라는 두 개의 새로운 맵이 추가됩니다.

```
map.resources {  
    "mapservices" :  
    [  
        {  
            "type"      : "geojson",  
            "name"      : "World (Bulit-in)",  
            "location"  : "built_in_world",  
            "id"        : "built_in_world"  
        },  
        {  
            "type"      : "geojson",  
            "name"      : "Blank (Built-in)",  
            "location"  : "built_in_blank",  
            "id"        : "built_in_blank"  
        },  
        {  
            "type"      : "geojson",  
            "name"      : "World - Local Map",  
            "location"  : "/mapfiles/world.json",  
            "useProxy"  : "false",  
            "id"        : "world_local"  
        },  
        {  
            "type"      : "geojson",  
            "name"      : "World - Online Map",  
            "location"  : "http://map.example.com/map/world.json",  
            "useProxy"  : "true",  
            "id"        : "world_online"  
        }  
    ]  
}
```

다음 모수를 맵 옵션에 사용할 수 있습니다.

표 4. 맵 서비스 모수

옵션	값	유형	필요 여부	설명
type	geojson	문자열	예	현재 이 옵션에는 geojson 값만 지원됩니다.

표 4. 맵 서비스 모수 (계속)

옵션	값	유형	필요 여부	설명
이름	맵 이름	문자열	예	맵의 이름을 입력하십시오. 이 이름은 맵 차트의 세부사항 분할창에 있는 맵 서비스 드롭 다운에 표시됩니다.
location	맵 위치	문자열	예	맵의 URL 또는 로컬 파일 위치를 입력하십시오. URL의 경우, <code>http://</code> 또는 <code>https://</code> 로 시작해야 합니다. 로컬 파일의 경우, 경로 /mapfiles/로 시작해야 합니다.
useProxy	true 또는 false	문자열	예	맵 위치가 URL일 경우 <code>true</code> 를 사용하고, 맵 위치가 로컬 파일일 경우 <code>false</code> 를 사용하십시오.
id	맵 ID	문자열	예	맵마다 고유한 ID를 사용하십시오.

참고:

- 기본 **World**(기본 제공) 및 **Blank**(기본 제공) 옵션을 제거하지 마십시오.
- 예에 표시된 대로 내용 블록을 쉼표로 구분하고 새 `map.resources` 섹션이 유효한 JSON 형식의 내용인지 확인하십시오.
- IBM SPSS Modeler Server(또는 서버가 없는 독립형 SPSS Modeler 클라이언트)에서 맵 지도에 액세스할 수 있어야 합니다.

4. IBM SPSS Modeler Server를 시작하십시오.

이제 맵 차트에 대해 작업하는 동안 세부사항 분할창에 사용자가 추가한 새로운 맵 서비스가 표시됩니다.

수학 곡선 차트

수학 곡선 차트는 사용자가 입력한 표현식에 따라 수학 방정식 곡선을 도표화합니다.

단순 수학 곡선 차트 작성

- 차트 작성기의 차트 유형 섹션에서 수학 곡선 아이콘을 클릭하십시오.

Math curve



캔버스가 업데이트되어 수학 곡선 차트 템플리트가 표시됩니다.

- X 값 시작 문자 필드에 X축의 시작값을 입력합니다.
- X 값 끝 문자 필드에 X축의 종료값을 입력합니다.
- 방정식 필드에 그래프 곡선을 도표화하는 방정식을 입력합니다. 추가 방정식을 포함시키려면 다른 열 추가를 클릭하십시오.

각 방정식은 x를 독립변수로 처리합니다. 방정식은 다음을 허용합니다.

- +, -, *, /, % 및 ^
- abs(x)
- ceil(x)
- floor(x)
- log(x)
- max(a,b,c...)
- min(a,b,c...)
- random()
- round(x)
- sqrt(x)
- sin
- cos
- exp
- tan
- atan
- atan2
- asin
- acos

추가 기능

X 값 시작 문자

X축 시작값입니다.

X 값 끝 문자

X축 종료값입니다.

방정식

그래프 곡선을 도표화하는 사용자 입력 방정식입니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

다중 차트

다중 차트는 여러 도표를 작성하는 옵션을 제공합니다. 차트는 동일하거나 서로 다른 유형일 수 있으며 동일한 데 이터 세트의 여러 변수를 포함할 수 있습니다.

단순 다중 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 **다중 차트** 아이콘을 클릭하십시오.

Multi-chart



캔버스가 업데이트되어 다중 차트 템플리트가 표시됩니다.

2. 서브차트를 추가하려면 **다른 서브차트 추가**를 클릭하십시오.

3. 유형 드롭 다운 목록에서 차트 유형을 선택하십시오.
4. 선택한 차트 유형에 따라 다음 옵션이 사용 가능합니다.

막대 및 원형 차트

드롭 다운 목록에서 범주 변수를 선택하십시오.

선 및 산점도 차트

드롭 다운 목록에서 X축 변수를 선택하십시오.

드롭 다운 목록에서 Y축 변수를 선택하십시오.

5. 추가 차트 유형을 포함시키려면 다른 서브차트 추가를 클릭하십시오.

추가 기능

제목

서브차트 제목입니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

다중 계열 차트

다중 계열 차트는 선형 차트와 유사하지만, y축에 변수를 여러 개 작성할 수 있다는 점이 다릅니다.

단순 다중 계열 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 다중 계열 아이콘을 클릭하십시오.



캔버스가 업데이트되어 다중 계열 차트 템플리트가 표시됩니다.

2. 변수를 X축 변수로 선택하십시오.
3. 최소 두 개의 척도변수를 Y축 변수로 선택하십시오.

추가 기능

X축

차트의 x축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

Y축

차트의 y축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

드롭 다운 목록에서 차트 유형(선, 막대 또는 산점도)을 선택하십시오.

차트에 열을 더 포함시키려면 다른 열 추가를 클릭하십시오.

계열 유형

Y축 방향을 정의하기 위한 옵션을 제공합니다. 사용 가능한 옵션은 다음과 같습니다.

- 기본값
- Y축 분리
- 보조 Y축 표시
- 이중 Y축

데이터 표준화

사용으로 설정할 경우 이 설정은 다중 데이터 세트 또는 다중 열의 데이터를 비교할 수 있도록 데이터를 정규 분포로 변환합니다. 이 설정은 빈도 및 변환 통계에 대한 100% 수직누적을 퍼센트로 작성합니다.

보조 Y축

- 차트의 보조 y축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.
- 드롭 다운 목록에서 차트 유형(선, 막대 또는 산점도)을 선택하십시오.
- 차트에 열을 더 포함시키려면 다른 열 추가를 클릭하십시오.

기본 제목

- 차트 제목입니다.

부제목

- 차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

- 차트 아래에 표시되는 차트 꼬리말입니다.

평형 차트

평형 차트는 고차원 기하학을 시각화하고 다변량 데이터를 분석할 때 유용합니다. 평형 차트는 시계열 데이터에 대한 선형 차트와 유사하지만, 축이 시점에 해당하지 않습니다(자연 순서가 없음).

단순 평형 차트 작성

- 차트 작성기의 차트 유형 섹션에서 평형 아이콘을 클릭하십시오.



캔버스가 업데이트되어 평형 차트 템플리트가 표시됩니다.

- 최소 두 개의 변수를 열 변수로 선택하십시오. 각 필드는 차트의 평형 수직축을 나타냅니다.

참고: 열 순서는 기능을 찾는 데 중요합니다. 일반 데이터 분석에서는 열을 여러 번 다시 정렬해야 할 수 있습니다.

추가 기능

열

차트의 y축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

열을 더 추가하려면 다른 열 추가를 클릭하십시오.

색상 맵

사용 가능한 색상 맵 변수가 나열됩니다. 이 변수는 색상 진행(지정된 열의 값 범위를 기준으로 함)을 사용하여 플롯 점에 자신을 표시합니다. 색상 맵은 코로플레스 맵이라고도 합니다.

기본 제목

- 차트 제목입니다.

부제목

- 차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

- 차트 아래에 표시되는 차트 꼬리말입니다.

원형 차트

원형 차트는 비율을 비교하는 데 유용합니다. 예를 들어 원형 차트를 사용하여 높은 비율의 여성이 특정 클래스에 등록했음을 설명할 수 있습니다.

단순 원형 차트 작성

- 차트 작성기의 차트 유형 섹션에서 원형 차트 아이콘을 클릭하십시오.



캔버스가 업데이트되어 원형 차트 템플리트가 표시됩니다.

- 범주 목록에서 범주형(명목 또는 순서) 변수를 선택하십시오. 이 변수의 범주에 따라 원형 차트의 조각 수가 결정됩니다.

3. 그래픽 요소에 대한 통계 요약함수를 선택하십시오. 원형 차트의 경우 일반적으로 개수 기반 통계 또는 합계를 원할 때 사용됩니다. 통계의 결과에 따라 각 조각의 크기가 결정됩니다.

추가 기능

범주

원형 차트의 조각 수를 결정하는 범주형(명목 또는 순서) 변수를 선택하십시오.

요약

그래픽 요소에 대한 통계 요약함수를 선택하십시오. 원형 차트의 경우 일반적으로 개수 기반 통계 또는 합계를 원할 때 사용됩니다. 통계의 결과에 따라 각 조각의 크기가 결정됩니다.

두 가지 유형의 통계 요약함수가 있습니다. **값** 변수를 지정해야 하는지 여부가 판별되기 때문에 구분하는 것이 중요합니다.

- **값 변수가 필요하지 않은 함수:** 변수가 필요하지 않은 함수입니다. 모든 개수 및 퍼센트 통계가 이 범주에 속합니다. 이 통계는 정의된 값 변수가 없을 때 사용할 수 있습니다.
- **값 변수가 필요한 함수:** 이 함수는 값 변수가 필요한 함수입니다. 예를 들어 평균 함수의 경우 평균을 계산할 때 변수가 필요합니다. 이 통계는 정의된 값 변수가 있을 때 사용할 수 있습니다.

값

이 필드는 척도변수가 필요한 **요약** 함수를 선택한 경우에 표시됩니다. 척도변수로 사용할 변수를 선택하십시오.

원 유형

사용 가능한 유형은 다음과 같습니다.

정규

원 세그먼트가 정규 조각으로 표시됩니다.

링

원 세그먼트가 링으로 표시됩니다. 이 유형은 도넛 차트라고도 합니다.

Rose

일반 반경을 사용하는 정규 원형 차트와 달리, 원 세그먼트의 크기가 값에 따라 달라집니다.

장미 영역

일반 반경을 사용하는 정규 원형 차트와 달리, 원 세그먼트의 크기가 영역에 따라 달라집니다.

장미 고리

일반 반경을 사용하는 정규 원형 차트와 달리, 원 세그먼트의 크기가 링에 따라 달라집니다.

반 장미

도표가 원도표의 절반으로 표시되는 점을 제외하고, 장미와 동일합니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

인구 피라미드 차트

인구 피라미드 차트("연령-성별 피라미드"라고도 함)는 일반적으로 연령과 성별을 기준으로 인구 정보를 표시하고 분석하는 데 사용됩니다.

단순 인구 피라미드 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 인구 피라미드 아이콘을 클릭하십시오.



캔버스가 업데이트되어 인구 피라미드 차트 템플릿이 표시됩니다.

2. 드롭 다운 목록에서 **Y축** 변수를 선택하십시오.

3. 드롭 다운 목록에서 **분할 기준** 변수를 선택하십시오.

추가 기능

Y축

차트의 *y*축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

분할 기준

분할 기준 변수에 범주별 셀이 있는 차트 테이블을 작성하는 범주형 변수를 선택하십시오. 그룹화와 마찬가지로, 분할 기준 변수는 기본적으로 변수 범주 각각에 대한 정보를 표시하여 차트에 차원을 더 추가합니다.

구간 너비

데이터를 그룹으로 분할하는 데 사용되는 간격의 크기를 슬라이더로 제어할 수 있습니다.

분포 곡선 표시

사용으로 설정할 경우 분포 적합 곡선이 차트에 표시됩니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

Q-Q 도표

Q-Q(quantile-quantile) 도표는 두 확률 분포에 대해 분위수를 도표화하여 두 확률 분포를 비교합니다. Q-Q 도표는 분포 형태를 비교하는 데 사용되며, 두 개의 분포에서 위치, 척도, 왜도 등의 특성이 얼마나 유사한지 또는 다른지를 그래픽으로 보여줍니다.

단순 Q-Q 도표 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 **Q-Q 도표** 아이콘을 클릭하십시오.



캔버스가 업데이트되어 Q-Q 도표 차트 템플릿이 표시됩니다.

2. 변수를 **X축** 변수로 선택하십시오.

추가 기능

X축

차트의 *x*축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

분포

드롭 다운 목록에 사용 가능한 모든 분포 방법이 나열됩니다.

분포 자동 적합

기본 설정입니다.

베타

지정한 형태 모수를 사용하여 베타 분포의 값을 반환합니다.

지수

지수 분포에서 값을 리턴합니다.

감마

지정한 형태와 척도 모수를 사용하여 감마 분포의 값을 반환합니다.

로그정규

지정한 모수를 사용하여 로그-정규 분포의 값을 반환합니다.

정규

지정한 평균과 표준 편차를 사용하여 정규 분포의 값을 반환합니다.

균일

최소값과 최대값 간 균일 분포의 값을 반환합니다.

스튜던트 T

지정한 자유도를 사용하여 스튜던트 T 분포의 값을 반환합니다.

도표 유형

Q-Q(quantile-quantile) 도표 또는 P-P(percent-percent) 도표를 선택하십시오.

자동 적합

사용으로 설정할 경우 선택한 분포에 대한 데이터 모수가 추정됩니다. 사용으로 설정할 경우 **모양** 및 **척도** 분포 값이 표시됩니다.

형태1

베타 분포의 형태1 값을 설정합니다. 이 설정은 자동 적합이 사용으로 설정되어 있지 않고 **베타가 분포로 선택된** 경우에만 사용할 수 있습니다.

형태2

베타 분포의 형태2 값을 설정합니다. 이 설정은 자동 적합이 사용으로 설정되어 있지 않고 **베타가 분포로 선택된** 경우에만 사용할 수 있습니다.

형태

선택한 분포의 형태 값을 설정합니다. 이 설정은 자동 적합이 사용으로 설정되어 있지 않고 **감마** 또는 **로그정규가 분포로 선택된** 경우에만 사용할 수 있습니다.

척도

선택한 분포의 척도 값을 설정합니다. 이 설정은 자동 적합이 사용으로 설정되어 있지 않고 **지수**, **감마** 또는 **로그정규가 분포로 선택된** 경우에만 사용할 수 있습니다.

평균

정규 분포의 평균 값을 설정합니다. 이 설정은 자동 적합이 사용으로 설정되어 있지 않고 **정규가 분포로 선택된** 경우에만 사용할 수 있습니다.

표준 편차

정규 분포의 표준 편차 값을 설정합니다. 이 설정은 **정규가 분포로 선택된** 경우에만 사용할 수 있습니다.

최소값

균일 분포의 최소값을 설정합니다. 이 설정은 자동 적합이 사용으로 설정되어 있지 않고 **균일이 분포로 선택된** 경우에만 사용할 수 있습니다.

최대값

균일 분포의 최대값을 설정합니다. 이 설정은 자동 적합이 사용으로 설정되어 있지 않고 **균일이 분포로 선택된** 경우에만 사용할 수 있습니다.

자유도(df)

스튜던트 t 분포에 대한 자유도 값을 설정합니다. 이 설정은 자동 적합이 사용으로 설정되어 있지 않고 **스튜던트 T가 분포로 선택된** 경우에만 사용할 수 있습니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

방사형 차트

방사형 차트는 여러 양적 변수를 비교하고, 값이 비슷한 변수를 시각화하거나 변수 간에 이상치가 있는지 표시하는 데 유용합니다. 방사형 차트는 일련의 스포크로 구성되며, 각 스포크는 단일 변수를 나타냅니다. 방사형 차트는 데이터 세트에서 점수가 높고 낮은 변수를 판별하는 데도 유용합니다.

단순 방사형 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 **방사형** 아이콘을 클릭하십시오.



캔버스가 업데이트되어 방사형 차트 템플리트가 표시됩니다.

2. 드롭 다운 목록에서 **열** 변수를 선택하십시오.

참고: 추가 열을 포함시키려면 다른 열 추가를 클릭하십시오. 세 개 이상의 열 변수를 정의해야 합니다.

추가 기능

범주

범주형(명목 또는 순서) 변수를 선택하십시오. 범주 값으로 **없음**을 선택하면 모든 값이 개별적으로 표시되며 요약 방법도 적용되지 않습니다.

요약

범주형 변수를 선택하면, 다음 요약 통계량이 사용 가능합니다.

수

총 케이스 수입니다.

합계

값의 합계입니다.

평균

합계를 케이스 수로 나눈 산술 평균 값입니다.

최대값

가장 큰(가장 높은) 값입니다.

최소값

가장 작은(가장 낮은) 값입니다.

방사형 레이아웃

방사형 차트의 배경 이미지 레이아웃을 판별합니다.

원형

선택할 경우 방사형 차트가 원형 레이아웃 위에 표시됩니다.

다각형

선택할 경우 방사형 차트가 다각형 레이아웃 위에 표시됩니다.

분할 기준

분할 기준 변수에 범주별 셀이 있는 차트 테이블을 작성하는 범주형 변수를 선택하십시오. 그룹화와 마찬가지로, 분할 기준 변수는 기본적으로 변수 범주 각각에 대한 정보를 표시하여 차트에 차원을 더 추가합니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

관계 차트

관계 차트는 변수의 상관 정도를 판별하는 데 유용합니다.

단순 관계 차트 작성

1. 차트 작성기의 **차트 유형** 섹션에서 관계 아이콘을 클릭하십시오.



캔버스가 업데이트되어 관계 차트 템플리트가 표시됩니다.

2. 최소 두 개의 변수를 **열** 변수로 선택하십시오.

참고: 열 변수를 더 포함시키려면 다른 열 추가를 클릭하십시오.

추가 기능

열

사용 가능한 데이터 세트 변수가 나열됩니다.

열을 더 추가하려면 다른 열 추가를 클릭하십시오.

선 유형

관련 데이터 점을 서로 연결하는 선 유형을 제어합니다.

곡선

선택할 경우 관련 데이터 점이 곡선으로 연결됩니다.

직선

선택할 경우 관련 데이터 포인트가 직선으로 연결됩니다.

레이블 임계값

정의된 값을 초과하는 값을 갖는 데이터 점의 레이블을 표시합니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

산점도 및 점도표

포인트 그래픽 요소로 작성된 여러 광범위한 범주의 차트가 있습니다.

- 산점도:** 다변량 데이터를 도표화하는 데 유용하며, 척도변수 간에 잠재적 관계를 판별하는 데 도움이 됩니다. 단순 산점도는 2차원 좌표계를 사용하여 두 개의 변수를 도표화합니다. 3차원 산점도는 3차원 좌표계를 사용하여 세 개의 변수를 도표화합니다. 추가 변수를 도표화해야 하는 경우, 오버레이 산점도 및 산점도 행렬 (SPLOM)을 시도할 수 있습니다. 오버레이 산점도는 각 대응 변수를 색상 또는 형태별로 구분하여 x-y 변수의 오버레이 대응을 표시합니다. SPLOM은 각 변수를 SPLOM의 다른 각 변수에 대해 도표화한 2차원 산점도 행렬을 작성합니다.
- 점도표:** 히스토그램과 마찬가지로, 단일 척도변수의 분포를 표시하는 데 유용합니다. 데이터가 구간화되지만, 구간당 하나의 값(예: 개수) 대신 각 구간의 모든 포인트가 표시되고 누적됩니다. 이 그래프는 밀도 도표라고도 합니다.
- 요약 포인트 도표:** 이 차트는 포인트가 막대 맨 위에 표시되는 것을 제외하고 막대형 차트와 거의 비슷합니다. 요약 포인트 도표는 막대형 차트와 매우 비슷하므로, 이 도표 작성에 대한 정보는 [60 페이지의 『막대형 차트』](#)의 내용을 참조하십시오.
- 하락-선 차트:** 이 차트는 특별한 유형의 요약 포인트 도표입니다. 포인트는 그룹화되고 선은 각 범주 내 포인트를 통해 작성됩니다. 하락-선 차트는 범주형 변수에 대한 통계를 비교하는 데 유용합니다.

단순 산점도 작성

- 차트 작성기의 차트 유형 섹션에서 산점도 아이콘을 클릭하십시오.



캔버스가 업데이트되어 산점도 차트 템플릿이 표시됩니다.

- 척도변수를 **X축** 변수로 선택하십시오.

- 척도변수를 **Y축** 변수로 선택하십시오. 산점도는 일반적으로 원시 값을 표시하기 때문에 통계를 지정할 필요가 없습니다.

추가 기능

X축

차트의 x축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

Y축

차트의 y축에 사용할 수 있는 데이터 세트 변수가 나열됩니다.

색상 맵

사용 가능한 색상 맵 변수가 나열됩니다. 이 변수는 색상 진행(지정된 열의 값 범위를 기준으로 함)을 사용하여 플롯 점에 자신을 표시합니다. 색상 맵은 코로플레스 맵이라고도 합니다.

크기 맵

사용 가능한 크기 맵 변수가 나열됩니다. 이 변수는 서로 다른 크기를 사용하여 플롯 점에 자신을 표시합니다.

모양 맵

사용 가능한 모양 맵 변수가 나열됩니다. 이 변수는 서로 다른 모양을 사용하여 플롯 점에 자신을 표시합니다.

회귀선 적합

회귀선 적합에서는 데이터 점이 하나의 선에 맞춰지는데 이 선이 데이터 점을 모두 통과하는 것은 아닙니다. 회귀선 적합은 데이터의 추세를 나타냅니다. 일부 회귀선 적합은 회귀분석을 기반으로 하고 다른 회귀선 적

합은 반복 가중치를 부여한 최소제곱을 기반으로 합니다. 드롭 다운 목록에서 회귀선 적합 옵션을 선택하십시오.

기울기 거품

이 전환 컨트롤은 거품형 차트에서 색상 기울기 및 3D 효과를 사용/사용 안함으로 설정합니다. 색상 맵 변수를 선택한 경우에는 이 설정이 사용 안 함으로 설정됩니다.

최소 거품 크기

최소 거품 크기를 설정합니다. 5와 20 사이의 값을 입력합니다.

최대 거품 크기

최대 거품 크기를 설정합니다. 20과 80 사이의 값을 입력합니다.

참조선 표시

사용으로 설정할 경우, X축 및 Y축 값을 기준으로 하는 차트에 참조선이 표시됩니다.

X축에 참조선 값 입력

참조선 표시가 사용으로 설정된 경우 이 설정을 사용하여 X축에 대한 특정 참조선 값을 지정할 수 있습니다. 추가 참조선 값을 지정하려면 다른 열 추가를 클릭하십시오.

Y축에 참조선 값 입력

참조선 표시가 사용으로 설정된 경우 이 설정을 사용하여 Y축에 대한 특정 참조선 값을 지정할 수 있습니다. 추가 참조선 값을 지정하려면 다른 열 추가를 클릭하십시오.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

산점도 행렬 차트

산점도 행렬은 다중 변수 간에 선형 상관관계가 있는지 판별하는 좋은 방법입니다.

산점도 행렬 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 산점도 행렬 아이콘을 클릭하십시오.



캔버스가 업데이트되어 산점도 행렬 차트 템플리트가 표시됩니다.

2. 최소 두 개 이상의 열 척도변수를 선택하십시오.

참고: 열 변수를 더 포함시키려면 다른 열 추가를 클릭하십시오.

선택한 각 변수가 다른 변수 각각에 대해 도표화되어 개별 산점도 행렬이 작성됩니다.

추가 기능

열

행렬 변수를 최소 두 개 선택하십시오. 변수는 숫자여야 합니다(날짜 형식 제외).

열을 더 추가하려면 다른 열 추가를 클릭하십시오.

색상 맵

사용 가능한 색상 맵 변수가 나열됩니다. 이 변수는 색상 진행(지정된 열의 값 범위를 기준으로 함)을 사용하여 플롯 점에 자신을 표시합니다. 색상 맵은 코로플레스 맵이라고도 합니다.

상관관계

사용으로 설정할 경우 선택한 변수에 대해 선형 상관관계 정보(강함, 중간, 약함)가 표시됩니다.

KDE 곡선 표시

사용으로 설정할 경우, 커널 밀도 추정값 곡선이 차트에 표시됩니다.

기울기 거품

이 전환 컨트롤은 거품형 차트에서 색상 기울기 및 3D 효과를 사용/사용 안함으로 설정합니다. 색상 맵 변수를 선택한 경우에는 이 설정이 사용 안 함으로 설정됩니다.

최소 거품 크기

최소 거품 크기를 설정합니다. 5와 20 사이의 값을 입력합니다.

최대 거품 크기

최대 거품 크기를 설정합니다. 20과 80 사이의 값을 입력합니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

햇살 차트

햇살 차트는 계층형 데이터 구조를 시각화할 때 유용합니다. 햇살 차트는 내부 원과 이를 둘러싸고 있는 더 깊은 계층 수준 고리로 구성됩니다. 각 세그먼트의 각도는 값에 비례하거나 내부 세그먼트 아래에서 동일하게 나뉩니다. 차트 세그먼트는 속해 있는 범주 또는 계층 구조 수준에 따라 색상이 지정됩니다.

단순 햇살 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 햇살 아이콘을 클릭하십시오.

Sunburst



캔버스가 업데이트되어 햇살 차트 템플리트가 표시됩니다.

2. 열 목록에서 범주형(명목 또는 순서) 변수를 선택하십시오. 이 변수의 범주에 따라 차트의 세그먼트 수가 결정됩니다.

3. 다른 열 추가를 클릭하고 열 목록에서 범주형(명목 또는 순서) 변수를 선택하십시오. 이 변수의 범주에 따라 차트 두 번째 고리의 세그먼트 수가 결정되며, 이는 더 높은 계층 구조 수준을 나타냅니다.

4. 그래픽 요소(개수 기반 통계 또는 합계)에 대한 통계 요약함수를 선택하십시오. 통계의 결과에 따라 각 세그먼트의 크기가 결정됩니다. 합계를 선택한 경우, 요약할 데이터 세트의 값을 나타낼 척도변수를 값 목록에서 선택하십시오.

5. 햇살 레이아웃 옵션을 기준 또는 분산형 중에서 선택하십시오.

추가 기능

열

차트의 세그먼트 수를 결정하는 범주형(명목 또는 순서) 변수를 선택하십시오.

요약

그래픽 요소(개수 기반 통계 또는 합계)에 대한 통계 요약함수를 선택하십시오. 통계의 결과에 따라 각 조각의 크기가 결정됩니다.

두 가지 유형의 통계 요약함수가 있습니다. **값** 변수를 지정해야 하는지 여부가 판별되기 때문에 구분하는 것이 중요합니다.

- **값 변수가 필요하지 않은 함수:** 이 함수는 **값** 변수가 필요하지 않은 함수입니다. 모든 개수 및 퍼센트 통계 가 이 범주에 속합니다. 이 통계는 정의된 **값** 변수가 없을 때 사용할 수 있습니다.
- **값 변수가 필요한 함수:** 이 함수는 **값** 변수가 필요한 함수입니다. 예를 들어, 합계 함수의 경우 합계를 계산 할 때 변수가 필요합니다.

값

이 필드는 척도변수가 필요한 **요약** 함수를 선택한 경우에 표시됩니다. 척도변수로 사용할 변수를 선택하십시오.

햇살 레이아웃

사용 가능한 옵션은 **기준** 및 **분기**입니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

시간 도표

시간 도표에서는 연속된 시간 구간으로 된 데이터 점을 보여줍니다. 구성하는 시계열은 숫자 값을 포함해야 하며 주기가 균일한 시간 범위에서 발생한다고 가정합니다. 시간 도표에서는 기본 통계량과 검정에서 시계열 데이터의 공정특성 변수를 예비 분석하므로, 모델링하기 전에 데이터에 관한 유용한 정보가 생성됩니다. 시간 도표에는 분해, ADF(Augmented Dickey-Fuller) 검정, 상관계수(ACF/PACF) 및 스펙트럼 분석과 같은 분석 방법이 포함됩니다.

단순 시간 도표 작성

1. 차트 작성기의 차트 유형 섹션에서 **시간 도표** 아이콘을 클릭하십시오.

Time plot



캔버스가 업데이트되어 시간 도표 차트 템플리트가 표시됩니다.

2. **값** 드롭 다운에서 y축 값을 선택하십시오.

추가 기능

날짜

해당되는 경우, 드롭 다운 목록에서 날짜를 선택하십시오. 각 관측값은 기본적으로 동일한 시간 구간으로 구분됩니다. 날짜 변수를 선택하는 경우 표본 재추출 옵션이 표시됩니다. 이 옵션을 사용하면 지정된 구간과 일치하는 값 필드를 통합할 수 있습니다.

시간 도표 알고리즘

시계열 데이터를 분석하는 데 사용하는 시간 도표 알고리즘입니다.

- **분해.** 시계열을 세 개의 구성요소(추세-순환, 계절 및 비정규)로 분해합니다. 분해는 가법 방식으로 수행됩니다.
- **ADF 검정.** ADF(Augmented Dickey-Fuller)는 계열에 단위 균이 있으며 계열이 정상이 아니라는 영가설을 검정합니다. 검정 결과에서 영가설을 거부하면 계열이 정상이거나 차분 모델을 사용하여 정상으로 표시될 수 있음을 나타냅니다.
- **ACF/PACF.** 계열의 상관관계입니다.
- **S스펙트럼 분석.** 빈도 도메인의 분석 도구입니다. 최고 빈도는 다이아몬드로 표시됩니다.

도표 위치 바꾸기

평면 측정 직사각형 좌표계 및 극 좌표계에서 도표의 위치를 반전시킵니다.

전환점 표시

도표에서 전환점을 표시하거나 숨깁니다. 시계열 분해의 추세-순환 구성요소를 기반으로 시계열에 전체 추세가 있는지 아니면 추세 패턴의 방향을 변경하는 전환점이 있는지 탐색하십시오.

이상치 표시

이상치를 표시하거나 숨깁니다. 시계열의 이상치는 시계열 분해의 비정규 구성요소에서 분석됩니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

테마 리버 차트

테마 리버는 시간 경과에 따른 변화를 보여주는 특수화된 플로우 그래프입니다.

단순 테마 리버 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 테마 리버 아이콘을 클릭하십시오.

Theme River



캔버스가 업데이트되어 테마 리버 차트 템플리트가 표시됩니다.

2. X축 변수를 선택하십시오.
3. 범주 변수를 선택하십시오.

제한사항: 범주 필드에 50개가 넘는 개별 범주가 있는 경우 상위 50개의 최대 범주만 차트에서 이벤트로 사용됩니다.

추가 기능

순서 기준

선택한 X축 값에 따라, 범주 순서가 범주 이름 또는 범주 값을 기준으로 하는지 여부를 지정할 수 있습니다.

범주 순서

선택한 X축 값에 따라, 범주 순서가 오름차순 또는 내림차순인지 아니면 데이터 세트에서 읽을 것인지 여부를 지정할 수 있습니다.

요약

통계 요약함수를 선택하십시오. 각 범주를 요약하는 데 사용되는 방법입니다.

두 가지 유형의 통계 요약함수가 있습니다. 값 변수를 지정해야 하는지 여부가 판별되기 때문에 구분하는 것이 중요합니다.

- **값 변수가 필요하지 않은 함수:** 변수가 필요하지 않은 함수입니다. 모든 개수 및 퍼센트 통계가 이 범주에 속합니다. 이 통계는 정의된 값 변수가 없을 때 사용할 수 있습니다.
- **값 변수가 필요한 함수:** 이 함수는 값 변수가 필요한 함수입니다. 예를 들어 평균 함수의 경우 평균을 계산 할 때 변수가 필요합니다. 이 통계는 정의된 값 변수가 있을 때 사용할 수 있습니다.

값

이 필드는 값 변수가 필요한 **요약** 함수를 선택한 경우에 표시됩니다. 값으로 사용할 변수를 선택하십시오.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

트리 차트

트리 차트는 트리와 유사한 구조로 계층을 표시합니다. 트리 차트의 구조는 루트 노드(상위 노드 없음), 분기라고 하는 선 연결(멤버 간의 관계와 연결을 나타냄), 리프 노드(하위 노드 없음)로 구성됩니다.

단순 트리 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 트리 아이콘을 클릭하십시오.

Tree



캔버스가 업데이트되어 트리 차트 템플리트가 표시됩니다.

2. 드롭 다운 목록에서 열 변수를 선택하십시오.

참고: 열 변수를 더 포함시키려면 다른 열 추가를 클릭하십시오.

추가 기능

열

차트 열을 표시하는 데 사용할 수 있는 데이터 세트 변수가 나열됩니다.

요약

그래픽 요소에 대한 통계 요약함수를 선택하십시오. 통계의 결과에 따라 그래픽 요소의 위치가 결정됩니다.

두 가지 유형의 통계 요약함수가 있습니다. 값 변수를 지정해야 하는지 여부가 판별되기 때문에 구분하는 것이 중요합니다.

- **값 변수가 필요하지 않은 함수:** 값 변수가 필요하지 않은 함수입니다. 모든 개수 및 퍼센트 통계가 이 범주에 속합니다. 이 통계는 정의된 값 변수가 없을 때 사용할 수 있습니다.
- **값 변수가 필요한 함수:** 이 함수는 값 변수가 필요한 함수입니다. 예를 들어 평균 함수의 경우 평균을 계산 할 때 변수가 필요합니다. 이 통계는 정의된 값 변수가 있을 때 사용할 수 있습니다.

값

이 필드는 값 변수가 필요한 **요약** 함수를 선택한 경우에 표시됩니다. 요약값의 기준으로 사용할 변수를 선택하십시오.

트리 레이아웃

왼쪽에서 오른쪽

루트 노드가 왼쪽에 표시되고 리프 노드는 오른쪽에 표시됩니다.

오른쪽에서 왼쪽

루트 노드가 오른쪽에 표시되고 리프 노드는 왼쪽에 표시됩니다.

위쪽에서 아래쪽

루트 노드가 위쪽에 표시되고 리프 노드는 아래쪽에 표시됩니다.

아래쪽에서 위쪽

루트 노드가 아래쪽에 표시되고 리프 노드는 위쪽에 표시됩니다.

방사형

루트 노드가 가운데 표시되고 리프 노드는 루트에서 방사형으로 뻗어나갑니다.

리프 깊이

리프 노드에 대한 드릴다운 수준 값을 설정합니다.

리프 레이블 표시

사용으로 설정할 경우 레이블이 리프 노드마다 표시됩니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

트리 맵 차트

트리 맵 차트는 트리 다이어그램의 계층 구조를 시각화하는 대체 방법이며, 범주별로 수량도 표시합니다. 트리 맵 차트는 특히 데이터의 패턴을 식별하는데 유용합니다. 트리 분기는 직사각형으로 표시되며, 각 하위 분기는 더 작은 직사각형으로 표시됩니다.

단순 트리 맵 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 트리 맵 아이콘을 클릭하십시오.



캔버스가 업데이트되어 트리 맵 차트 템플릿이 표시됩니다.

2. 드롭 다운 목록에서 열 변수를 선택하십시오.

참고: 열 변수를 더 포함시키려면 다른 열 추가를 클릭하십시오.

추가 기능

열

차트 열을 표시하는 데 사용할 수 있는 데이터 세트 변수가 나열됩니다.

요약

그래픽 요소에 대한 통계 요약함수를 선택하십시오. 통계의 결과에 따라 그래픽 요소의 위치가 결정됩니다.

두 가지 유형의 통계 요약함수가 있습니다. 값 변수를 지정해야 하는지 여부가 판별되기 때문에 구분하는 것이 중요합니다.

- **값 변수가 필요하지 않은 함수:** 값 변수가 필요하지 않은 함수입니다. 모든 개수 및 퍼센트 통계가 이 범주에 속합니다. 이 통계는 정의된 값 변수가 없을 때 사용할 수 있습니다.
- **값 변수가 필요한 함수:** 이 함수는 값 변수가 필요한 함수입니다. 예를 들어 평균 함수의 경우 평균을 계산할 때 변수가 필요합니다. 이 통계는 정의된 값 변수가 있을 때 사용할 수 있습니다.

값

이 필드는 값 변수가 필요한 **요약** 함수를 선택한 경우에 표시됩니다. 요약값의 기준으로 사용할 변수를 선택하십시오.

리프 깊이

리프 노드에 대한 드릴다운 수준 값을 설정합니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

t-SNE 차트

t-SNE(T-distributed Stochastic Neighbor Embedding)는 시각화를 위한 머신 학습 알고리즘입니다. t-SNE 차트는 유사한 오브젝트는 근접 지점 옆에 모델링하고 유사하지 않은 오브젝트는 높은 확률의 먼 지점에 모델링하는 방식으로, 고차원 오브젝트 각각을 2차원 또는 3차원 지점 옆에 모델링합니다.

단순 t-SNE 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 **t-SNE** 아이콘을 클릭하십시오.



캔버스가 업데이트되어 t-SNE 차트 템플리트가 표시됩니다.

2. 혼란, 학습률, 최대 반복 값을 설정하십시오.
3. 선택적으로 색상 맵 변수를 선택하십시오.

추가 기능

혼란

각 데이터 점의 근접 이웃 수에 대해 경험적 추측을 구축할 숫자를 설정합니다. 이는 데이터의 로컬 측면과 글로벌 측면의 균형을 맞추기 위한 것입니다.

학습률

이 값은 반복 시마다 가중치 크기 변경사항을 지정하여 학습 속도에 영향을 미칩니다.

최대 반복

수행할 최대 반복 수입니다.

색상 맵

사용 가능한 색상 맵 변수가 나열됩니다. 이 변수는 색상 진행(지정된 열의 값 범위를 기준으로 함)을 사용하여 플롯 점에 자신을 표시합니다. 색상 맵은 코로플레스 맵이라고도 합니다.

기본 제목

차트 제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

단어 클라우드 차트

단어 클라우드 차트는 데이터를 단어로 표시하며, 개별 단어의 크기와 배치는 가중값에 따라 달라집니다.

단순 단어 클라우드 차트 작성

1. 차트 작성기의 차트 유형 섹션에서 단어 클라우드 아이콘을 클릭하십시오.



캔버스가 업데이트되어 단어 클라우드 차트 템플리트가 표시됩니다.

2. 변수를 소스 변수로 선택하십시오. 각 변수 범주가 가중값을 기준으로 차트에 표시됩니다.
3. 차트의 형태 값을 선택하십시오. 생성되는 차트 데이터가 선택한 형태로 표시됩니다.

추가 기능

소스

차트의 소스로 사용할 수 있는 데이터 세트 변수가 나열됩니다. 각 변수 범주가 가중값을 기준으로 차트에 표시됩니다.

형태

사용 가능한 차트 형태가 나열됩니다. 생성되는 차트 데이터가 선택한 형태로 표시됩니다.

기본 제목

차트 제목입니다.

부제목

차트 제목 바로 아래에 표시되는 차트 부제목입니다.

꼬리말

차트 아래에 표시되는 차트 꼬리말입니다.

대시보드

여러 차트를 동시에 볼 수 있는 차트 대시보드 레이아웃을 작성할 수 있습니다. 레이아웃을 템플리트로 저장한 다음 저장한 차트를 레이아웃의 위치로 끌어다 놓으십시오.

- 데이터 노드를 마우스 오른쪽 단추로 클릭하고 **데이터 보기**를 선택한 후 **동작** 섹션에서 대시보드 제어를 클릭하십시오.



그림 14. 대시보드 제어

대시보드는 다음과 같은 설정을 표시하고 제공합니다.

템플리트

이 섹션에서는 레이아웃 템플리트를 새로 작성하거나 사전 정의된 템플리트 또는 저장된 템플리트 중에서 선택할 수 있습니다.

레이아웃 템플리트 선택

사용 가능한 레이아웃 템플리트 중에서 선택하거나 새 템플리트를 사용하여 시작하십시오.

동작

선택한 레이아웃 템플리트를 편집하려면 **레이아웃 편집** 아이콘을 클릭하십시오. 또한 대시보드 파일을 가져올 수도 있습니다. 완료되면, 템플리트를 저장하십시오.

완료된 경우 **레이아웃 편집 모드 나가기** 아이콘을 클릭하십시오.

컨텐츠

이 섹션을 사용하여 항목을 대시보드 레이아웃으로 끌어다 놓을 수 있습니다.

저장된 차트 선택

이 섹션은 저장된 차트 목록을 제공합니다. 대시보드 레이아웃에서 원하는 위치로 차트를 끌어다 놓으십시오.

오브젝트 선택

HTML 텍스트 또는 이미지를 대시보드 레이아웃으로 끌어올 수도 있습니다.

글로벌 시각화 기본 설정

제목, 범위 슬라이더, 격자선 및 마우스 추적에 대한 기본 설정을 대체할 수 있습니다. 다른 색 구성표 템플리트를 지정할 수도 있습니다.

- 데이터 노드를 마우스 오른쪽 단추로 클릭하고 **데이터 보기**를 선택한 후 **동작** 섹션에서 글로벌 시각화 기본 설정 제어를 클릭하십시오.



그림 15. 글로벌 시각화 기본 설정 제어

글로벌 시각화 기본 설정 대화 상자는 다음과 같은 설정을 표시하고 제공합니다.

제목

이 섹션은 글로벌 차트 제목 설정을 제공합니다.

글로벌 제목

모든 차트에 대해 글로벌 제목을 사용 또는 사용 안함으로 설정합니다.

글로벌 1차 제목

글로벌 1차 차트 제목의 표시를 사용 또는 사용 안함으로 설정합니다. 사용으로 설정할 경우 여기에 입력한 최상위 레벨 차트 제목이 모든 차트에 적용되며, 결과적으로 각 차트의 개별 **1차 제목** 설정을 대체합니다.

글로벌 부제목

글로벌 차트 부제목의 표시를 사용 또는 사용 안함으로 설정합니다. 사용으로 설정할 경우 여기에 입력한 차트 부제목이 모든 차트에 적용되며, 결과적으로 각 차트의 개별 **부제목** 설정을 대체합니다.

기본 제목

모든 차트에 대해 기본 제목을 사용 또는 사용 안함으로 설정합니다.

도구

이 섹션은 차트 동작을 제어하는 옵션을 제공합니다.

범위 슬라이더

각 차트에 대해 범위 슬라이더를 사용 또는 사용 안함으로 설정합니다. 사용으로 설정할 경우 각 차트의 아래에 제공되는 범위 슬라이더를 통해 표시되는 차트 데이터의 양을 제어할 수 있습니다.

격자선

X축(세로) 및 Y축(가로) 격자선의 표시를 제어합니다.

마우스 추적

사용으로 설정할 경우 커서를 차트 위의 원하는 위치에 두면 차트 데이터와 관련하여 마우스 커서 위치가 표시됩니다.

도구 상자

각 차트에 대해 도구 상자를 사용 또는 사용 안함으로 설정합니다. 차트 유형에 따라 화면의 오른쪽에 있는 도구 상자는 확대/축소, 이미지로 저장, 복원, 데이터 선택, 선택사항 지우기 등의 도구를 제공합니다.

ARIA

사용으로 설정할 경우, 장애가 있는 사용자가 웹 컨텐츠와 웹 애플리케이션에 보다 쉽게 액세스할 수 있습니다.

널값 필터링

널 차트 데이터 필터링을 사용 또는 사용 안함으로 설정합니다.

X축을 원점에

사용으로 설정할 경우 X축이 다른 원점 위치에 있습니다. 사용 안함으로 설정할 경우 X축이 항상 0에서 시작합니다.

Y축을 원점에

사용으로 설정할 경우 Y축이 다른 원점 위치에 있습니다. 사용 안함으로 설정할 경우 Y축이 항상 0에서 시작합니다.

X축 레이블 표시

X축 레이블을 사용 또는 사용 안함으로 설정합니다.

Y축 레이블 표시

Y축 레이블을 사용 또는 사용 안함으로 설정합니다.

X축 선 표시

X축 선을 사용 또는 사용 안함으로 설정합니다.

Y축 선 표시

Y축 선을 사용 또는 사용 안함으로 설정합니다.

테마

그룹화 또는 스택 변수가 있는 차트에서 사용되는 색상을 변경하려면 템플리트를 선택하십시오. 선택된 템플리트 파일에 정의된 요소 속성은 해당 요소 속성의 기본 템플리트 설정을 대체합니다.

2. 설정을 저장하려면 **적용**을 클릭하고, 변경을 무시하려면 **취소**를 클릭하십시오.

제 7 장 출력 작업

일부 스트림을 실행할 때 모델 너깃 노드의 **모델 탭** 또는 **고급 탭**을 통해 뷰어에서 결과를 사용할 수 있습니다. 뷰어에서 보려는 출력으로 쉽게 이동할 수 있습니다. 출력을 조작하고 원하는 출력을 정확히 포함하는 문서를 작성할 수도 있습니다. 또한 일부 그래프 출력의 경우 뷰어가 사용됩니다.

뷰어는 IBM SPSS Modeler에서 다음 출력에 사용됩니다.

- TCM 모델 너깃
- STP 모델 너깃
- TwoStep-AS 군집 모델 너깃
- GSAR 모델 너깃
- 맵 시각화 그래프 노드

뷰어

결과는 뷰어에 표시됩니다. 뷰어를 사용하여 다음을 수행할 수 있습니다.

- 결과 찾아보기
- 선택한 테이블과 차트를 표시하거나 숨깁니다.
- 선택한 항목을 이동하여 결과의 표시 순서를 변경합니다.
- 항목을 뷰어와 다른 애플리케이션 간에 이동합니다.

뷰어는 두 개의 분할창으로 나뉘어져 있습니다.

- 왼쪽 분할창에는 내용의 개요 보기 있습니다.
- 오른쪽 분할창에는 통계표, 차트, 텍스트 출력이 들어 있습니다.

개요에서 항목을 눌러 해당하는 테이블이나 차트로 바로 이동할 수 있습니다. 개요 분할창의 오른쪽 경계선을 끌어 개요 분할창의 너비를 변경합니다.

결과 표시 및 숨김

뷰어에서 전체 프로시저의 개별 테이블이나 결과를 선택적으로 표시하거나 숨길 수 있습니다. 이 방법은 내용 창에서 볼 수 있는 출력 양을 줄이려고 할 때 유용합니다.

테이블 및 차트 숨김

1. 뷰어의 개요 분할창에서 항목의 책 모양 아이콘을 두 번 클릭하십시오.

또는

2. 항목을 클릭하여 선택하십시오.

3. 메뉴에서 다음을 선택하십시오.

보기 > 숨김

또는

4. 개요 도구 모음에서 닫힌 책 모양(숨김) 아이콘을 클릭하십시오.

열린 책 모양(표시) 아이콘이 활성 아이콘이 되며 해당 항목이 현재 숨겨져 있음을 나타냅니다.

프로시저 결과 숨김

1. 개요 분할창에서 프로시저 이름 왼쪽의 상자를 클릭하십시오.

프로시저의 결과를 모두 숨기고 개요 보기 를 접습니다.

출력 이동, 삭제, 복사

출력을 항목이나 항목 그룹의 복사, 이동, 삭제 등의 방법으로 재정렬합니다.

뷰어에서 출력 이동

1. 개요 또는 내용 창에서 항목을 선택하십시오.
2. 선택한 항목을 다른 위치로 끌어다 놓습니다.

1단계에서 잘라낸 항목이 선택한 위치로 이동합니다.

뷰어에서 출력 삭제

1. 개요 또는 내용 창에서 항목을 선택하십시오.
2. **Delete** 키를 누르십시오.

또는

3. 메뉴에서 다음을 선택하십시오.

편집 > 삭제

초기 맞춤 변경

기본적으로 모든 결과는 처음에는 왼쪽 맞춤으로 표시됩니다. 새로운 출력 항목의 초기 맞춤 변경

1. 메뉴에서 다음을 선택하십시오.

편집 > 옵션

2. 뷰어 탭을 클릭하십시오.
3. 초기 출력결과 상태 그룹에서 항목 유형(예: 피벗 테이블, 차트, 텍스트 출력)을 선택합니다.
4. 원하는 맞춤 옵션을 선택하십시오.

출력 항목의 맞추기 변경

1. 개요 또는 내용 창에서 맞추려는 항목을 선택합니다.
2. 메뉴에서 다음을 선택하십시오.

형식 > 왼쪽 맞춤

또는

형식 > 가운데 맞춤

또는

형식 > 오른쪽 맞춤

뷰어 개요

개요 분할창은 뷰어 문서의 목차를 제공합니다. 개요 분할창을 사용하여 결과를 탐색하고 표시를 제어할 수 있습니다. 개요 분할창의 동작은 대부분 내용 창에 영향을 줍니다.

- 개요 분할창에서 항목을 선택하면 내용 창에서 해당하는 항목도 표시됩니다.
- 개요 분할창에서 항목을 이동하면 내용 창에서 해당하는 항목도 이동됩니다.
- 개요 보기 를 접으면 접혀진 수준의 모든 항목으로부터 결과를 숨깁니다.

개요 표시 제어

개요 표시 형식을 제어하려면 다음을 수행할 수 있습니다.

- 개요 보기 를 펼치거나 접습니다.

- 선택한 항목에 대한 개요 수준을 바꿉니다.
- 개요 표시에서 항목 크기를 바꿉니다.
- 개요 표시에 사용하는 글꼴을 바꿉니다.

개요 보기 접기 및 펼치기

1. 접거나 펼치려는 개요 항목의 왼쪽에 있는 상자를 클릭하십시오.

또는

2. 개요에 있는 항목을 클릭하십시오.

3. 메뉴에서 다음을 선택하십시오.

보기 > 접기

또는

보기 > 펼치기

개요 수준 변경

1. 개요 분할창에 있는 항목을 클릭하십시오.

2. 메뉴에서 다음을 선택하십시오.

편집 > 개요 > 올리기

또는

편집 > 개요 > 내리기

개요 항목의 크기 변경

1. 메뉴에서 다음을 선택하십시오.

보기 > 개요 크기

2. 개요 크기(작게, 중간 또는 크게)를 선택하십시오.

개요에서 글꼴 변경

1. 메뉴에서 다음을 선택하십시오.

보기 > 개요 글꼴...

2. 글꼴을 선택하십시오.

뷰어에 항목 추가 및 편집

뷰어에서는 제목, 새로운 텍스트, 차트 또는 다른 애플리케이션의 자료 등의 항목을 편집하고 추가할 수 있습니다.

제목 또는 텍스트 추가

테이블이나 차트에 연결되어 있지 않은 텍스트 항목을 뷰어에 추가할 수 있습니다.

1. 제목이나 텍스트 앞에 올 테이블, 차트, 기타 오브젝트를 클릭하십시오.

2. 메뉴에서 다음을 선택하십시오.

삽입 > 새 제목

또는

삽입 > 새 텍스트

3. 새 오브젝트를 두 번 클릭하십시오.

4. 텍스트를 입력하십시오.

텍스트 파일 추가

1. 뷰어의 개요 영역이나 내용 창에서 텍스트 앞에 올 테이블, 차트 또는 기타 오브젝트를 클릭하십시오.
2. 메뉴에서 다음을 선택하십시오.

삽입 > 텍스트 파일...

3. 텍스트 파일을 선택하십시오.

텍스트를 편집하려면 이를 두 번 클릭하십시오.

뷰어에 오브젝트 붙여넣기

다른 애플리케이션의 오브젝트를 뷰어에 붙여넣을 수 있습니다. 뒤에 붙여넣기 또는 선택하여 붙여넣기를 사용할 수 있습니다. 붙여넣기 유형은 뷰어에서 현재 선택된 오브젝트 뒤에 새 오브젝트를 붙여넣습니다. 붙여넣은 오브젝트의 형식을 선택하려면 선택하여 붙여넣기를 사용하십시오.

뷰어에서 찾기 및 바꾸기

1. 뷰어에서 정보를 찾거나 바꾸려면 메뉴에서 다음을 선택합니다.

편집 > 찾기

또는

편집 > 바꾸기

찾기 및 바꾸기를 사용하여 다음 작업을 수행할 수 있습니다.

- 전체 문서 또는 선택한 항목만 검색합니다.
- 현재 위치에서 위나 아래로 검색합니다.
- 두 영역을 모두 검색하거나 내용 또는 개요 분할창으로 검색을 제한합니다.
- 숨겨진 항목을 검색합니다. 여기에는 내용 창의 숨겨진 항목(예: 기본적으로 숨겨지는 참고 테이블) 및 피벗 테이블의 숨겨진 행과 열이 포함됩니다.
- 대소문자가 매치하도록 검색 기준을 제한합니다.
- 전체 셀 내용이 매치하도록 피벗 테이블의 검색 기준을 제한합니다.
- 피벗 테이블의 검색 기준을 꼬리말 표식으로만 제한합니다. 뷰어의 선택에 피벗 테이블이 아닌 다른 것이 포함될 경우 이 옵션을 사용할 수 없습니다.

숨겨진 항목 및 피벗 테이블 레이어

- 다차원 피벗 테이블의 현재 표시된 레이어 아래에 있는 레이어는 숨겨진 것으로 간주되지 않으며 숨겨진 항목이 검색에 포함되어 있지 않을 때도 검색 영역에 포함됩니다.
- 숨겨진 항목은 내용 창의 숨겨진 항목(내용 창에서 닫힌 책 모양 아이콘이 있는 항목 또는 개요 분할창의 접힌 블록 내에 포함된 항목) 및 기본적으로 숨겨졌거나(예: 기본적으로 숨겨진 빈 열과 행) 테이블을 편집하고 특정 행이나 열을 선택적으로 숨겨 수동으로 숨겨진 행과 열을 포함합니다. 숨겨진 항목은 숨겨진 항목 포함을 명시적으로 선택하는 경우에만 검색에 포함됩니다.
- 두 경우 모두 검색 텍스트나 값이 포함되어 있는 숨겨졌거나 표시되지 않은 요소는 발견되면 표시되지만 그 후에 항목은 원래 상태로 돌아갑니다.

피벗 테이블의 값 범위 검색

피벗 테이블의 특정 값 범위에 속하는 값을 찾으려면 다음을 수행합니다.

1. 피벗 테이블을 활성화하거나 뷰어에서 하나 이상의 피벗 테이블을 선택합니다. 피벗 테이블만 선택했는지 확인합니다. 다른 오브젝트가 선택된 경우 범위 옵션을 사용할 수 없습니다.
2. 메뉴에서 다음을 선택하십시오.

편집 > 찾기

3. 범위 탭을 클릭하십시오.

4. 범위 유형: 사이, 이상 또는 이하에서 선택합니다.
 5. 값을 선택하거나 범위를 정의하는 값을 선택합니다.
 - 어느 한쪽 값에 숫자가 아닌 문자가 포함될 경우 두 값 모두 문자열로 처리됩니다.
 - 양쪽 값이 숫자일 경우 숫자 값만 검색됩니다.
 - 범위 템을 사용하여 값을 바꿀 수 없습니다.
- 이 기능은 레거시 테이블에 사용할 수 없습니다. 자세한 정보는 [118 페이지의 『레거시 테이블』](#)의 내용을 참조하십시오.

다른 애플리케이션으로 출력 복사

출력 오브젝트를 다른 워드프로세서나 스프레드시트 같은 애플리케이션으로 복사하고 붙여넣을 수 있습니다. 여러 가지 형식으로 출력을 붙여 넣을 수 있습니다. 대상 애플리케이션 및 선택한 출력 오브젝트에 따라 다음 형식의 일부나 모두를 사용할 수 있습니다.

메타파일. WMF 및 EMF 메타파일 형식. 이러한 형식은 Windows 운영 체제에서만 사용할 수 있습니다.

RTF(서식있는 텍스트). 선택한 다중 오브젝트, 텍스트 출력 및 피벗 테이블은 RTF 형식으로 복사하고 붙여넣을 수 있습니다. 피벗 테이블의 경우 대부분의 애플리케이션에서 테이블을 다른 애플리케이션에서 편집 가능한 테이블로 붙여 넣을 수 있음을 의미합니다. 문서 너비보다 너무 큰 피벗 테이블은 피벗 테이블 옵션 설정에 따라 문서 너비에 맞게 줄 바꿈 또는 축소되거나 변경되지 않고 그대로 있게 됩니다. 자세한 정보는 [119 페이지의 『피벗 테이블 옵션』](#)의 내용을 참조하십시오.

참고: 매우 넓은 테이블은 Microsoft Word에서 올바르게 표시되지 않을 수 있습니다.

이미지. JPG 및 PNG 이미지 형식.

BIFF. 피벗 테이블 및 텍스트 출력은 BIFF 형식으로 스프레드시트에 붙여 넣을 수 있습니다. 피벗 테이블의 숫자는 숫자 정밀도를 유지합니다. 이 형식은 Windows 운영 체제에서만 사용할 수 있습니다.

텍스트. 피벗 테이블 및 텍스트 출력은 텍스트로 복사하고 붙여 넣을 수 있습니다. 이 프로세스는 텍스트만 허용하거나 전송할 수 있는 이메일과 같은 애플리케이션에서 유용합니다.

Microsoft Office Graphic Object. 이 형식을 지원하는 차트를 Microsoft Office 애플리케이션에 복사하여 해당 애플리케이션에서 기본 Microsoft Office 차트로 편집할 수 있습니다. SPSS Statistics/SPSS Modeler 차트와 Microsoft Office 차트 사이의 차이점으로 인해 SPSS Statistics/SPSS Modeler 차트의 일부 기능은 복사된 버전에 유지되지 않습니다. Microsoft Office Graphic Object 형식으로 선택한 여러 차트를 복사하는 기능은 지원되지 않습니다.

대상 애플리케이션이 여러 사용 가능한 형식을 지원하는 경우, 선택하여 붙여넣기 메뉴 항목을 갖고 있어 형식을 선택할 수 있도록 하거나 사용 가능한 형식 목록을 자동으로 표시할 수 있습니다.

참고: 상자도표 및 히스토그램을 복사하여 붙여넣으려면 Microsoft Office 버전 16 이상이 필요합니다.

여러 출력 오브젝트를 복사하여 붙여넣기

다중 출력 오브젝트를 다른 애플리케이션으로 붙여넣을 경우 다음 제한이 적용됩니다.

- **RTF 형식.** 대부분의 애플리케이션에서 피벗 테이블은 해당 애플리케이션에서 편집할 수 있는 테이블로 붙여 넣어집니다. 차트, 트리 및 모델 보기는 이미지로 붙여 넣어집니다.
- **메타파일 및 이미지 형식.** 선택한 모든 출력 오브젝트는 다른 애플리케이션에 하나의 오브젝트로 붙여 넣어집니다.
- **BIFF 형식.** 차트, 트리 및 모델 보기가 제외됩니다.

선택하여 복사

대량의 출력을 복사 및 붙여넣을 때 특히 매우 큰 피벗 테이블은 편집 > 선택하여 복사를 사용하여 클립보드로 복사되는 형식 수를 제한하여 작업 속도를 개선할 수 있습니다.

선택한 형식을 클립보드에 복사할 기본 형식 세트로 저장할 수도 있습니다. 이 설정은 세션 전체에 지속됩니다.

복사 유형

출력 뷰어에서 마우스 오른쪽 단추로 선택된 오브젝트를 클릭하고 편집 > 복사 유형을 선택하여 가장 대중적인 형식(예: 모두, 이미지 또는 Microsoft Office 그래픽 오브젝트)으로 복사하십시오. 편집 > 복사를 선택하면 모두 복사됩니다. 선택된 오브젝트에 대해 복사 유형이 회색으로 표시되거나 표시되지 않으면 해당되는 특정 오브젝트에 대해 이 복사 형식을 사용할 수 없습니다.

대화형 출력

대화형 출력 오브젝트는 여러 개의 관련 출력 오브젝트를 포함합니다. 한 오브젝트에서의 선택사항은 다른 오브젝트에 표시되거나 강조 표시되는 내용을 변경할 수 있습니다. 예를 들어, 테이블에서 행을 선택하면 맵의 영역이 강조 표시되거나 다른 범주의 차트가 표시될 수도 있습니다.

대화형 출력 오브젝트는 텍스트, 색상, 글꼴 또는 테이블 경계선 변경 등과 같은 편집 기능을 지원하지 않습니다. 개별 오브젝트는 대화형 오브젝트에서 뷰어로 복사될 수 있습니다. 대화형 출력에서 복사된 테이블은 피벗 테이블 편집기에서 편집할 수 있습니다.

대화형 출력으로부터 오브젝트 복사

파일 > 뷰어로 복사는 개별 출력 오브젝트를 뷰어 창으로 복사합니다.

- 사용 가능한 옵션은 대화형 출력의 내용에 따라 다릅니다.
- 차트 및 맵은 차트 오브젝트를 작성합니다.
- 테이블은 피벗 테이블 편집기에서 편집할 수 있는 피벗 테이블을 작성합니다.
- 스냅샷은 현재 보기의 이미지를 작성합니다.
- 모델은 현재 대화형 출력 오브젝트의 사본을 작성합니다.

편집 > 오브젝트 복사는 개별 출력 오브젝트를 클립보드로 복사합니다.

- 복사된 오브젝트를 뷰어에 붙여넣는 것은 파일 > 뷰어로 복사에 해당합니다.
- 오브젝트를 다른 애플리케이션에 붙여넣으면 오브젝트를 이미지로 붙여넣습니다.

확대/축소 및 팬

맵의 경우 보기 > 확대/축소를 사용하여 맵의 보기 를 확대/축소할 수 있습니다. 확대/축소된 맵 보기 내에서 보기 > 팬을 사용하여 보기 를 이동할 수 있습니다.

인쇄 설정

파일 > 인쇄 설정은 대화형 오브젝트를 인쇄하는 방법을 제어합니다.

- 볼 수 있는 보기만 인쇄. 현재 표시되는 보기만을 인쇄합니다. 이 옵션은 기본 설정입니다.
- 모든 보기 인쇄. 대화형 출력에 포함된 모든 보기 를 인쇄합니다.
- 선택한 옵션은 또한 출력 오브젝트를 내보내기 위한 기본 동작을 결정합니다.

출력 내보내기

내보내기 출력 결과는 뷰어 결과를 HTML, 텍스트, Word/RTF, Excel, PowerPoint(PowerPoint 97 이상 필요) 또는 PDF 형식으로 저장합니다. 여러 가지 서로 다른 그래픽 형식으로 차트를 내보낼 수도 있습니다.

참고: PowerPoint로 내보내기는 Windows 운영 체제에서만 사용할 수 있습니다.

출력 내보내기

- 대화 상자 내의 아무 곳이나 클릭하여 뷰어를 활성 대화 상자로 설정하십시오.
- 도구 모음에서 내보내기 단추를 클릭하거나 출력 창을 마우스 오른쪽 단추로 클릭하고 내보내기 를 선택하십시오.

3. 파일 이름(또는 차트의 접두문자)을 입력하고 내보내기 형식을 선택하십시오.

내보낼 오브젝트. 뷰어의 모든 오브젝트, 보이는 모든 오브젝트, 또는 선택한 오브젝트만 내보낼 수 있습니다.

문서 유형. 다음 옵션을 사용할 수 있습니다.

- **Word/RTF(*.doc):** 피벗 테이블은 모든 형식 속성이 온전한 상태로 Word 테이블로 내보냅니다(예: 셀 경계선, 글꼴 유형, 배경색). 텍스트 출력은 서식있는 RTF로 내보내집니다. 차트, 트리 다이어그램 및 모델 보기는 PNG 형식에 포함됩니다. Microsoft Word는 너무 넓은 테이블을 제대로 표시하지 않을 수도 있습니다.
- **Excel 97-2004(*.xls)/Excel 2007 이상(*.xlsx):** 피벗 테이블 행, 열 및 셀은 모든 형식 속성이 온전한 상태로 Excel 행, 열 및 셀로 내보냅니다(예: 셀 경계, 글꼴 유형 및 배경색). 텍스트 출력은 모든 글꼴 속성을 그대로 유지한 채 내보냅니다. 텍스트 출력의 각 줄은 줄의 전체 내용이 한 셀이 있는 Excel 파일의 행입니다. 차트, 트리 다이어그램 및 모델 보기는 PNG 형식에 포함됩니다. 출력은 Excel 97-2004 또는 Excel 2007 이상으로 내보낼 수 있습니다.
- **HTML(*.htm):** 피벗 테이블을 HTML 테이블로 내보냅니다. 텍스트 출력은 형식이 미리 설정된 HTML로 내보냅니다. 차트, 트리 다이어그램 및 모델 보기는 선택된 그래픽 형식으로 문서에 임베드됩니다. HTML 5와 호환 가능한 브라우저는 HTML 형식으로 내보내는 출력을 보는 데 필요합니다.
- **포터블 문서 형식(*.pdf):** 모든 출력이 인쇄 미리보기에 표시되는 그대로 내보내지고 모든 형식 속성이 유지됩니다.
- **텍스트 - 일반/UTF8/UTF16 (*.txt):** 텍스트 출력 형식은 일반 텍스트, UTF-8 및 UTF-16을 포함합니다. 피벗 테이블은 탭으로 구분된 형식이나 공백으로 구분된 형식으로 내보낼 수 있습니다. 모든 텍스트 출력은 공백-구분 형식으로 내보냅니다. 차트, 트리 다이어그램 및 모델 보기의 경우 줄은 각 그래픽마다 텍스트 파일에 삽입되어 이미지 파일 이름을 표시합니다.
- **없음(그래픽만)** 사용 가능한 내보내기 형식에는 EPS, JPEG, TIFF, PNG 및 BMP가 포함됩니다. Windows 운영 체제에서는 EMF(확장 메타파일) 형식도 사용할 수 있습니다.

포함 폴더 열기. 내보내기가 작성한 파일을 포함하는 폴더를 엽니다.

HTML 옵션

HTML 내보내기에는 HTML 5와 호환 가능한 브라우저가 필요합니다.

출력을 HTML 형식으로 내보낼 때 사용할 수 있는 옵션은 다음과 같습니다.

피벗 테이블의 레이어. 기본적으로 피벗 테이블 레이어의 포함 또는 제외는 각 피벗 테이블의 테이블 특성으로 제어됩니다. 이 설정을 무시하고 모든 레이어를 포함하거나 현재 표시된 레이어를 제외한 모든 레이어를 제외할 수 있습니다. 자세한 정보는 [114 페이지의 『테이블 특성: 인쇄』](#) 주제를 참조하십시오.

대화형으로 레이어된 테이블 내보내기. 레이어된 테이블은 뷰어에 나타날 때 표시되고 표시된 레이어를 브라우저에서 대화식으로 변경할 수 있습니다. 이 옵션이 선택되지 않으면 각 테이블 레이어는 별도의 테이블로 표시됩니다.

테이블을 HTML로. 내보낸 피벗 테이블에 포함된 유형 정보를 제어합니다.

- **유형 및 고정 열 너비로 내보내기.** 모든 피벗 테이블 유형 정보(글꼴 유형, 배경색 등) 및 열 너비가 보존됩니다.
- **유형 없이 내보내기.** 피벗 테이블이 기본 HTML 테이블로 변환됩니다. 유형 속성이 유지되지 않습니다. 열 너비가 자동으로 결정됩니다.

꼬리말과 캡션을 포함합니다. 모든 피벗 테이블 꼬리말과 캡션의 포함과 제외를 제어합니다.

모델 보기. 기본적으로 모델 보기 포함 또는 제외 여부는 각 모델의 모델 특성에 의해 제어됩니다. 이 설정을 무시하고 모든 보기와 포함하거나 현재 표시된 보기와 제외한 모든 레이어를 제외할 수 있습니다. (참고: 테이블을 포함하여 모든 모델 보기는 그래픽으로 내보냅니다.)

참고: HTML의 경우 내보낸 차트의 이미지 파일 형식을 제어할 수도 있습니다. 자세한 정보는 [103 페이지의 『그래픽 형식 옵션』](#)의 내용을 참조하십시오.

HTML 내보내기 옵션 설정

1. 내보내기 형식으로 **HTML**을 선택하십시오.
2. 옵션 변경을 클릭하십시오.

웹 보고서 옵션

웹 보고서는 대부분의 브라우저와 호환 가능한 대화식 문서입니다. 뷰어에서 사용 가능한 피벗 테이블의 대화식 기능의 대부분은 웹 보고서에서도 사용 가능합니다.

보고서 제목. 보고서의 헤더에 표시되는 제목. 기본적으로, 파일 이름을 사용합니다. 파일 이름 대신에 사용할 사용자 정의 제목을 지정할 수 있습니다.

형식. 보고서 형식에는 두 개의 옵션이 있습니다.

- **SPSS 웹 보고서(HTML 5).** 이 형식은 HTML 5와 호환 가능한 브라우저가 필요합니다.
- **Cognos Active Report(mht).** 이 형식에는 MHT 형식 파일 또는 Cognos Active Report 애플리케이션을 지원하는 브라우저가 필요합니다.

오브젝트 제외. 선택한 오브젝트 유형을 보고서에서 제외할 수 있습니다.

- **텍스트.** 로그가 아닌 텍스트 오브젝트입니다. 이 옵션에는 활성 데이터 세트에 대한 정보를 포함하는 텍스트 오브젝트가 포함됩니다.
- **로그.** 실행된 명령 구문의 목록을 포함하는 텍스트 오브젝트입니다. 로그 항목에는 뷰어 출력을 생성하지 않는 명령이 마주치는 경고 및 오류 메시지를 포함합니다.
- **참고 테이블.** 통계 및 차트 프로시저의 출력에는 참고 테이블이 포함됩니다. 이 테이블에는 사용된 데이터 세트에 대한 정보, 결측값 및 프로시저를 실행하는 데 사용된 명령 구문이 포함됩니다.
- **경고 및 오류 메시지.** 통계 및 차트 프로시저로부터의 경고 및 오류 메시지

웹 보고서와 매치하도록 테이블 및 차트를 재유형화. 이 옵션은 표준 웹 보고서 양식을 모든 테이블과 차트에 적용합니다. 이 설정은 뷰어에 표시된 대로 출력에 있는 모든 글꼴, 색상 또는 다른 유형을 대체합니다. 표준 웹 보고서 유형을 수정할 수 없습니다.

웹 서버 연결. IBM SPSS Statistics Web Report Application Server를 실행 중인 하나 이상의 애플리케이션 서버의 URL 위치를 포함할 수 있습니다. 웹 애플리케이션 서버는 테이블을 피벗하고, 차트를 편집하고 수정된 웹 보고서를 저장하기 위한 기능을 제공합니다.

- 웹 보고서에 포함하려는 각 애플리케이션 서버마다 **사용**을 선택하십시오.
- 웹 보고서에 URL 스페이 포함된 경우, 웹 보고서는 추가 편집 기능을 제공하기 위해 해당 애플리케이션 서버에 연결합니다.
- 여러 URL을 지정하는 경우 웹 보고서는 서버가 지정된 순서대로 각 서버에 연결하려고 시도합니다.

IBM SPSS Statistics Web Report Application Server는 <http://www.ibm.com/developerworks/spssdevcentral>에서 다운로드할 수 있습니다.

Word 옵션

결과를 Word 형식으로 내보낼 때 사용할 수 있는 옵션은 다음과 같습니다.

구분점 및 집단 보존 구분점을 정의한 경우에는 이러한 설정이 Word 테이블에 보존됩니다.

꼬리말과 캡션 포함. 모든 피벗표 꼬리말과 캡션의 포함 또는 제외를 제어합니다.

모델 보기. 기본적으로 모델 보기 포함 또는 제외 여부는 각 모델의 모델 특성에 의해 제어됩니다. 이 설정을 무시하고 모든 보기와 포함하거나 현재 표시된 보기와 제외한 모든 레이어를 제외할 수 있습니다. (참고: 테이블을 포함하여 모든 모델 보기는 그래픽으로 내보냅니다.)

내보내기를 위한 페이지 설정. 내보낼 문서의 페이지 크기 및 여백을 정의할 수 있는 대화 상자를 엽니다. 줄바꿈이나 축소 동작 설정에 사용하는 문서 너비는 페이지 너비에서 오른쪽 및 왼쪽 여백을 뺀 값입니다.

Word 내보내기 옵션 설정

1. 내보내기 형식으로 **Word/RTF**를 선택합니다.
2. 옵션 변경을 클릭하십시오.

Excel 옵션

결과를 Excel 형식으로 내보낼 때 사용할 수 있는 옵션은 다음과 같습니다.

워크시트 또는 워크북을 작성하거나 기존 워크시트를 수정. 기본적으로 새 워크북이 작성됩니다. 특정 이름이 이미 존재할 경우에는 이 이름을 덮어쓰게 됩니다. 옵션을 선택하여 워크시트를 생성할 때 지정된 이름을 가진 워크시트가 지정된 파일에 이미 존재할 경우에는 기존 이름을 덮어쓰게 됩니다. 옵션을 선택하여 기존 워크시트를 수정하려면 워크시트 이름도 지정해야 합니다. (이는 워크시트를 작성하기 위한 선택사항입니다.) 워크시트 이름은 31자를 초과할 수 없으며 슬래시(/) 또는 백슬래시(\), 대괄호, 물음표, 별표 등이 포함될 수 없습니다.

Excel 97-2004로 내보낼 때 기존의 워크시트를 수정할 경우 차트, 모델 보기 및 트리 다이어그램은 내보낸 결과에 포함되지 않습니다.

워크시트에서의 위치. 내보낸 파일에 대한 워크시트 내의 위치를 제어합니다. 기본적으로 내보낸 파일은 첫 번째 행에서 시작하여 기존 내용을 수정하지 않고 모든 내용이 들어 있는 마지막 열 다음에 추가됩니다. 기존의 워크시트에 새 열을 추가하는 방식이 좋습니다. 마지막 열 다음에 내보낸 결과를 추가하는 것은 기존의 워크시트에 새 행을 추가하는 데 유용한 방식입니다. 특정 셀의 위치에서부터 내보내는 출력을 추가하면 내보낸 파일이 추가되는 영역에 있는 기존의 모든 내용을 덮어쓰게 됩니다.

피벗 테이블의 레이어. 기본적으로 피벗 테이블 레이어의 포함 또는 제외는 각 피벗 테이블의 테이블 특성으로 제어됩니다. 이 설정을 무시하고 모든 레이어를 포함하거나 현재 표시된 레이어를 제외한 모든 레이어를 제외할 수 있습니다. 자세한 정보는 [114 페이지의 『테이블 특성: 인쇄』](#) 주제를 참조하십시오.

꼬리말과 캡션 포함. 모든 피벗 테이블 꼬리말과 캡션의 포함 또는 제외를 제어합니다.

모델 보기. 기본적으로 모델 보기 포함 또는 제외 여부는 각 모델의 모델 특성에 의해 제어됩니다. 이 설정을 무시하고 모든 보기를 포함하거나 현재 표시된 보기를 제외한 모든 레이어를 제외할 수 있습니다. (참고: 테이블을 포함하여 모든 모델 보기는 그래픽으로 내보냅니다.)

Excel 내보내기 옵션 설정

1. 내보내기 형식으로 **Excel**을 선택하십시오.
2. 옵션 변경을 클릭하십시오.

PowerPoint 옵션

다음 옵션을 PowerPoint에 사용할 수 있습니다.

피벗 테이블의 레이어. 기본적으로 피벗 테이블 레이어의 포함 또는 제외는 각 피벗 테이블의 테이블 특성으로 제어됩니다. 이 설정을 무시하고 모든 레이어를 포함하거나 현재 표시된 레이어를 제외한 모든 레이어를 제외할 수 있습니다. 자세한 정보는 [114 페이지의 『테이블 특성: 인쇄』](#) 주제를 참조하십시오.

넓은 피벗 테이블. 정의된 문서 너비에 비해 너무 넓은 테이블 처리를 제어합니다. 기본적으로 테이블이 너비에 맞게 줄바꿈됩니다. 테이블이 여러 섹션으로 나뉘며 행 레이블이 테이블의 각 섹션에 대해 반복됩니다. 또는 넓은 테이블을 축소하거나 그대로 둘 수 있으며 정의된 문서의 너비보다 크게 확장할 수 있습니다.

꼬리말과 캡션을 포함합니다. 모든 피벗 테이블 꼬리말과 캡션의 포함과 제외를 제어합니다.

뷰어 개요 항목을 슬라이드 제목으로 사용. 내보내기로 작성한 각 슬라이드에 제목을 포함합니다. 각 슬라이드에는 뷰어에서 내보내는 항목이 하나씩 포함됩니다. 제목은 뷰어의 개요 분할창에 있는 항목에 대한 개요 항목을 통해 구성됩니다.

모델 보기. 기본적으로 모델 보기 포함 또는 제외 여부는 각 모델의 모델 특성에 의해 제어됩니다. 이 설정을 무시하고 모든 보기를 포함하거나 현재 표시된 보기를 제외한 모든 레이어를 제외할 수 있습니다. (참고: 테이블을 포함하여 모든 모델 보기는 그래픽으로 내보냅니다.)

내보내기를 위한 페이지 설정. 내보낼 문서의 페이지 크기 및 여백을 정의할 수 있는 대화 상자를 엽니다. 줄바꿈이나 축소 동작 결정에 사용하는 문서 너비는 페이지 너비에서 오른쪽 및 왼쪽 여백을 뺀 값입니다.

PowerPoint 내보내기 옵션 설정

1. 내보내기 형식으로 **PowerPoint**를 선택합니다.
2. 옵션 변경을 클릭하십시오.

참고: PowerPoint로 내보내기는 Windows 운영 체제에서만 사용 가능합니다.

PDF 옵션

다음 옵션을 PDF에 사용할 수 있습니다.

책갈피 포함. 이 옵션을 선택하면 뷰어 개요 항목에 해당하는 책갈피가 PDF 문서에 포함됩니다. 뷰어 개요 분할 창에서와 마찬가지로 책갈피를 사용하면 출력 오브젝트 수가 매우 많은 문서도 훨씬 쉽게 탐색할 수 있습니다.

글꼴 포함. 어떤 컴퓨터에서건 PDF 문서가 동일하게 표시되도록 글꼴을 삽입합니다. 이 옵션을 사용하지 않는 경우 PDF 문서를 보거나 인쇄하는 데 사용하는 컴퓨터에 이 문서를 작성할 때 사용했던 글꼴 중 일부가 설치되어 있지 않으면 대체 글꼴을 사용해야 하므로 최상의 결과물을 얻지 못할 수도 있습니다.

피벗 테이블의 레이어. 기본적으로 피벗 테이블 레이어의 포함 또는 제외는 각 피벗 테이블의 테이블 특성으로 제어됩니다. 이 설정을 무시하고 모든 레이어를 포함하거나 현재 표시된 레이어를 제외한 모든 레이어를 제외할 수 있습니다. 자세한 정보는 [114 페이지의 『테이블 특성: 인쇄』](#) 주제를 참조하십시오.

모델 보기. 기본적으로 모델 보기 포함 또는 제외 여부는 각 모델의 모델 특성에 의해 제어됩니다. 이 설정을 무시하고 모든 보기의 포함하거나 현재 표시된 보기의 제외한 모든 레이어를 제외할 수 있습니다. (참고: 테이블을 포함하여 모든 모델 보기는 그래픽으로 내보냅니다.)

PDF 내보내기 옵션 설정

1. 내보내기 형식으로 **포터블 문서 형식**을 선택합니다.

2. 옵션 변경을 클릭하십시오.

PDF 출력에 영향을 주는 기타 설정

페이지 설정/페이지 속성. PDF 문서의 페이지 크기, 방향, 여백, 페이지 머리말과 꼬리말의 내용 및 출력, 인쇄된 문자 크기는 페이지 설정과 페이지 속성 옵션으로 제어합니다.

테이블 특성/테이블모양. 가로로 길거나 세로로 긴 테이블의 척도 및 테이블 레이어의 인쇄는 각 테이블에 대한 테이블 특성을 통해 제어합니다. 이러한 특성은 테이블모양에도 저장할 수 있습니다.

기본값/현재 프린터. PDF 문서의 해상도(DPI)는 기본값 프린터 또는 현재 선택된 프린터의 현재 해상도 설정입니다. 이는 페이지 설정을 사용하여 변경할 수 있습니다. 최대 해상도는 1200DPI입니다. 프린터 설정이 이보다 높더라도 PDF 문서는 1200DPI로 설정됩니다.

참고: 고해상도 문서를 저해상도 프린터에서 인쇄하면 인쇄 품질이 떨어질 수 있습니다.

텍스트 옵션

텍스트 출력 옵션 설정

다음 옵션을 텍스트 내보내기에 사용할 수 있습니다.

피벗 테이블 형식. 피벗 테이블은 템으로 구분된 형식이나 공백으로 구분된 형식으로 내보낼 수 있습니다. 공백으로 구분된 형식의 경우 다음을 제어할 수도 있습니다.

- 열 너비.** 자동맞춤은 열 내용을 줄바꿈하지 않으며 각 열의 너비는 해당 열에서 가장 넓은 테이블이나 값 너비와 동일합니다. 사용자 정의 설정은 테이블에 있는 모든 열에 적용되는 최대 열 너비를 설정하며 이 너비를 초과하는 값은 해당 열에서 다음 줄로 줄바꿈됩니다.
- 행/열 경계 문자.** 행과 열 경계를 작성하는 데 사용하는 문자를 제어합니다. 행과 열 경계를 표시하지 않으려면 값으로 빈 공백을 입력하십시오.

피벗 테이블의 레이어. 기본적으로 피벗 테이블 레이어의 포함 또는 제외는 각 피벗 테이블의 테이블 특성으로 제어됩니다. 이 설정을 무시하고 모든 레이어를 포함하거나 현재 표시된 레이어를 제외한 모든 레이어를 제외할 수 있습니다. 자세한 정보는 [114 페이지의 『테이블 특성: 인쇄』](#) 주제를 참조하십시오.

꼬리말과 캡션을 포함합니다. 모든 피벗 테이블 꼬리말과 캡션의 포함과 제외를 제어합니다.

모델 보기. 기본적으로 모델 보기 포함 또는 제외 여부는 각 모델의 모델 특성에 의해 제어됩니다. 이 설정을 무시하고 모든 보기의 포함하거나 현재 표시된 보기의 제외한 모든 레이어를 제외할 수 있습니다. (참고: 테이블을 포함하여 모든 모델 보기는 그래픽으로 내보냅니다.)

텍스트 내보내기 옵션 설정

1. 내보내기 형식으로 **텍스트**를 선택합니다.

2. 옵션 변경을 클릭하십시오.

이미지만 옵션

다음 옵션을 이미지만 내보내기에 사용할 수 있습니다.

1. 탭의 아무 곳이나 클릭하여 출력 탭을 활성 탭으로 설정하십시오.
2. 내보낼 차트 항목을 선택하고 수직 생략 기호 컨트롤을 클릭한 후 메뉴에서 **내보내기**를 선택하십시오. 또한 도구 모음에서 **출력 내보내기** 단추를 클릭할 수도 있습니다. 내보내기 출력결과 대화 상자가 표시됩니다.

문서 설정

유형

문서 유형으로 이미지만을 선택하십시오.

이미지 형식

내보내는 이미지의 이미지 형식을 선택하십시오.

PNG

선택할 경우 이미지를 Portable Network Graphics(*.png) 형식으로 내보냅니다. 이 설정은 이미지 척도(%) 및 색 농도를 선택하는 옵션을 제공합니다.

JPG

선택할 경우 이미지를 Joint Photographic Experts Group(*.jpg) 형식으로 내보냅니다. 이 설정은 이미지 척도(%) 및 그레이스케일로 변환을 선택하는 옵션을 제공합니다.

BMP

선택할 경우 이미지를 비트맵(*.bmp) 형식으로 내보냅니다. 이 설정은 이미지 척도(%) 및 파일 크기를 줄이기 위해 이미지 압축을 선택하는 옵션을 제공합니다.

포함 폴더 열기

(선택사항) 내보내기 후 내보내기 저장 파일 위치 폴더를 열려면 이 설정을 선택하십시오.

내보낼 오브젝트

포함될 오브젝트

모든, 표시된 또는 선택된(기본 설정) 출력 오브젝트를 내보내려면 선택하십시오.

3. 문서 설정을 지정한 후 **내보내기**를 클릭하십시오.

그래픽 형식 옵션

HTML 및 텍스트 문서에서 차트 내보내기만 하는 경우에는 그래픽 형식을 선택할 수 있고 각 그래픽 형식에 대해서는 다양한 옵션 설정을 제어할 수 있습니다.

내보낸 차트의 그래픽 형식 및 옵션을 선택

1. **HTML, 텍스트 또는 없음(그래픽만 해당)**을 문서 유형으로 선택하십시오.
2. 드롭 다운 목록에서 그래픽 파일 형식을 선택하십시오.
3. 옵션 변경을 클릭하여 선택한 그래픽 파일 형식의 옵션을 변경하십시오.

JPEG 차트 내보내기 옵션

이미지 크기(%)

최대 200 퍼센트의 원래 차트 크기 백분율.

그레이스케일로 변환

색상을 회색 음영으로 변환합니다.

BMP 차트 내보내기 옵션

이미지 크기(%)

최대 200 퍼센트의 원래 차트 크기 백분율.

파일 크기를 줄이기 위해 이미지 압축

이미지 품질에 영향을 주지 않고 작은 파일을 작성하는 무손실 압축 기법입니다.

PNG 차트 내보내기 옵션

이미지 크기(%)

최대 200 퍼센트의 원래 차트 크기 백분율.

색 농도

내보낸 차트에서 사용하는 색상 수를 결정합니다. 임의의 깊이로 저장된 차트에는 실제로 사용된 최소 색상 수와 해당 깊이가 허용하는 최대 색상 수가 포함됩니다. 예를 들어, 차트를 세 가지 색(빨강, 흰색, 검정)을 사용하여 작성하고 16색으로 저장하면 차트는 세 가지 색상으로 표시됩니다.

참고: 차트에서 사용한 색상 수가 해당 깊이에 대한 색상 수를 초과하는 경우 색상이 혼합되어 차트의 색상을 복제합니다.

EMF 및 TIFF 차트 내보내기 옵션

이미지 크기. 최대 200 퍼센트의 원래 차트 크기 백분율.

참고: EMF(향상된 메타파일) 형식은 Windows 운영 체제에서만 사용 가능합니다.

EPS 차트 내보내기 옵션

이미지 크기. 원래 이미지 크기(최대 200 퍼센트)의 백분율로 크기를 지정하거나 이미지 너비를 픽셀(너비 값과 가로 세로 비율로 결정되는 높이)로 지정할 수 있습니다. 내보낸 이미지는 항상 원본에 비례합니다.

TIFF 미리보기 이미지 포함. EPS 이미지를 화면에 표시할 수 없는 애플리케이션에서 EPS 이미지를 볼 수 있도록 이를 TIFF 형식의 미리보기와 함께 저장합니다.

글꼴. EPS 이미지에서의 글꼴 처리 방법을 제어합니다.

- **글꼴 참조 사용.** 차트에 사용된 글꼴을 출력 장치에서도 사용할 수 있으면 해당 글꼴을 사용합니다. 그렇지 않으면 출력 장치에서 대체 글꼴을 사용합니다.
- **글꼴을 곡선으로 바꾸기.** 글꼴을 포스트스크립트 곡선 데이터로 변경합니다. 텍스트 자체는 EPS 그래픽을 편집할 수 있는 애플리케이션의 텍스트로는 더 이상 편집할 수 없습니다. 이 옵션은 차트를 작성하는 데 사용한 글꼴을 출력 장치에서 사용할 수 없는 경우에 유용합니다.

뷰어 인쇄

뷰어 창의 내용을 인쇄하는 옵션은 두 가지가 있습니다.

모든 가시출력결과. 내용 창에 현재 표시되어 있는 항목만 인쇄합니다. 숨겨진 항목(개요 분할창에 닫힌 책 모양 아이콘이 있는 항목 또는 접힌 개요 레이어에 숨겨진 항목)은 인쇄되지 않습니다.

선택. 개요 및/또는 내용 창에서 현재 선택된 항목만 인쇄합니다.

출력과 차트 인쇄

1. 뷰어를 활성 창으로 만드십시오(창의 임의의 위치 클릭).
2. 메뉴에서 다음을 선택하십시오.

파일 > 인쇄...

3. 원하는 인쇄 설정을 선택합니다.
4. 확인을 클릭하여 인쇄합니다.

인쇄 미리보기

인쇄 미리보기로 뷰어 문서의 각 페이지에 인쇄할 내용을 볼 수 있습니다. 인쇄 미리보기로 다음을 포함하여 뷰어의 내용 창에서 찾아 볼 수 없는 항목을 볼 수 있으므로 뷰어 문서를 실제로 인쇄하기 전에 인쇄 미리보기로 확인하는 것이 좋습니다.

- 페이지 구분
- 피벗 테이블의 은닉총

- 넓은 테이블의 구분
- 각 페이지에 인쇄되는 헤더 및 꼬리말

어떤 출력이 뷰어에 현재 선택되어 있으면 미리보기는 선택한 출력만 나타냅니다. 모든 출력에 대한 미리보기를 실행하려면 뷰어에 선택된 출력이 없는지 확인하십시오.

페이지 속성: 머리말 및 꼬리말

머리말과 꼬리말은 각 페이지의 상단과 하단에 인쇄되는 정보입니다. 머리말 및 꼬리말로 사용할 텍스트를 입력할 수 있습니다. 삽입할 대화 상자 가운데에 있는 도구 모음을 사용하여 다음을 입력할 수도 있습니다.

- 날짜 및 시간
- 페이지 번호
- 뷰어 파일 이름
- 개요 머리말 테이블
- 페이지 제목 및 부제목
- 기본값 설정은 새로운 뷰어 문서의 기본 설정으로 여기에 지정한 설정을 사용합니다. (참고: 그러면 머리말/꼬리말 템과 기본 설정의 옵션 템 둘 모두에서 현재 설정이 됩니다.)
- 개요 머리말 테이블은 각 페이지에 있는 첫 번째 항목에 대한 첫 번째, 두 번째, 세 번째 또는 네 번째 수준의 개요 머리말을 나타냅니다.
- 페이지 제목과 부제목은 현재 페이지 제목과 부제목을 인쇄합니다. 이것은 뷰어 삽입 메뉴의 새 페이지 제목 또는 TITLE 및 SUBTITLE 명령을 사용하여 작성할 수 있습니다. 페이지 제목이나 부제목을 지정하지 않았다면 이 설정은 무시됩니다.

참고: 새 페이지 제목과 부제목의 글꼴 특성은 옵션 대화 상자의 뷰어 템에서 제어됩니다(편집 메뉴에서 옵션을 선택하여 액세스). 기존의 페이지 제목과 부제목에 대한 글꼴 특성을 변경하려면 뷰어에서 제목을 편집하면 됩니다.

인쇄된 페이지에서 머리말과 꼬리말이 어떻게 표시되는지 확인하려면 파일 메뉴에서 인쇄 미리보기를 선택합니다.

페이지 헤더 및 꼬리말 삽입

- 뷰어를 활성 창으로 만드십시오(창의 임의의 위치 클릭).
 - 메뉴에서 다음을 선택하십시오.
- 파일 > 헤더 및 꼬리말...
- 각 페이지에 표시할 헤더 및/또는 꼬리말을 입력합니다.

페이지 속성: 옵션

이 대화 상자는 인쇄된 차트 크기, 인쇄된 출력 항목 사이의 공간 및 페이지 번호 매김을 제어합니다.

- 차트 인쇄 크기. 정의된 페이지 크기에 상대적인 차트 인쇄 크기를 제어합니다. 차트의 가로 세로 비율(너비 대 높이 비율)은 차트 인쇄 크기의 영향을 받지 않습니다. 차트의 전체 인쇄 크기는 높이와 너비로 제한됩니다. 차트 외부의 경계가 페이지의 왼쪽과 오른쪽 경계선에 도달하면 차트 크기가 페이지 높이를 더 채우기 위해서 더 이상 커지지 않습니다.
- 항목 간 간격. 인쇄된 항목 간의 공백을 제어합니다. 각 피벗 테이블, 차트 및 텍스트 오브젝트는 개별 항목입니다. 이 설정은 뷰어의 항목 표시에 영향을 미치지 않습니다.
- 페이지 번호 시작. 지정한 번호부터 순차적으로 페이지 번호를 시작합니다.
- 기본값 지정. 이 옵션은 새 뷰어 문서의 기본 설정으로 여기에 지정되는 설정을 사용합니다. (이 옵션은 현재 머리말/꼬리말 설정 및 옵션 설정을 기본값으로 만드는 점을 유의하십시오.)

차트 인쇄 크기, 페이지 매김, 인쇄된 항목 간 공백 변경 방법

- 뷰어를 활성 창으로 만드십시오(창의 임의의 위치 클릭).

2. 메뉴에서 다음을 선택하십시오.
파일 > 페이지 속성...
3. 옵션 탭을 클릭하십시오.
4. 설정을 변경하고 확인을 클릭하십시오.

출력 저장

뷰어의 내용은 저장할 수 있습니다.

- **출력 오브젝트(*.cou).** 이 형식은 그래프, 탭, 주석(Annotation) 등을 포함하여 전체 출력 컨테이너를 저장합니다. 이 형식은 IBM SPSS Modeler에서 열리고 표시되고, 프로젝트에 추가되고, IBM SPSS Collaboration and Deployment Services Repository를 사용하여 게시되고 추적됩니다. 이 형식은 IBM SPSS Statistics과 호환되지 않습니다.
- **뷰어 파일(*.spv).** 뷰어 창에 파일을 표시하는 데 사용되는 형식입니다. IBM SPSS Modeler에서 모델 너깃으로부터 이 형식을 저장하면 모델 탭의 뷰어 내용만 저장됩니다.

웹 보고서를 저장하기 위한 옵션을 제어하거나 결과를 다른 형식으로 저장하려면(예: 텍스트, Word, Excel), 파일 메뉴에서 내보내기를 사용하십시오.

뷰어 문서 저장

1. 뷰어 창에서 다음을 선택합니다.

파일 > 저장

2. 문서 이름을 입력하고 저장을 클릭하십시오.

선택적으로 다음을 수행할 수 있습니다.

파일을 잠가 IBM SPSS Smartreader에서의 편집 방지.

뷰어 문서가 잠기면 피벗 테이블을 조작할 수 있습니다(행과 열 바꾸기, 표시된 레이어 변경 등). 그러나 IBM SPSS Smartreader에서 출력을 편집하거나 변경사항을 뷰어 문서에 저장할 수 없습니다(뷰어 문서에서 작업하기 위한 별도의 제품). 이 설정은 IBM SPSS Statistics 또는 IBM SPSS Modeler에서 열린 뷰어 문서에는 영향을 미치지 않습니다.

파일을 비밀번호로 암호화

비밀번호로 문서를 암호화하여 뷰어 문서에 저장된 기밀 정보를 보호할 수 있습니다. 암호화되면 비밀번호를 제공해야만 문서를 열 수 있습니다. IBM SPSS Smartreader 사용자 또한 파일을 열기 위해서는 비밀번호를 제공해야 합니다.

뷰어 문서를 암호화하는 방법은 다음과 같습니다.

- a. 다른 이름으로 출력 저장 대화 상자에서 **비밀번호로 파일 암호화**를 선택합니다.
- b. 저장을 클릭합니다.
- c. 파일 암호화 대화 상자에서 비밀번호를 제공하고 비밀번호 확인 텍스트 상자에 비밀번호를 다시 입력합니다. 비밀번호는 10자로 제한되며 대소문자를 구분합니다.

경고: 비밀번호는 분실되면 복구할 수 없습니다. 비밀번호를 분실하면 파일을 열 수 없습니다.

강력한 비밀번호 작성

- 8자 이상을 사용하십시오.
- 비밀번호에 숫자, 기호 및 문장 부호까지 포함하십시오.
- "123" 및 "abc"처럼 연속된 숫자나 문자를 피하며 "111aaa"처럼 반복하지 않습니다.
- 생일 또는 별명과 같은 개인 정보를 사용하여 비밀번호를 만들지 마십시오.
- 비밀번호를 주기적으로 변경하십시오.

참고: 암호화된 파일을 IBM SPSS Collaboration and Deployment Services Repository에 저장하는 것은 지원되지 않습니다.

암호화된 파일 수정

- 암호화된 파일을 열어 파일을 수정하고 파일 > 저장을 선택할 경우 수정된 파일은 같은 비밀번호로 저장됩니다.
- 파일을 열고 암호화 단계를 반복한 후 파일 암호화 대화 상자에 다른 비밀번호를 지정하여 비밀번호를 변경할 수 있습니다.
- 파일을 열고 파일 > 다른 이름으로 저장을 선택한 후 다른 이름으로 출력 저장 대화 상자에서 **비밀번호로 파일 암호화**를 선택 해제하여 암호화된 파일의 암호화되지 않은 버전을 저장할 수 있습니다.

참고: 암호화된 데이터 파일 및 출력 문서는 버전 21 이전의 IBM SPSS Statistics 버전에서는 열 수 없습니다. 암호화된 구문 파일은 버전 22 이전에서는 열 수 없습니다.

필수 모델 정보를 출력 문서와 함께 저장

이 옵션은 일부 대화식 기능을 사용하기 위해서는 보조 정보가 필요한 출력 문서에 모델 뷰어 항목이 있는 경우에만 적용됩니다. 보조 정보를 요구하는 대화식 기능 및 이러한 모델 뷰어 항목 목록을 표시하려면 **자세한 정보**를 클릭하십시오. 이 정보를 출력 문서와 함께 저장하면 문서 크기가 대폭 커집니다. 이 정보를 저장하지 않기로 선택한 경우에도 이러한 출력 항목을 열 수는 있지만 지정된 대화식 기능은 사용할 수 없습니다.

피벗 테이블

피벗 테이블

여러 결과는 대화형으로 피벗할 수 있는 테이블로 나타낼 수 있습니다. 즉, 행, 열 및 레이어를 다시 배열할 수 있습니다.

피벗 테이블 조작

피벗 테이블 조작에는 다음과 같은 옵션이 있습니다.

- 행과 열 전치
- 행과 열 이동
- 다차원 레이어 작성
- 행과 열의 그룹 설정 및 해제
- 행, 열 및 기타 정보 표시 및 숨김
- 행과 열 레이블 회전
- 항에 대한 정의 찾기

피벗 테이블 활성화

피벗 테이블을 조작하거나 수정하기 전에 테이블을 활성화해야 합니다. 테이블 활성화 방법

- 테이블을 두 번 클릭하십시오.
- 또는
- 테이블을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 **내용 편집**을 선택하십시오.
- 하위 메뉴에서 **뷰어에서 또는 별도의 창에서**를 선택하십시오.

테이블 피벗

테이블에는 행, 열 및 레이어의 3차원이 있습니다. 차원은 여러 요소를 포함할 수 있습니다(또는 어떠한 요소도 포함하지 않을 수 있음). 차원 간에 또는 차원 내에서 요소를 이동하여 테이블의 구성을 변경할 수 있습니다. 요소를 이동하려면 끌어다 원하는 곳에 놓습니다.

차원 내에서 요소의 표시 순서 변경

테이블 차원(행, 열 또는 레이어) 내에서 요소의 표시 순서 변경

- 피벗 트레이가 설정되지 않은 경우에는 피벗 테이블 메뉴에서 다음을 선택합니다.

피벗 > 피벗 트레이

2. 피벗 트레이의 차원 내에 요소를 끌어다 놓습니다.

차원 요소 내에서 열 및 행 이동

1. 테이블 자체에서(피벗 트레이가 아님) 이동하려는 행이나 열의 레이블을 클릭하십시오.
2. 레이블을 새 위치에 끌어다 놓으십시오.

행과 열 전치

행과 열을 뒤집기만 하려는 경우 피벗 트레이를 사용하는 간단한 방법이 있습니다.

1. 메뉴에서 다음을 선택하십시오.

피벗 > 행과 열 전치

열 차원으로 행 요소를 모두 끌 때와 행 요소로 열 요소를 모두 끌 때 같은 효과가 나타납니다.

행 또는 열 그룹화

1. 그룹화할 행이나 열에 대해 레이블을 선택합니다. 마우스로 끌거나 Shift와 함께 마우스로 눌러 여러 레이블을 선택합니다.
2. 메뉴에서 다음을 선택하십시오.

편집 > 그룹

그룹 레이블이 자동으로 삽입됩니다. 레이블 텍스트를 편집하려면 해당 그룹 레이블을 두 번 클릭합니다.

참고: 기존 그룹에 행이나 열을 추가하려면 먼저 현재 그룹에 있는 항목을 그룹 해제해야 합니다. 그런 다음 추가 항목을 포함하는 새 그룹을 작성할 수 있습니다.

행이나 열 그룹 해제

그룹 해제를 사용하면 그룹 레이블이 자동으로 삭제됩니다.

행이나 열 레이블 회전

테이블에서 가장 안쪽의 열 레이블과 가장 바깥쪽의 행 레이블에 대해 수평 및 수직 출력 간에 레이블을 회전할 수 있습니다.

1. 메뉴에서 다음을 선택하십시오.

형식 > 내부 열 레이블 회전

또는

형식 > 외부 행 레이블 회전

가장 안쪽의 열 레이블과 가장 바깥쪽의 행 레이블만 회전할 수 있습니다.

행 정렬

피벗 테이블의 행을 정렬하려면 다음을 수행합니다.

1. 테이블을 활성화합니다.
 2. 정렬을 사용하려는 열에서 셀을 선택합니다. 선택한 행 그룹만 정렬하려면 정렬을 사용하려는 열에서 두 개 이상의 연속 셀을 선택합니다.
 3. 메뉴에서 다음을 선택하십시오.
- ## 편집 > 행 정렬
4. 하위 메뉴에서 오름차순 또는 내림차순을 선택합니다.
- 행 차원에 그룹이 포함된 경우, 정렬은 선택을 포함하는 그룹에만 영향을 미칩니다.
- 108 IBM SPSS Modeler 18.3 사용자 안내서

- 그룹 경계를 걸쳐 정렬할 수 없습니다.
- 행 차원에 항목이 둘 이상인 테이블은 정렬할 수 없습니다.

행 및 열 삽입

피벗 테이블에 행 또는 열을 삽입하려면 다음을 수행하십시오.

- 테이블을 활성화합니다.
- 테이블에서 셀을 선택합니다.
- 메뉴에서 다음을 선택하십시오.

앞에 삽입

또는

뒤에 삽입

하위 메뉴에서 다음을 선택하십시오.

행

또는

열

- 셀이 비어있기 때문에 자동으로 숨겨지는 것을 방지하기 위해 새 행이나 열의 각 셀에 더하기 부호(+)를 삽입합니다.
- 중첩되거나 레이어 차원인 테이블에서는 모든 해당하는 차원 수준에 열이나 행이 삽입됩니다.

변수 및 변수값 레이블 표시 제어

변수에 구체적인 변수 또는 변수값 레이블이 들어 있을 경우, 변수 이름과 레이블 및 피벗 테이블의 데이터 값과 변수값 레이블 표시를 제어할 수 있습니다.

- 피벗 테이블을 활성화합니다.
- 메뉴에서 다음을 선택하십시오.

보기 > 변수 레이블

또는

보기 > 값 레이블

- 하위 메뉴에서 다음 옵션 중 하나를 선택합니다.
 - 이름 또는 값. 변수 이름(또는 값)만이 표시됩니다. 설명 레이블은 표시되지 않습니다.
 - 레이블.설명 레이블만 표시됩니다. 변수 이름(또는 값)은 표시되지 않습니다.
 - 모두. 이름(또는 값) 및 설명 레이블 둘 모두가 표시됩니다.

출력 언어 변경

피벗 테이블의 출력 언어를 변경하려면 다음을 수행하십시오.

- 테이블을 활성화합니다.
- 메뉴에서 다음을 선택하십시오.

보기 > 언어

- 사용 가능한 언어 중 하나를 선택합니다.

언어를 변경하면 테이블 제목, 행과 열 레이블 및 꼬리말 텍스트 등과 같이 애플리케이션이 생성한 텍스트에만 영향을 미칩니다. 변수 이름과 설명 변수 및 값 레이블은 영향을 받지 않습니다.

큰 테이블 탐색

탐색창을 사용하여 큰 테이블을 탐색하려면 다음을 수행하십시오.

1. 테이블을 활성화합니다.
2. 메뉴에서 다음을 선택하십시오.

보기 > 탐색

변경 실행 취소

활성화된 피벗 테이블에 작성한 마지막 변경 또는 모든 변경을 실행 취소할 수 있습니다. 두 동작은 모두 테이블의 가장 최근 활성화 이후에 작성된 변경에만 적용됩니다.

마지막 변경 실행 취소:

1. 메뉴에서 다음을 선택하십시오.

편집 > 실행 취소

모든 변경 실행 취소:

2. 메뉴에서 다음을 선택하십시오.

편집 > 복원

레이어 작업

각 범주나 범주 조합에 대해 별도의 2차원 테이블을 표시할 수 있습니다. 이러한 테이블은 레이어에서 누적된 것으로 간주되며 맨 위 레이어만 볼 수 있습니다.

레이어 작성 및 표시

레이어 작성

1. 피벗 트레이가 설정되지 않은 경우에는 **피벗 테이블** 메뉴에서 다음을 선택합니다.

피벗 > 피벗 트레이

2. 행이나 열 차원의 요소를 레이어 차원으로 끕니다.

요소를 레이어 차원으로 이동하면 다차원 테이블이 작성되지만 단 하나의 2차원 "조각"만이 표시됩니다. 표시되는 테이블은 위쪽 레이어에 대한 테이블입니다. 예를 들어, 예/아니오 범주형 변수가 레이어 차원에 있는 경우에는 다차원 테이블에는 2개의 레이어가 있습니다. 하나는 "예" 범주이고 하나는 "아니오" 범주입니다.

표시된 레이어 변경

1. 레이어의 드롭 다운 목록에서 범주를 선택합니다(피벗 트레이가 아닌 피벗 테이블 자체에 있음).

레이어 범주로 이동

레이어 범주로 이동을 사용하여 피벗 테이블에서 레이어를 바꿀 수 있습니다. 이 대화 상자는 많은 레이어가 있거나 선택한 레이어에 많은 범주가 있을 때 특히 유용합니다.

항목 표시 및 숨김

다음을 비롯한 많은 셀 유형을 숨길 수 있습니다.

- 차원 레이블
- 행이나 열에서 레이블 셀과 데이터 셀을 포함하는 범주
- 범주 레이블(데이터 셀을 숨김없이)
- 꼬리말, 제목, 캡션

테이블에서 행과 열을 숨김

테이블에서 숨겨진 행과 열을 표시

1. 메뉴에서 다음을 선택하십시오.

보기 > 모든 범주 표시

테이블에 숨겨진 모든 행과 열을 표시합니다. (현재 테이블의 테이블 특성에서 **비어있는 행과 열 숨김**을 선택하는 경우 빈 행 또는 빈 열 전체가 숨겨집니다.)

차원 레이블 숨김 및 표시

1. 차원 레이블이나 차원 내 범주 레이블을 선택합니다.
2. 보기 메뉴나 팝업 메뉴에서 차원 레이블 숨김 또는 차원 레이블 표시를 선택합니다.

테이블 제목 숨김 및 표시

제목 숨김

1. 피벗 테이블을 활성화합니다.
2. 제목을 선택합니다.
3. 보기 메뉴에서 숨김을 선택합니다.

숨겨진 제목 표시:

4. 보기 메뉴에서 모두 표시를 선택합니다.

테이블모양

테이블모양은 테이블의 외형을 정의하는 일련의 특성입니다. 이전에 정의된 테이블모양을 선택하거나 사용자만의 테이블모양을 작성할 수 있습니다.

- 테이블모양 적용 전과 후에, 셀 특성을 사용하여 개별 셀이나 셀 그룹의 셀 형식을 변경할 수 있습니다. 편집된 셀 형식은 새 테이블모양을 적용할 때에도 온전한 상태로 남아 있습니다.
- 또는 모든 셀을 현재 테이블모양으로 정의되는 셀 형식으로 재설정할 수 있습니다. 이 옵션은 편집된 모든 셀을 재설정합니다. 표시된 대로가 테이블모양 파일 목록에서 선택된 경우 편집된 셀은 현재 테이블 특성으로 재설정됩니다.
- 테이블 특성 대화 상자에 정의된 테이블 특성만 테이블모양에 저장됩니다. 테이블모양은 개별 셀 수정을 포함하지 않습니다.

테이블모양 적용

1. 피벗 테이블을 활성화합니다.
2. 메뉴에서 다음을 선택하십시오.

형식 > 테이블모양...

3. 파일 목록에서 테이블모양을 선택합니다. 다른 디렉토리에서 파일을 선택하려면 **찾아보기**를 클릭하십시오.
4. 확인을 클릭하여 선택한 피벗 테이블에 테이블모양을 적용합니다.

테이블모양 편집 또는 작성

1. 테이블모양 대화 상자의 파일 목록에서 테이블모양을 선택합니다.
 2. **모양 편집**을 클릭합니다.
 3. 원하는 특성에 대한 테이블 특성을 조정하고 **확인**을 클릭합니다.
 4. **모양 저장**을 클릭하여 편집한 테이블모양을 저장하거나 다른 이름으로 저장을 클릭하여 새 테이블모양으로 저장합니다.
- 테이블모양을 편집하면 선택한 피벗 테이블만 영향을 받습니다. 편집된 테이블모양은 테이블을 선택하고 테이블모양을 다시 적용하지 않는 한 해당 테이블모양을 사용하는 다른 테이블에 적용되지 않습니다.

- 테이블 특성 대화 상자에 정의된 테이블 특성만 테이블모양에 저장됩니다. 테이블모양은 개별 셀 수정을 포함하지 않습니다.

테이블 특성

테이블 특성을 사용하면 테이블의 일반 특성을 설정하고, 및 테이블의 여러 부분에 대한 셀 유형을 설정하고 할 수 있습니다. 다음을 수행할 수 있습니다.

- 빈 행이나 열 숨김과 인쇄 특성 조정 같은 일반 특성을 제어합니다.
- 꼬리말 표식 형식과 위치를 제어합니다.
- 행과 열 테이블 및 테이블의 다른 영역에 대해 데이터 영역에서 셀의 특정 형식을 결정합니다.
- 테이블의 각 영역 경계를 형성하는 선의 너비와 색상을 제어합니다.

피벗 테이블 특성 변경

1. 메뉴에서 다음을 선택하십시오.

 형식 > 테이블 특성...

2. 템(일반, 꼬리말, 셀 형식, 경계선 또는 인쇄)을 선택하십시오.

3. 원하는 옵션을 선택하십시오.

4. 확인 또는 적용을 클릭하십시오.

새 특성은 선택한 피벗 테이블에 적용됩니다.

테이블 특성: 일반

여러 특성이 테이블에 전체적으로 적용됩니다. 다음을 수행할 수 있습니다.

- 빈 행과 열을 표시하거나 숨깁니다. (빈 행이나 열에는 데이터 셀에 아무 것도 없습니다.)
- 왼쪽 위 코너에 배치하거나 중첩할 수 있는 행 테이블의 위치를 제어합니다.
- 열 너비 최대값과 최소값을 제어합니다. 포인트로 표시됩니다.

일반 테이블 특성 변경

1. 일반 템을 클릭하십시오.

2. 원하는 옵션을 선택하십시오.

3. 확인 또는 적용을 클릭하십시오.

표시할 행 수 설정

참고: 이 기능은 레거시 테이블에만 적용됩니다.

기본적으로, 행 수가 많은 테이블은 100개 행 섹션으로 표시됩니다. 테이블에 표시되는 행 수를 제어하려면 다음을 수행합니다.

1. 행별 테이블 표시를 선택하십시오.

2. 표시할 행 수 설정을 클릭하십시오.

또는

3. 활성화된 피벗 테이블의 보기 메뉴에서 행별 테이블 표시 및 표시할 행 수 설정을 선택합니다.

표시할 행 수. 한 번에 표시되는 최대 행 수를 제어합니다. 탐색 제어를 사용하여 테이블의 여러 섹션으로 이동할 수 있습니다. 최소 값은 10입니다. 기본값은 100입니다.

과부/고아 줄 공차. 테이블에서 가장 안쪽에 있는 행 차원의 최대 행 수를 제어하여 표시되는 테이블 보기 전체에 분할합니다. 예를 들어, 가장 안쪽 행 차원의 각 그룹에 6개의 범주가 있는 경우 값을 6으로 지정하면 표시된 보기에서 그룹이 분할되지 않습니다. 이렇게 설정하면 표시된 보기의 총 행 수가 표시되도록 지정된 최대 행 수를 초과합니다.

테이블 특성: 참고

테이블 특성 대화 상자의 참고 탭은 꼬리말 형식과 테이블 주석 텍스트를 제어합니다.

꼬리말. 꼬리말 표식 특성에는 텍스트와 관련된 유형과 위치가 포함됩니다.

- 꼬리말 표식의 유형은 숫자(1, 2, 3, ...) 또는 문자(a, b, c, ...)입니다.

- 꼬리말 표식은 위첨자나 아래첨자로 텍스트에 첨부됩니다.

주석 텍스트. 각 테이블에 주석 텍스트를 추가할 수 있습니다.

- 브이에서 마우스를 테이블 위에 올리면 주석 텍스트가 도구 팁에 표시됩니다.

- 테이블에 초점이 맞으면 스크린 리더가 주석 텍스트를 읽습니다.

- 브이의 도구 팁은 주석의 처음 200문자만을 표시하지만 스크린 리더는 전체 텍스트를 읽습니다.

- 출력을 HTML로 내보낼 때 주석 텍스트는 alt 텍스트로 사용합니다.

테이블 특성: 셀 형식

형식의 경우 테이블은 제목, 레이어, 코너 레이블, 행 레이블, 열 레이블, 데이터, 캡션 및 꼬리말의 영역으로 구분됩니다. 각 테이블 영역에서는 관련된 셀 형식을 수정할 수 있습니다. 셀 형식에는 텍스트 특성(글꼴, 크기, 색상, 유형 등), 수평 및 수직맞춤, 배경색 및 내부 셀 여백 등이 포함됩니다.

셀 형식은 영역(정보의 범주)에 적용됩니다. 이들은 개별 셀의 특성이 아닙니다. 이러한 차이는 테이블을 피벗할 때 중요한 사항이 됩니다.

예를 들어, 다음 함수는

- 열 레이블의 셀 형식으로 굵은 글꼴을 지정하는 경우 열 레이블은 열 차원에 현재 표시된 정보에 관계없이 굵게 나타납니다. 열 차원에서 다른 차원으로 항목을 이동하는 경우 열 레이블의 굵은 특성을 유지하지 않습니다.
- 현재 피벗 테이블의 셀을 선택하고 도구 모음에서 굵게 단추를 눌러 열 레이블을 굵게 표시하면 셀 내용은 이동하는 차원에 관계없이 계속 굵게 표시되지만, 열 차원으로 옮긴 다른 항목에 대해서는 열 레이블이 굵게 유지되지 않습니다.

셀 형식 변경

1. 셀 형식 탭을 선택합니다.
2. 드롭 다운 목록에서 영역을 선택하거나 보기 영역을 클릭하십시오.
3. 영역 특성을 선택합니다. 선택은 보기에 반영됩니다.
4. 확인 또는 적용을 클릭하십시오.

행 색상 변경

테이블의 데이터 영역에 있는 행의 배경 및/또는 텍스트 색상을 변경하려면 다음을 수행합니다.

1. 영역 드롭 다운 목록에서 데이터를 선택합니다.
2. 배경색 그룹에서 행 색상 변경을 선택합니다.
3. 행의 배경 및 텍스트를 변경하는 데 사용할 색상을 선택합니다.

행 색상 변경은 테이블의 데이터 영역에만 적용됩니다. 행 또는 열 레이블 영역에는 적용되지 않습니다.

테이블 특성: 경계선

테이블의 각 경계선 위치에서 선 유형과 색상을 선택할 수 있습니다. 유형을 없음으로 선택하면 선택한 위치에서 선이 보이지 않게 됩니다.

테이블 경계선 변경

1. 경계선 탭을 클릭하십시오.
2. 목록에서 해당 이름을 누르거나 보기 영역에서 선을 눌러 경계 위치를 선택하십시오.
3. 선 유형을 선택하거나 없음을 선택하십시오.
4. 색상을 선택하십시오.

5. 확인 또는 적용을 클릭하십시오.

테이블 특성: 인쇄

인쇄할 피벗 테이블에서 다음 특성을 제어할 수 있습니다.

- 모든 레이어나 테이블의 위쪽 레이어만 인쇄하고 각 레이어를 별도의 페이지에 인쇄합니다.
- 테이블을 인쇄할 수 있도록 가로 또는 세로로 크기를 줄여 용지 크기에 맞춥니다.
- 정의된 페이지 크기에 비해 테이블이 너무 넓거나 너무 길면 테이블의 인쇄 영역에 포함할 행과 열의 최소값을 제어하여 과부/고아 줄을 조정합니다.

참고: 위에 다른 출력이 있으므로 테이블이 너무 길어서 현재 페이지에 맞출 수 없지만 정의된 페이지 길이에는 맞는 경우에는 테이블은 과부/고아 설정과 관계없이 새 페이지에 자동으로 인쇄됩니다.

- 한 페이지에 들어가지 못하는 테이블에는 '계속' 표시를 합니다. 각 페이지의 위와 아래에 '계속' 표시를 할 수 있습니다. 어떤 옵션도 선택하지 않으면 '계속' 표시는 나타나지 않습니다.

피벗 테이블 인쇄 특성 제어

- 인쇄 탭을 클릭하십시오.
- 원하는 인쇄 옵션을 선택하십시오.
- 확인 또는 적용을 클릭하십시오.

셀 특성

셀 특성은 선택한 셀에 적용됩니다. 글꼴, 값 형식, 맞춤, 여백 및 색상을 변경할 수 있습니다. 셀 특성은 테이블 특성을 무시하므로 테이블 특성을 변경하는 경우 개별적으로 적용된 셀 특성은 변경되지 않습니다.

셀 특성 변경:

- 테이블에서 셀을 선택합니다.
- 형식 메뉴 또는 팝업 메뉴에서 셀 특성을 선택하십시오.

글꼴 및 배경

글꼴 및 배경 탭은 테이블에서 선택한 셀에 대한 글꼴 유형과 색상 및 배경 색상을 제어합니다.

형식 값

형식 값 탭은 선택한 셀에 대한 값 형식을 제어합니다. 숫자, 날짜, 시간, 통화 등에 대한 형식을 선택하고 표시할 소수점 자릿수를 조정할 수 있습니다.

맞춤 및 여백

맞춤 및 여백 탭은 값의 수평과 수직 맞춤 및 선택한 셀에 대한 위, 아래, 왼쪽, 오른쪽 여백을 제어합니다. 혼합 수평 맞춤은 유형에 따라 각 셀의 내용을 맞춥니다. 예를 들어, 날짜는 오른쪽으로 맞추고 텍스트 값은 왼쪽으로 맞춥니다.

꼬리말 및 캡션

테이블에 꼬리말과 캡션을 추가할 수 있습니다. 꼬리말이나 캡션을 숨기고, 꼬리말 표식을 변경하고, 꼬리말 번호를 재지정할 수도 있습니다.

꼬리말 및 캡션 추가

테이블에 캡션 추가

- 삽입 메뉴에서 캡션을 선택합니다.

테이블의 항목에 꼬리말을 첨부할 수 있습니다. 꼬리말 추가:

- 활성화된 피벗 테이블 내에서 제목, 셀 또는 캡션을 클릭하십시오.

2. 삽입 메뉴에서 **꼬리말**을 선택하십시오.
3. 제공된 영역에 꼬리말 텍스트를 삽입하십시오.

캡션 숨김 또는 표시

캡션 숨김

1. 캡션을 선택합니다.
2. 보기 메뉴에서 숨김을 선택합니다.

숨겨진 캡션 표시

1. 보기 메뉴에서 모두 표시를 선택합니다.

테이블에서 꼬리말 숨김 또는 표시

꼬리말 숨김

1. 꼬리말 참조를 포함하는 셀을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 **꼬리말 숨김**을 선택하십시오.

또는

2. 테이블의 꼬리말 영역에서 꼬리말을 선택하고 팝업 메뉴에서 숨김을 선택하십시오.

참고: 레거시 테이블의 경우 테이블의 꼬리말 영역을 선택하고, 팝업 메뉴에서 **꼬리말 편집**을 선택한 다음 숨기려는 꼬리말의 표시 특성을 선택 취소(지우기)하십시오.

셀에 여러 꼬리말이 포함된 경우에는 후자의 방법을 사용하여 꼬리말을 선택적으로 숨기십시오.

테이블에서 모든 꼬리말 숨김

1. 테이블의 꼬리말 영역에서 모든 꼬리말을 선택하고(클릭과 끌기 또는 Shift+클릭을 사용하여 꼬리말 선택) 보기 메뉴에서 숨김을 선택하십시오.

숨겨진 꼬리말 표시

1. 보기 메뉴에서 모든 꼬리말 표시를 선택합니다.

꼬리말 표식

꼬리말 표식은 꼬리말을 표시할 때 사용할 수 있는 문자를 변경합니다. 기본적으로 표준 꼬리말 표식은 테이블 특성 설정에 따라 순차 문자 또는 숫자입니다. 특수 표식을 할당할 수도 있습니다. 특수 표식은 꼬리말 번호를 재지정하거나 표준 표식의 숫자와 문자를 전환할 경우 영향을 받지 않습니다. 표준 표식의 번호 또는 문자 표시와 꼬리말 표식의 아래첨자 또는 위첨자 위치는 테이블 특성 대화 상자의 꼬리말 탭으로 제어됩니다.

꼬리말 표식 변경

1. 꼬리말을 선택합니다.
2. 형식 메뉴에서 **꼬리말 표식**을 선택하십시오.

특수 표식은 2문자로 제한됩니다. 특수 표식이 있는 꼬리말은 테이블의 꼬리말 영역에서 순차 문자나 숫자가 있는 꼬리말보다 앞섭니다. 따라서 특수 표식으로 변경하면 꼬리말 목록을 다시 정렬할 수 있습니다.

꼬리말에 다시 번호 매기기

행, 열, 레이어를 전환하여 테이블을 피벗하면 꼬리말 순서가 달라질 수 있습니다. 꼬리말 번호를 재지정하려면 다음을 수행합니다.

1. 형식 메뉴에서 **꼬리말에 다시 번호 매기기**를 선택하십시오.

레거시 테이블의 꼬리말 편집

레거시 테이블의 경우, 꼬리말 편집 대화 상자를 사용하여 꼬리말 텍스트와 글꼴 설정을 입력 및 수정하고 꼬리말 표식을 변경하며 선택적으로 꼬리말을 숨기거나 삭제할 수 있습니다.

레거시 테이블에 새 꼬리말을 삽입할 경우 꼬리말 편집 대화 상자가 자동으로 열립니다. 꼬리말 편집 대화 상자를 사용하여 기존 꼬리말 편집(새 꼬리말 작성하지 않음)

표식. 기본적으로 표준 꼬리말 표식은 테이블 특성 설정에 따라 순차 문자 또는 숫자입니다. 특수 표식을 지정하려면 표식 열에 새 표식 값을 입력합니다. 특수 표식은 꼬리말 번호를 재지정하거나 표준 표식의 숫자와 문자를 전환할 경우 영향을 받지 않습니다. 표준 표식의 번호 또는 문자 표시와 꼬리말 표식의 아래첨자 또는 위첨자 위치는 테이블 특성 대화 상자의 꼬리말 탭으로 제어됩니다. 자세한 정보는 [113 페이지의 『테이블 특성: 참고』](#)의 내용을 참조하십시오.

특수 표식을 다시 표준 표식으로 변경하려면 꼬리말 편집 대화 상자에서 표식을 마우스 오른쪽 단추로 클릭하고, 팝업 메뉴에서 **꼬리말 표식**을 선택하고 꼬리말 표식 대화 상자에서 표준 표식을 선택하십시오.

꼬리말. 꼬리말의 내용입니다. 표시는 현재 글꼴 및 배경 설정을 반영합니다. 글꼴 설정은 형식 하위 대화 상자를 사용하여 개별 꼬리말에 대해 변경할 수 있습니다. 자세한 정보는 [116 페이지의 『꼬리말 글꼴 및 색상 설정』](#)의 내용을 참조하십시오. 하나의 배경 색상이 모든 꼬리말에 적용되고 셀 특성 대화 상자의 글꼴 및 배경 탭에서 변경할 수 있습니다. 자세한 정보는 [114 페이지의 『글꼴 및 배경』](#) 주제를 참조하십시오.

표시. 모든 꼬리말은 기본적으로 표시됩니다. 꼬리말을 숨기려면 표시 확인 상자를 선택 취소(지우기)하십시오.

꼬리말 글꼴 및 색상 설정

레거시 테이블의 경우, 형식 대화 상자를 사용하여 하나 이상의 선택한 꼬리말에 대해 글꼴군, 유형, 크기 및 색상을 변경할 수 있습니다.

1. 꼬리말 편집 대화 상자에서 꼬리말 눈금에 있는 하나 이상의 꼬리말을 선택(클릭)합니다.
2. 형식 단추를 클릭하십시오.

선택한 글꼴군, 유형, 크기 및 색상이 선택한 모든 꼬리말에 적용됩니다.

배경색, 맞춤 및 여백은 셀 특성 대화 상자에서 설정하여 모든 꼬리말에 적용할 수 있습니다. 이러한 설정은 개별 꼬리말에 대해 변경할 수 없습니다. 자세한 정보는 [114 페이지의 『글꼴 및 배경』](#) 주제를 참조하십시오.

데이터 셀 너비

데이터 셀 너비 설정은 모든 데이터 셀 너비를 동일하게 하는 데 사용합니다.

모든 데이터 셀의 너비 설정

1. 메뉴에서 다음을 선택하십시오.
형식 > 데이터 셀 너비 설정...
2. 셀 너비 값을 입력합니다.

열 너비 변경

1. 열 경계를 클릭하고 끕니다.

피벗 테이블에서 숨겨진 경계선 표시

경계선이 보이지 않는 테이블에서는 숨겨진 경계선을 표시할 수 있습니다. 이렇게 하면 열 너비 변경 같은 작업을 쉽게 수행할 수 있습니다.

1. 보기 메뉴에서 **격자선**을 선택하십시오.

피벗 테이블에서 행, 열과 셀을 선택

전체 행이나 열 또는 지정된 데이터 세트 및 테이블 셀을 선택할 수 있습니다.

여러 셀 선택

선택 > 데이터 및 테이블 셀

피벗 테이블 인쇄

여러 요소가 인쇄된 피벗 테이블이 표시되는 방식에 영향을 미칠 수 있으며 이러한 요소는 피벗 테이블 속성을 변경하여 제어할 수 있습니다.

- 다차원 피벗 테이블(레이어가 있는 테이블)의 경우 모든 레이어를 인쇄하거나 맨 위(표시된) 레이어만 인쇄할 수 있습니다. 자세한 정보는 [114 페이지의 『테이블 특성: 인쇄』](#)의 내용을 참조하십시오.
- 길거나 넓은 피벗 테이블의 경우 페이지에 맞게 테이블을 자동으로 크기 조정하거나 페이지 구분과 테이블 구분 위치를 제어할 수 있습니다. 자세한 정보는 [114 페이지의 『테이블 특성: 인쇄』](#)의 내용을 참조하십시오.

파일 메뉴의 인쇄 미리보기를 사용하여 인쇄된 피벗 테이블의 모습을 확인하십시오.

넓거나 긴 테이블에 대한 테이블 구분 제어

정의된 페이지 크기 내에 인쇄하기에는 너무 넓거나 긴 피벗 테이블은 자동으로 분할되고 다중 섹션에 인쇄됩니다. 다음을 수행할 수 있습니다.

- 큰 테이블이 분할되는 행과 열 위치를 제어합니다.
- 테이블을 분할할 때 분할되지 않도록 해야 하는 행과 열을 지정합니다.
- 정의된 페이지 크기에 맞게 큰 테이블의 크기를 조정합니다.

인쇄된 피벗 테이블을 위해 행과 열 구분을 지정합니다.

- 피벗 테이블을 활성화합니다.
- 구분을 삽입하려는 위치 왼쪽의 열에서 셀을 클릭하거나 구분을 삽입하려는 행 앞의 행의 셀을 클릭하십시오.
- 메뉴에서 다음을 선택하십시오.

형식 > 중단점 > 수직 중단점

또는

형식 > 중단점 > 수평 중단점

- 피벗 테이블을 활성화합니다.
- 구분을 삽입하려는 위치 왼쪽의 열에서 셀을 클릭하거나 구분을 삽입하려는 행 앞의 행의 셀을 클릭하십시오.
- 메뉴에서 다음을 선택하십시오.

형식 > 중단점 > 수직 중단점

또는

형식 > 중단점 > 수평 중단점

연결할 행이나 열을 지정:

- 연결 할 행 레이블이나 열 레이블을 선택합니다. 여러 행이나 열 레이블을 선택하려면 클릭한 후 끌거나 Shift +클릭하십시오.
- 메뉴에서 다음을 선택하십시오.

형식 > 중단점 > 연결

중단점을 보고 그룹을 연결:

- 메뉴에서 다음을 선택하십시오.

형식 > 중단점 > 중단점 표시

중단점은 수직선 또는 수평선으로 나타납니다. 그룹 연결은 어두운 경계선으로 둘러싸인 회색 직사각형 영역으로 나타납니다.

참고: 중단점 표시와 그룹 연결은 레거시 테이블에서는 지원되지 않습니다.

중단점 및 연결 그룹 지우기

중단점 지우기

- 수직 중단점 왼쪽 열의 셀을 클릭하거나 수평 중단점 위 행의 셀을 클릭하십시오.

2. 메뉴에서 다음을 선택하십시오.

형식 > 중단점 > 중단점 또는 그룹 지우기

연결 그룹 지우기

3. 그룹을 지정하는 열 레이블 또는 행 레이블을 선택합니다.

4. 메뉴에서 다음을 선택하십시오.

형식 > 중단점 > 중단점 또는 그룹 지우기

모든 중단점 및 연결 그룹은 행이나 열을 피벗하거나 다시 정렬할 때 자동으로 지워집니다. 이 동작은 레거시 테이블에 적용되지 않습니다.

피벗 테이블에서 차트 작성

1. 피벗 테이블을 두 번 클릭하여 이를 활성화하십시오.

2. 차트에 표시할 행, 열 또는 셀을 선택하십시오.

3. 선택한 영역의 임의의 위치에 마우스 오른쪽 단추로 클릭하십시오.

4. 팝업 메뉴에서 그래프 작성을 선택하고 차트 유형을 선택하십시오.

레거시 테이블

테이블을 레거시 테이블(릴리스 19에서 완전히 기능하는 테이블로 알려짐)로서 렌더링하기로 선택할 수 있고 그러면 이는 20 이전의 IBM SPSS Statistics 릴리스와 완전히 호환 가능합니다. 레거시 테이블은 느리게 렌더링할 수 있으며 20 이전의 릴리스와 호환성이 필요한 경우에만 권장됩니다. 레거시 테이블을 작성하는 방법에 대한 자세한 정보는 [119 페이지의 『피벗 테이블 옵션』](#)의 내용을 참조하십시오.

옵션

옵션

옵션은 여러 설정을 제어합니다.

옵션 설정 변경

1. 메뉴에서 다음을 선택합니다.

편집 > 옵션...

2. 변경할 설정의 탭을 클릭합니다.

3. 설정을 변경합니다.

4. 확인 또는 적용을 클릭합니다.

일반 옵션

최대 스레드 수

결과를 계산할 때 멀티스레드 프로시저가 사용할 스레드 수입니다. 자동 설정은 사용 가능한 처리 코어 수를 기준으로 합니다. 멀티스레드 프로시저가 실행되는 동안 다른 애플리케이션에서 더 많은 처리 지원을 사용할 수 있도록 하려면 낮은 값을 지정하십시오. 분산 분석 모드에서는 이 옵션을 사용할 수 없습니다.

출력

소수점 값에 선행 0 표시 - 소수점 부분으로만 구성된 숫자 값에 선행 0을 표시합니다. 예를 들어 선행 0을 표시할 경우 값 .123은 0.123으로 표시됩니다. 통화 또는 퍼센트 형식의 숫자 값에는 이 설정이 적용되지 않습니다. 데이터를 외부 파일에 저장할 때 선행 0은 포함되지 않습니다. 단, 고정 ASCII 파일 (*.dat)은 예외입니다.

측정 시스템. 피벗 테이블 인쇄 작업에서 셀 여백, 셀 너비, 테이블 간 간격 등의 속성을 지정하는 데 사용하는 측정 시스템(포인트, 인치, 센티미터)입니다.

뷰어 옵션

뷰어 출력 표시 옵션은 설정을 변경한 후에 생성되는 새 출력에만 영향을 미칩니다. 뷰어에 이미 표시된 출력은 이러한 설정의 변경에 영향을 받지 않습니다.

초기 출력결과 상태

프로시저를 실행할 때마다 어떤 항목이 자동으로 표시되거나 숨겨지는지와 항목이 처음에 정렬되는 방법을 제어합니다. 다음 항목의 표시를 제어할 수 있습니다. 로그, 경고, 참고, 제목, 피벗 테이블, 차트, 트리 다이어그램 및 텍스트 출력입니다. 로그에서 명령의 표시를 켜거나 끌 수도 있습니다. 로그에서 명령 구문을 복사하고 이를 구문 파일에 저장할 수 있습니다.

참고: 모든 결과 항목은 뷰어에서 왼쪽 맞춤으로 표시됩니다. 인쇄된 출력의 정렬만 맞춤 설정으로 영향을 받습니다. 가운데 맞춤과 오른쪽 맞춤 항목은 작은 기호로 식별됩니다.

제목

새 결과 제목의 글꼴 유형, 크기, 색상 등을 제어합니다. 글꼴 크기 목록은 일련의 사전 정의된 크기를 제공하지만, 지원되는 다른 크기 값을 수동으로 입력할 수 있습니다.

페이지 제목

새 페이지 제목 및 TITLE 및 SUBTITLE 명령 구문을 통해 생성되거나 삽입 메뉴의 새 페이지 제목을 통해 작성된 페이지 제목의 글꼴 유형, 크기 및 색상을 제어합니다. 글꼴 크기 목록은 일련의 사전 정의된 크기를 제공하지만, 지원되는 다른 크기 값을 수동으로 입력할 수 있습니다.

텍스트 출력

텍스트 출력에 사용되는 글꼴입니다. 텍스트 출력에는 단일자간(고정 피치) 글꼴을 사용하는 것이 적합합니다. 비례 글꼴을 선택하면 테이블 출력은 제대로 정렬되지 않습니다. 글꼴 크기 목록은 일련의 사전 정의된 크기를 제공하지만, 지원되는 다른 크기 값을 수동으로 입력할 수 있습니다.

기본 페이지 설정

인쇄의 출력방향과 여백을 위한 기본 옵션을 제어합니다.

피벗 테이블 옵션

피벗 테이블 옵션은 피벗 테이블의 표시에 대해 다양한 옵션을 설정합니다.

테이블모양

파일 목록에서 테이블모양을 선택하고 확인 또는 적용을 클릭하십시오. IBM SPSS Statistics에서 제공하는 테이블모양 중 하나를 사용하거나, 형식 메뉴의 테이블모양을 선택하여 피벗 테이블 편집기에서 테이블모양을 직접 작성할 수 있습니다.

- **찾아보기.** 다른 디렉토리에서 테이블모양을 선택할 수 있습니다.
- **테이블모양 디렉토리 설정.** 기본값 테이블모양 디렉토리를 변경할 수 있습니다. 찾아보기를 사용하여 사용할 디렉토리를 탐색하고 해당 디렉토리에서 테이블모양을 선택한 다음 테이블모양 디렉토리 설정을 선택합니다.

참고: IBM SPSS Statistics의 이전 버전에서 작성된 테이블모양은 버전 16.0 이상에서는 사용할 수 없습니다.

열 너비

이러한 옵션은 피벗 테이블에서 열 너비의 자동 조절을 제어합니다.

- **레이블만 수정.** 열 너비를 열 레이블 너비에 맞춥니다. 이 프로시저를 사용하면 간결한 테이블이 생성되지만 레이블보다 넓은 데이터 값은 잘릴 수 있습니다.
- **모든 테이블에 대해 레이블과 데이터 수정.** 열 너비를 열 레이블 또는 가장 큰 데이터 값 중 더 큰 값으로 조정합니다. 이 프로시저를 사용하면 테이블 너비가 더 넓지만 모든 값이 빠짐없이 표시됩니다.

기본 편집 모드

이 옵션은 뷰어 창이나 별도의 창에서 피벗 테이블의 활성화를 제어합니다. 기본적으로 피벗 테이블을 두 번 클릭하면 매우 큰 테이블을 제외하고 뷰어 창에서 테이블이 활성화됩니다. 개별 창에서 피벗 테이블을 활성화할 것을 선택하거나 뷰어 창에서 피벗 테이블을 크게 열거나 개별 창에서 피벗 테이블을 크게 열도록 하는 크기 설정을 선택할 수 있습니다.

넓은 테이블을 서식있는 텍스트 형식으로 클립보드에 복사

피벗 테이블을 Word/RTF 형식으로 붙여넣을 경우 문서 너비보다 큰 테이블은 문서 너비에 맞게 줄바꿈 또는 축소되거나 변경되지 않고 그대로 있게 됩니다.

출력 옵션

출력 옵션은 출력 옵션 수의 기본 설정을 제어합니다.

스크린 리더 내게 필요한 옵션. 스크린 리더가 피벗 테이블과 열 레이블을 읽는 방법을 제어합니다. 각 데이터 셀의 전체 행과 열 레이블을 읽거나 테이블에서 데이터 셀 간에 이동할 때 변경하는 레이블만 읽을 수 있습니다.

제 8 장 결측값 처리

결측값의 개요

데이터 마이닝의 데이터 준비 단계 동안 종종 데이터에서 결측값을 대체하려고 할 수 있습니다. 결측값은 알려지지 않고, 수집되지 않거나 잘못 입력된 데이터 세트의 값입니다. 일반적으로 이러한 값은 필드에 유효하지 않습니다. 예를 들어, 성별 필드에는 M 및 F 값이 포함되어야 합니다. 필드에 Y 또는 Z 값이 있다면 이를 유효하지 않은 값으로 가정하여 공백으로 해석해야 합니다. 마찬가지로 연령 필드의 음수값은 의미가 없으므로 공백으로 해석되어야 합니다. 때때로 무응답을 보여주기 위해 질문지에 이렇게 명백하게 잘못된 값을 고의로 입력하거나 필드를 공백으로 두기도 합니다. 때로는 이러한 공백을 자세히 조사하여 누군가가 나이 제공을 거부하는 등 무응답한 것이 특정 결과를 예측하는 데 있어서 요인인지 여부를 판별하려고 할 수 있습니다.

몇몇 모델링 기술은 다른 기술보다 결측 데이터를 보다 잘 처리합니다. 예를 들어, C5.0 및 Apriori는 유형 노드에서 "결측"으로 명시적으로 선언된 값에 잘 대처합니다. 다른 모델링 기법은 결측값을 다루는 데 문제가 있고 학습 시간이 길고 정확성이 떨어지는 모델을 생성합니다.

IBM SPSS Modeler에 의해 인지되는 결측값에는 몇 가지 유형이 있습니다.

- **널 또는 시스템 결측값.** 이들은 데이터베이스나 소스 파일에 공백으로 남겨졌고 소스 또는 유형 노드에서 "결측"으로 정의되지 않은 문자열이 아닌 값입니다. 시스템 결측값은 **\$null\$**로 표시됩니다. 빈 문자열은 특정 데이터베이스에 의해 널로 처리되더라도 IBM SPSS Modeler에서는 널로 간주되지 않음을 유의하십시오.
- **빈 문자열 및 공백.** 빈 문자열 값과 공백(눈에 보이는 문자가 없는 문자열)은 널값과는 별개로 처리됩니다. 빈 문자열은 대부분의 경우에서 공백과 동일하게 처리됩니다. 예를 들어, 소스나 유형 노드에서 공백을 공란으로 처리하는 옵션을 선택한 경우 이 설정은 빈 문자열에도 적용됩니다.
- **공백 또는 사용자 정의 결측값.** 이들은 소스 노드 또는 유형 노드에서 결측으로 명백하게 정의되어 있는 unknown, 99 또는 -1 등과 같은 값입니다. 또는 널과 공백을 공란으로 처리하기로 선택할 수도 있는데 그러면 이들은 특수 처리용으로 플래그가 지정되고 대부분의 계산에서 제외됩니다. 예를 들어, @BLANK 함수를 사용하여 이들 값 및 다른 유형의 결측값을 공란으로 처리할 수 있습니다.

혼합 데이터 읽기. 숫자 저장 공간(정수, 실수, 시간, 시간소인 또는 날짜)이 있는 필드를 읽을 때 숫자가 아닌 값은 널 또는 시스템 결측값으로 설정됨을 유의하십시오. 이는 몇몇 애플리케이션과는 달리 필드 내에서 혼합된 저장 유형을 허용하지 않기 때문입니다. 이를 피하려면 혼합 데이터가 있는 필드는 소스 노드나 외부 애플리케이션에서 필요에 따라 저장 유형을 변경하여 문자열로 읽어야 합니다.

Oracle에서 빈 문자열 읽기. Oracle 데이터베이스에서 읽거나 Oracle 데이터베이스에 쓸 때 IBM SPSS Modeler 및 대부분의 다른 데이터베이스와 달리 Oracle은 빈 문자열 값을 널값과 같다고 간주하고 저장합니다. 이는 Oracle 데이터베이스에서 추출된 동일한 데이터가 파일 또는 다른 데이터베이스에서 추출된 경우와 다르게 동작하고 데이터가 다른 결과를 리턴할 수 있음을 의미합니다.

결측값 처리

사용자의 비즈니스나 도메인 지식에 비추어 결측값을 처리하는 방법을 결정해야 합니다. 학습 시간을 줄이고 정확도를 높이기 위해 데이터 세트에서 공백을 제거하려고 할 수 있습니다. 반면, 공백값이 있으면 새로운 비즈니스 기회나 추가적인 통찰로 이어질 수도 있습니다. 최상의 기법을 선택하려면 데이터의 다음 측면을 고려해야 합니다.

- 데이터 세트의 크기
- 공백을 포함하는 필드 수
- 결측 정보의 양

일반적으로 두 가지의 방법을 따를 수 있습니다.

- 결측값이 있는 필드나 레코드를 제외할 수 있습니다.
- 다양한 방법을 사용하여 결측값을 대치, 바꾸거나 강요할 수 있습니다.

이러한 두 방법은 모두 데이터 검토 노드를 사용하여 대부분 자동화할 수 있습니다. 예를 들어, 결측값이 너무 많은 필드가 모델링에서 유용하게 되도록 이를 제외하는 필터 노드를 생성하거나 남아 있는 필드의 일부 또는 모두의 결측값을 대체하는 수퍼노드를 생성할 수 있습니다. 이는 감사의 실제 힘이며 이를 통해, 데이터의 현재 상태를 평가할 수 있을 뿐만 아니라 평가를 기반으로 조치를 취할 수도 있습니다.

결측값이 있는 레코드 처리

대다수의 결측값이 소수의 레코드에 집중되면 이러한 레코드를 제외하기만 하면 됩니다. 예를 들어, 은행은 일반적으로 대출 고객에 대해서 자세하고 완전한 레코드를 유지합니다. 그러나 은행이 자체 직원에게 대출을 승인할 때는 덜 제한적이라면 직원 대출에 대해 수집한 데이터에는 몇몇 공백 필드가 있을 수 있습니다. 이런 경우에는 이러한 결측값을 다루기 위한 두 가지 옵션이 있습니다.

- 선택 노드를 사용하여 직원 레코드를 제거할 수 있습니다.
- 데이터 세트가 크면 공백이 있는 모든 레코드를 삭제할 수 있습니다.

결측값이 있는 필드 처리

결측값의 대부분이 소수의 필드에 집중되면 이들을 레코드 수준이 아니라 필드 수준에서 해결할 수 있습니다. 이 방법을 사용하면 결측값 처리를 위한 방법을 결정하기 전에 특정 필드의 상대적 중요성을 실험할 수 있습니다. 필드가 모델링에서 중요하지 않은 경우에는 결측값이 얼마나 많은지와는 관계없이 유지할 가치가 없을 수 있습니다.

예를 들어, 시장 조사 회사는 50개의 질문을 포함하는 일반 질문지에서 데이터를 수집할 수 있습니다. 질문 중 두 개는 연령과 정치적 신념 등 대다수의 사람이 응답을 꺼려하는 정보를 다룹니다. 이 경우, *Age* 및 *Political_persuasion*에는 결측값이 많습니다.

필드 측정 수준

사용할 방법을 결정할 때 결측값이 있는 필드의 측정 수준 또한 고려해야 합니다.

수치 필드. 연속과 같은 수치 필드 유형의 경우에는 수치 필드에 공백이 포함되면 대다수의 모델이 작동하지 않으므로 모델을 작성하기 전에 항상 수치가 아닌 값을 제거해야 합니다.

범주형 필드. 명목 및 플래그 등과 같은 범주형 필드의 경우 결측값을 변경하는 것은 필수는 아니지만 모델의 정확도를 높여줍니다. 예를 들어, 성별 필드를 사용하는 모델은 *Y*와 *Z* 같은 의미없는 값이 있어도 잘 작동하지만 *M*과 *F*를 제외한 모든 값을 제거하면 모델의 정확도가 증가합니다.

필드 선별 또는 제거

결측값이 너무 많은 필드를 선별하는 데에는 몇몇 옵션이 있습니다.

- 데이터 검토 노드를 사용하여 품질을 기반으로 필드를 필터링할 수 있습니다.
- 필드선택 노드를 사용하여 지정된 결측값 비율보다 높은 필드를 선별하고 지정된 목표와 관련된 중요도를 기반으로 필드의 등급을 지정할 수 있습니다.
- 필드를 제거하는 대신에 유형 노드를 사용하여 필드 역할을 **없음**으로 설정할 수 있습니다. 그러면 필드가 데이터 세트에 보관되지만 모델링 프로세스에서는 제외됩니다.

시스템 결측값이 있는 레코드 처리

시스템 결측값이란?

시스템 결측값은 알려지지 않거나 적용되지 않는 데이터 값을 나타냅니다. 데이터베이스에서 이러한 값은 종종 **NULL** 값으로 언급됩니다.

시스템 결측값은 공백값과는 다릅니다. 공백값은 일반적으로 유형 노드에 특정 값 또는 값의 범위로 정의되어 있고 사용자 정의 결측으로 간주될 수 있습니다. 공백값은 모델링 컨텍스트에서 다르게 처리됩니다.

시스템 결측값 구성

시스템 결측값은 데이터 소스에서 읽혀지는 데이터에 있을 수 있습니다(예: 데이터베이스 테이블에 **NULL** 값이 포함될 수 있음).

시스템 결측값은 표현식에서 **undef** 값을 사용하여 구성될 수 있습니다. 예를 들어, 다음 CLEM 표현식은 30 이하의 경우 연령을 리턴하고 30보다 큰 경우 결측값을 리턴합니다.

```
if Age > 30 then undef else Age endif
```

결측값은 또한 외부 결합이 수행될 때, 숫자가 0으로 나눠질 때, 음수의 제곱근이 계산될 때 및 기타 상황에서도 작성될 수 있습니다.

시스템 결측값 표시

시스템 결측값은 테이블과 다른 출력에 \$null\$로 표시됩니다.

시스템 결측값 검정

인수 값이 시스템 결측값인 경우 true를 리턴하도록 특수 함수 **@NULL**을 사용하십시오. 예:

```
if @NULL(MyFieldName) then 'It is null' else 'It is not null' endif
```

함수에 전달된 시스템 결측값

함수에 전달된 시스템 결측값은 일반적으로 출력에 결측값을 전파합니다. 예를 들어, f1 필드 값이 특정 행에서 시스템 결측값이면 log(f1) 표현식 또한 해당 행을 시스템 결측값으로 평가합니다. 예외는 **@NULL** 함수입니다.

산술 연산자와 관련된 표현식에서 시스템 결측값

시스템 결측값을 포함하는 값에 산술 연산자를 적용하면 시스템 결측값이 생깁니다. 예를 들어, f1 필드 값이 특정 행에서 시스템 결측값이면 f1 + 10 표현식 또한 해당 행을 시스템 결측값으로 평가합니다.

논리 연산자와 관련된 표현식에서 시스템 결측값

논리 연산자와 관련된 표현식에서 시스템 결측값에 대해 작업할 때 3차 값 논리(true, false 및 missing)의 규칙이 적용되고 진리표에서 설명될 수 있습니다. *not*, *and* 및 *or*의 공통 논리 연산자에 대한 진리표는 다음 테이블에 표시됩니다.

표 5. NOT에 대한 진리표	
피연산자	NOT 피연산자
참	거짓
거짓	참
결측	결측

Operand1	Operand2	Operand1 AND Operand2
참	참	참
참	거짓	거짓
참	결측	결측
거짓	참	거짓
거짓	거짓	거짓
거짓	결측	거짓
결측	참	결측

표 6. AND에 대한 진리표 (계속)

Operand1	Operand2	Operand1 AND Operand2
결측	거짓	거짓
결측	결측	결측

표 7. OR에 대한 진리표

Operand1	Operand2	Operand1 OR Operand2
참	참	참
참	거짓	참
참	결측	참
거짓	참	참
거짓	거짓	거짓
거짓	결측	결측
결측	참	참
결측	거짓	결측
결측	결측	결측

비교 연산자와 관련된 표현식에서 시스템 결측값

시스템 결측값과 비시스템 결측값을 비교할 때 출력은 참 또는 거짓 결과 대신에 시스템 결측값으로 평가합니다. 시스템 결측값은 서로 비교할 수 있습니다. 두 개의 시스템 결측값은 동일한 것으로 간주됩니다.

if/then/else/endif 표현식에서 시스템 결측값

조건식을 사용할 때 조건식이 시스템 결측값을 리턴하면 else 절의 값이 조건식에서 리턴됩니다.

선택 노드의 시스템 결측값

특정 레코드의 경우 선택 표현식이 결측값으로 평가되면 레코드는 선택 노드의 출력이 아닙니다. 이 조치는 포함 및 삭제 모드 둘 모두에 적용됩니다.

병합 노드의 시스템 결측값

키를 사용하여 병합할 때 키 필드에 시스템 결측값이 있는 모든 레코드는 병합되지 않습니다.

통합의 시스템 결측값

열에서 데이터를 통합할 때 결측값은 계산에 포함되지 않습니다. 예를 들어, 3개의 값 {1, 2 및 undef}가 있는 열에서 열의 값 합계는 3으로 계산되고, 평균 값은 1.5로 계산됩니다.

결측값 대치 또는 채우기

결측값이 몇 개밖에 없는 경우에는 공백을 대체하기 위해 값을 삽입하는 것이 유용할 수도 있습니다. 이를 데이터 검토 보고서에서 수행할 수 있고 이를 통해 특정 필드에 옵션을 적합하게 지정한 다음 여러 방법을 사용하여 값을 대치하는 SuperNode를 생성할 수 있습니다. 이는 가장 유용한 방법이고 단일 노드에서 대량 필드의 처리를 지정 할 수도 있습니다.

결측값을 대치하기 위해 다음 방법을 사용할 수 있습니다.

고정됨. 고정된 값을 대체합니다(필드 평균, 범위의 중심점 또는 사용자가 지정하는 상수).

변량. 보통 또는 균일 분포를 기반으로 변량 값을 대체합니다.

표현식. 사용자 정의 표현식을 지정할 수 있습니다. 예를 들어, 전역값 설정 노드에 의해 생성된 글로벌 변수로 값을 대체할 수 있습니다.

알고리즘. C&RT 알고리즘을 기반으로 모델에 의해 예측된 값을 대체합니다. 이 방법을 사용하여 대체된 각 필드의 경우, 공백과 널을 모델에 의해 예측된 값으로 대체하는 채움 노드와 함께 별도의 C&RT 모델이 있습니다. 그러면 필터 노드가 모델에 의해 생성된 예측 필드를 제거하는 데 사용합니다.

또는 특정 필드의 값을 강제 실행하려면 유형 노드를 사용하여 필드 유형이 합법적인 값만을 다루는지 확인한 다음 공백값을 대체해야 하는 필드의 선택 열을 **강제 실행**으로 설정하십시오.

CLEM 함수 결측값

결측값을 처리하는 데에는 몇 가지 함수를 사용합니다. 다음 함수는 종종 결측값을 버리거나 채우기 위해 선택 및 채움 노드에 사용합니다.

- `count_nulls(LIST)`
- `@BLANK(FIELD)`
- `@NULL(FIELD)`
- `undef`

`@` 함수는 `@FIELD` 함수와 함께 사용하여 하나 이상의 필드에서 공백이나 널값을 식별할 수 있습니다. 필드는 공백이나 널값이 있을 때 간단히 플래그를 지정하거나 대체 값으로 채우거나 다양한 다른 작업에 사용할 수 있습니다.

다음과 같이 필드 목록 전체에 걸쳐 널을 계산할 수 있습니다.

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

필드 목록을 입력으로 받아들이는 함수를 사용할 때 다음 예제에 표시된 대로 특수 함수 `@FIELDS_BETWEEN` 및 `@FIELDS_MATCHING`을 사용할 수 있습니다.

```
count_nulls(@FIELDS_MATCHING('card*'))
```

`undef` 함수를 사용하여 **\$null\$**로 표시되는 시스템 결측값이 있는 필드를 채울 수 있습니다. 예를 들어, 숫자 값을 대체하기 위해 다음과 같이 조건문을 사용할 수 있습니다.

```
if not(Age > 17) or not(Age < 66) then undef else Age endif
```

이는 **\$null\$**로 표시되는 시스템 결측값이 있는 범위가 아닌 모든 것을 대체합니다. `not()` 함수를 사용하여 음수를 포함하여 다른 모든 숫자 값을 포착할 수 있습니다. 자세한 정보는 [173 페이지의 『공백 및 널값 처리 함수』](#)의 내용을 참조하십시오.

레코드 삭제에 대한 참고

선택 노드를 사용하여 레코드를 제거할 때 구문이 3값 논리를 사용하고 선택 명령문에 널값을 자동으로 포함함을 유의하십시오. 널값(시스템 결측값)을 선택 표현식에서 제외하려면 표현식에서 `and not`을 사용하여 이를 명시적으로 지정해야 합니다. 예를 들어, 처방약의 유형이 Drug C인 모든 레코드를 선택하고 포함하려면 다음 선택 명령문을 사용합니다.

```
Drug = 'drugC' and not(@NULL(Drug))
```

이러한 상황에서 제외된 널값의 이전 버전

제 9 장 CLEM 표현식 작성

CLEM 정보

CLEM(Control Language for Expression Manipulation)은 IBM SPSS Modeler 스트림을 따라 플로우하는 데이터를 분석하고 조작하기 위한 강력한 언어입니다. 데이터 마이너는 비용과 수입 데이터에서 수익을 파생하는 것과 같은 단순한 작업 또는 웹 로그 데이터를 유용한 정보가 있는 필드와 레코드 세트로 변환하는 것과 같은 복잡한 작업을 수행하기 위해 스트림 작업에서 CLEM을 집중적으로 사용합니다.

CLEM은 IBM SPSS Modeler 내에서 다음을 위해 사용합니다.

- 레코드 필드에서 조건을 비교하고 평가합니다.
- 새 필드의 값을 파생합니다.
- 기존 필드의 새 값을 파생합니다.
- 레코드 시퀀스에 대한 원인.
- 데이터를 레코드에서 보고서에 삽입합니다.

CLEM 표현식은 IBM SPSS Modeler에서 데이터 준비에는 필수적이며 레코드와 필드 작업(선택, 균형, 채움)부터 차트와 출력(분석, 보고서, 테이블)까지 광범위한 노드에서 사용할 수 있습니다. 예를 들어, 파생 노드에서 CLEM을 사용하여 비율과 같은 수식을 기반으로 새 필드를 작성할 수 있습니다.

CLEM 표현식은 또한 글로벌 검색과 바꾸기 작업에 사용할 수도 있습니다. 예를 들어, @NULL(@FIELD) 표현식은 채움 노드에서 시스템 결측값을 정수값 0으로 바꾸는 데 사용할 수 있습니다. (공란이라고도 불리는 사용자 결측값을 바꾸려면 @BLANK 함수를 사용하십시오.)

보다 복잡한 CLEM 표현식 또한 작성할 수 있습니다. 예를 들어, 다음 표현식을 사용하여 작성된 새 값 범주 등과 같이 규칙 조건 세트를 기반으로 새 필드를 파생시킬 수 있습니다. If: CardID = @OFFSET(CardID, 1), Then: @OFFSET(ValueCategory, 1), Else: 'exclude'

이 예제는 @OFFSET 함수를 사용하여 다음과 같이 나타냅니다. "주어진 레코드의 CardID 필드 값이 이전 레코드와 같은 경우에는 이전 레코드의 이름이 ValueCategory인 필드의 값을 리턴합니다. 그렇지 않으면 문자열 "exclude"를 지정합니다. 즉, 인접 레코드의 CardID가 동일하면 동일한 값 범주가 지정되어야 합니다. (제외 문자열이 있는 레코드는 나중에 선택 노드를 사용하여 모을 수 있습니다.)

CLEM 예제

CLEM으로 가능한 표현식 유형은 물론 올바른 구문을 설명하기 위해 예제 표현식이 이어집니다.

단순 표현식

수식은 After 및 Before 필드 값을 기반으로 새 필드를 파생시키는 이 수식만큼 단순할 수 있습니다.

```
(After - Before) / Before * 100.0
```

필드 값을 참조할 때 필드 이름에는 따옴표가 없습니다.

마찬가지로 다음 표현식은 salary 필드 각 값의 로그를 리턴하기만 합니다.

```
log(salary)
```

복합 표현식

또한 표현식은 길고 더 복잡할 수 있습니다. 다음 표현식은 두 필드(\$KX-Kohonen 및 \$KY-Kohonen) 값이 지정된 범위 내에 있을 경우 참을 리턴합니다. 필드 이름에 특수 문자가 포함되어 있기 때문에 여기서 필드 이름에는 작은따옴표가 있습니다.

```
('$KX-Kohonen' >= -0.2635771036148072 and '$KX-Kohonen' <= 0.3146203637123107
```

```
and '$KY-Kohonen' >= -0.18975617885589602 and  
'$KY-Kohonen' <= 0.17674794197082522) -> T
```

문자열 함수와 같은 몇 개의 함수에는 올바른 구문을 사용하여 몇 개의 모수를 입력해야 합니다. 다음 예제에서 `subscrs` 함수는 `produce_ID` 필드의 첫 번째 문자를 리턴하는데 사용되어 항목이 유기농인지 유전자 변형인지 일반인지 여부를 표시합니다. 표현식 결과는 `-> 'result'`로 설명됩니다.

```
subscrs(1,produce_ID) -> 'c'
```

마찬가지로 다음 표현식이 있습니다.

```
stripchar('3','123') -> '12'
```

문자는 단일 역따옴표 안에서 항상 캡슐화됨에 유의하는 것이 중요합니다.

표현식에서 함수 조합

CLEM 표현식은 함수 조합으로 구성되는 경우가 빈번합니다. 다음 표현식은 `subscr` 및 `lowertoupper`를 조합하여 `produce_ID`의 첫 번째 문자를 리턴하고 이를 대문자로 변환합니다.

```
lowertoupper(subscr(1,produce_ID)) -> 'C'
```

이 동일한 표현식을 간단하게 다음과 같이 작성할 수 있습니다.

```
lowertoupper(produce_ID(1)) -> 'C'
```

일반적으로 사용하는 또 다른 함수 조합은 다음과 같습니다.

```
locchar_back('n', (length(web_page)), web_page)
```

이 표현식은 필드 값의 마지막 문자부터 거꾸로 읽어 `web_page` 필드 값에 'n' 문자를 위치시킵니다. `length` 함수도 포함시켜 이 표현식은 7과 같은 정적 숫자(7자 미만인 값의 경우 올바르지 않음)를 사용하지 않고 현재 값의 길이를 동적으로 계산합니다.

특수 함수

수많은 특수 함수(앞에 @ 기호가 붙음)를 사용할 수 있습니다. 일반적으로 사용하는 함수는 다음과 같습니다.

```
@BLANK('referrer ID') -> T
```

특수 함수를 조합에서 사용하는 경우가 빈번한데, 한 번에 두 개 이상의 필드에서 공백에 플래그를 표시하는 일반적으로 사용하는 방법입니다.

```
@BLANK(@FIELD) -> T
```

추가 예제가 CLEM 문서 전반에 설명되어 있습니다. 자세한 정보는 [143 페이지의 『CLEM 참조 개요』](#)의 내용을 참조하십시오.

값 및 데이터 유형

CLEM 표현식은 값, 필드 이름, 연산자 및 함수로부터 구성된 수식과 유사합니다. 가장 단순한 CLEM 표현식은 값 또는 필드 이름입니다. 유효한 값의 예제는 다음과 같습니다.

```
31.79  
'banana'
```

필드 이름의 예:

```
Product_ID  
'$P-NextField'
```

여기서 *Product*는 장바구니 데이터 세트의 필드 이름이고, '\$P-NextField'는 모수의 이름이고, 표현식의 값은 이름 지정된 필드의 값입니다. 일반적으로 필드 이름은 문자로 시작하고 숫자와 밑줄(_)을 포함할 수도 있습니다. 이름을 따옴표 내에 놓으면 이러한 규칙을 따르지 않는 이름을 사용할 수 있습니다. CLEM 값은 다음 중 하나일 수 있습니다.

- 문자열--예: "c1", "Type 2", "a piece of free text"
- 정수--예: 12, 0, -189
- 실수--예: 12.34, 0.0, -0.0045
- 날짜/시간 필드--예: 05/12/2002, 12/05/2002, 12/05/02

다음 요소를 사용하는 것도 가능합니다.

- 문자 코드--예: 'a' or 3
- 항목 목록--예: [1 2 3], ['Type 1' 'Type 2']

문자 코드와 목록은 보통 필드 값으로서 발생하지 않습니다. 일반적으로 이들은 CLEM 함수의 인수로서 사용합니다.

인용 규칙

CLEM 표현식에서 사용하는 필드, 값, 모수 및 문자열을 결정할 때 소프트웨어는 유연하지만, 다음 일반 규칙은 표현식을 작성할 때 사용할 "우수 사례" 목록을 제공합니다.

- **문자열**—문자열을 쓸 때 항상 큰따옴표를 사용하십시오("Type 2" 또는 "value"). 작은따옴표를 대신 사용할 수 있지만 인용된 필드와 혼동될 위험이 있습니다.
- **문자**—항상 '와 같이 하나의 역따옴표를 사용하십시오. 예를 들어, `stripchar('d', "drugA")` 함수에서 d 문자를 참고하십시오. 이에 대한 유일한 예외는 문자열에서 특정 문자를 가리키기 위해 정수를 사용하는 경우입니다. 예를 들어, `lowertoupper("drugA"(5))` → "A" 함수에서 5 문자를 참고하십시오. 참고: 표준 영국 및 미국 키보드에서 역따옴표 문자의 키(억음 부호, 유니코드 0060)는 Esc 키 바로 아래에서 찾을 수 있습니다.
- **필드**—필드는 일반적으로 CLEM 표현식에서 사용할 때 인용을 사용하지 않습니다(`subscr(2, arrayID)`) → CHAR). 필요한 경우 작은따옴표를 사용하여 공백이나 기타 특수 문자를 묶을 수 있습니다('Order Number'). 인용되었지만 데이터 세트에 정의되지 않은 필드는 문자열로 잘못 읽힙니다.
- **모수**—항상 작은따옴표('\$P-threshold')를 사용합니다.

표현식 및 조건

CLEM 표현식은 다음과 같은 결과를 리턴할 수 있습니다(예: 새 값을 파생시킬 때).

```
Weight * 2.2
Age + 1
sqrt(Signal-Echo)
```

또는 다음과 같은 참 또는 거짓을 평가할 수 있습니다(조건을 기준으로 선택할 때).

```
Drug = "drugA"
Age < 16
not(PowerFlux) and Power > 2000
```

다음과 같이 CLEM 표현식에서 연산자와 함수를 임의로 결합할 수 있습니다.

```
sqrt(abs(Signal)) * max(T1, T2) + Baseline
```

대괄호와 연산자 우선순위는 표현식이 평가되는 순서를 결정합니다. 이 예에서는 평가의 순서는 다음과 같습니다.

- `abs(Signal)`가 평가되고, `sqrt`가 해당 결과에 적용됩니다.
- `max(T1, T2)`가 평가됩니다.

- 두 개의 결과를 곱합니다. **x**가 +보다 우선순위가 높습니다.
- 마지막으로, **Baseline**이 결과에 추가됩니다.

우선순위의 내림차순(즉, 가장 먼저 수행되는 작업부터 가장 마지막에 수행되는 작업)은 다음과 같습니다.

- 함수 인수
- 함수 호출
- **xx**
- **x / mod div rem**
- **+ -**
- **> < >= <= /== == = /=**

우선순위를 대체하려거나 평가 순서에 의심이 든다면 다음과 같이 소괄호를 사용하여 이를 명백하게 만들 수 있습니다.

```
sqrt(abs(Signal)) * (max(T1, T2) + Baseline)
```

스트림, 세션 및 수퍼노드 모수

CLEM 표현식과 스크립팅에서 사용할 모수를 정의할 수 있습니다. 사실상 모수는 현재 스트림, 세션 또는 수퍼노드와 함께 저장되고 지속되는 사용자 정의 변수이며 스크립팅을 통해서는 물론 사용자 인터페이스에서도 액세스 할 수 있습니다. 예를 들어, 스트림을 저장하면 해당 스트림에 대해 설정된 모수도 저장됩니다(이는 선언된 스크립트에서만 사용할 수 있는 로컬 스크립트 변수와 모수를 구별함). 모수는 스크립트에 하드 코드화하지 않아도 되는 필드 및 값에 대한 정보를 제공하여 스크립팅에서 스크립트 동작을 제어하는데 사용하는 경우가 많습니다.

모수 범위는 모수가 설정된 위치에 따라 다릅니다.

- 스트림 모수는 스트림 스크립트 또는 스트림 특성 대화 상자에서 설정할 수 있으며 스트림의 모든 노드에 사용 할 수 있습니다. 표현식 작성기의 모수 목록에 표시됩니다.
- 세션 모수는 독립형 스크립트 또는 세션 모수 대화 상자에서 설정할 수 있습니다. 현재 세션에서 사용하는 모든 스트림(관리자 분할창의 스트림 탭에 나열된 모든 스트림)에 사용할 수 있습니다.

수퍼노드에 대해서도 모수를 설정할 수 있는데, 이 경우 해당 수퍼노드 내에 캡슐화된 노드에만 표시됩니다.

CLEM 표현식에서 모수 사용

모수는 CLEM 표현식에서 **\$P-pname**으로 표시됩니다. 여기서 **pname**은 모수의 이름입니다. CLEM 표현식에서 사용하면 모수는 작은 따옴표 안에 두어야 합니다(예: '**\$P-scale**').

사용 가능한 모수는 표현식 작성기를 사용하여 쉽게 볼 수 있습니다. 현재 모수 보기:

1. CLEM 표현식을 승인하는 대화 상자에서 표현식 작성기 단추를 클릭하십시오.
2. 필드 목록에서 모수를 선택하십시오.

CLEM 표현식에 삽입하기 위해 목록에서 모수를 선택할 수 있습니다. 자세한 정보는 [138 페이지의 『필드, 모수 및 글로벌 변수 선택』](#)의 내용을 참조하십시오.

문자열에 대한 작업

문자열에 사용 가능한 많은 연산이 있는데, 다음과 같습니다.

- 문자열을 대문자 또는 소문자로 변환 - **uppertolower(CHAR)**
- 문자열 변수에서 지정된 문자(예: 'ID_' 또는 '\$') 제거 - **stripchar(CHAR,STRING)**
- 문자열 변수의 길이(문자 수) 판별 - **length(STRING)**
- 문자열 값의 알파벳 순서 검사 - **alphabefore(STRING1, STRING2)**
- 값에서 선행 또는 후행 공백 제거 - **trim(STRING), trim_start(STRING) 또는 trimend(STRING)**

- 문자열에서 처음 또는 마지막 n 개의 문자 추출 - `startstring(LENGTH, STRING)` 또는 `endstring(LENGTH, STRING)`. 예를 들어, 제품 이름을 4자리 ID 코드와 결합(ACME CAMERA-D109)하는 *item* 필드가 있다고 가정하십시오. 4자리 코드만 포함하는 새 필드를 작성하려면 파생 노드에 다음 수식을 지정하십시오.

```
endstring(4, item)
```

- 특정 패턴 매치 - `STRING matches PATTERN`. 예를 들어, 직위 어딘가에 "market"이 포함되어 있는 사용자를 선택하려면 선택 노드에 다음을 지정할 수 있습니다.

```
job_title matches "*market*"
```

- 문자열 내 하위 문자열의 인스턴스 모두 바꾸기 - `replace(SUBSTRING, NEWSUBSTRING, STRING)`. 예를 들어, 텍스트 마이닝 이전에 지원되지 않는 문자(예: 수직 파이프(|))를 세미콜론으로 모두 바꾸려면 채움 노드 `replace` 함수를 사용하십시오. 필드 채우기에서 문자가 나타날 수 있는 필드를 모두 선택하십시오. **바꾸기:** 조건에 대해 항상을 선택하고 **바꿀 문자열:** 아래에 다음 조건을 지정하십시오.

```
replace(' | ', ';' , @FIELD)
```

- 특정 하위 문자열의 존재를 기준으로 플래그 필드 파생. 예를 들어, 파생 노드에서 문자열 함수를 사용하여 다음과 같은 표현식으로 각 반응에 대해 별도의 플래그 필드를 생성할 수 있습니다.

```
hassubstring(museums, "museum_of_design")
```

자세한 정보는 [158 페이지의 『문자열 함수』](#)의 내용을 참조하십시오.

공백 및 결측값 처리

공백 또는 결측값을 대체하는 것은 데이터 마이너를 위한 공통적인 데이터 준비 작업입니다. CLEM은 공백 처리를 자동화하기 위한 많은 도구를 제공합니다. 채움 노드는 공백에 대한 작업을 하는 가장 일반적인 위치입니다. 그러나 CLEM 표현식을 허용하는 모든 노드에서 다음 함수를 사용할 수 있습니다.

- `@BLANK(FIELD)`을 사용하여 특정 필드(예: *Age*)에 대해 값이 공백인 레코드를 판별할 수 있습니다.
- `@NULL(FIELD)`을 사용하여 지정된 필드에 대해 시스템 결측값인 레코드를 판별할 수 있습니다. IBM SPSS Modeler에서 시스템 결측값은 `$null$` 값으로 표시됩니다.

자세한 정보는 [173 페이지의 『공백 및 널값 처리 함수』](#)의 내용을 참조하십시오.

숫자에 대한 작업

IBM SPSS Modeler에서는 숫자 값에 대한 다수의 표준 연산을 사용할 수 있는데, 다음과 같습니다.

- 지정된 각의 사인 계산 — `sin(NUM)`
- 숫자 필드의 자연로그 계산 — `log(NUM)`
- 두 숫자의 합계 계산 — `NUM1 + NUM2`

자세한 정보는 [153 페이지의 『숫자 함수』](#)의 내용을 참조하십시오.

시간 및 날짜에 대한 작업

시간 및 날짜 형식은 데이터 소스와 로케일에 따라 다릅니다. 날짜 및 시간 형식은 각 스트림에 고유하며 스트림 특성 대화 상자에서 설정됩니다. 다음 예제는 날짜/시간 필드에 대한 작업에 일반적으로 사용하는 함수입니다.

경과 시간 계산

다음 함수와 유사한 함수 패밀리를 사용하여 기준 날짜부터 경과한 시간을 쉽게 계산할 수 있습니다. 이 함수는 기준 날짜부터 날짜 문자열 DATE가 나타내는 날짜까지의 시간(월)을 실수로 리턴합니다. 이는 30.0일을 1개월로 한 대략적인 숫자입니다.

```
date_in_months(Date)
```

날짜/시간 값 비교

다음 함수와 유사한 함수를 사용하여 레코드 전반에서 날짜/시간 필드 값과 비교할 수 있습니다. 이 함수는 날짜 문자열 DATE1이 날짜 문자열 DATE2가 나타내는 날짜 이전의 날짜를 나타내는 경우 참 값을 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다.

```
date_before(Date1, Date2)
```

차이 계산

다음과 같은 함수를 사용하여 두 개의 시간과 두 개의 날짜 간의 차이도 계산할 수 있습니다.

```
date_weeks_difference(Date1, Date2)
```

이 함수는 날짜 문자열 DATE1이 나타내는 날짜부터 날짜 문자열 DATE2가 나타내는 날짜까지의 시간(주)을 실수로 리턴합니다. 이는 7.0일을 1주로 합니다. DATE2가 DATE1 이전인 경우 이 함수는 음수를 리턴합니다.

오늘 날짜

@TODAY 함수를 사용하여 데이터 세트에 현재 날짜를 추가할 수 있습니다. 오늘 날짜는 스트림 특성 대화 상자에서 선택된 날짜 형식을 사용하여 지정된 필드 또는 새 필드에 문자열로 추가됩니다. 자세한 정보는 [163 페이지의 『날짜 및 시간 함수』](#)의 내용을 참조하십시오.

다중 필드 요약

CLEM 언어는 다중 필드 전반에서 요약 통계를 리턴하는 많은 함수를 포함합니다. 이러한 함수는 특히 하나의 질문에 대한 다중반응이 다중 필드에 저장될 수 있는 설문조사 데이터 분석 시 유용합니다. 자세한 정보는 [133 페이지의 『다중반응 데이터에 대한 작업』](#)의 내용을 참조하십시오.

비교 함수

min_n과 max_n 함수를 사용하여 다중 필드 전반에서 값을 비교할 수 있습니다. 예를 들어, 다음과 같습니다.

```
max_n(['card1fee' 'card2fee' 'card3fee' 'card4fee'])
```

또한 다수의 계산 함수를 사용하여 특정 기준을 충족시키는 값 수를 확보할 수 있습니다(해당 값이 다중 필드에 저장된 경우에도 적용됨). 예를 들어, 5년 넘게 보유한 카드 수를 계산하려면 다음을 사용하십시오.

```
count_greater_than(5, ['cardtenure' 'card2tenure' 'card3tenure'])
```

같은 필드 세트에서 널 값을 계산하려면 다음을 사용하십시오.

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

이 예제에서는 카드를 보유한 사람 수는 계산하지 않고 보유 중인 카드 수를 계산합니다. 자세한 정보는 [151 페이지의 『비교 함수』](#)의 내용을 참조하십시오.

지정된 값이 다중 필드에서 나타나는 횟수를 계산하려면 count_equal 함수를 사용할 수 있습니다. 다음 예제에서는 목록에서 Y 값이 포함된 필드 수를 계산합니다.

```
count_equal("Y", [Answer1, Answer2, Answer3])
```

목록의 필드에 다음 값이 주어지면 이 함수는 표시된 대로 Y 값에 대한 결과를 리턴합니다.

표 8. 함수 값

Answer1	Answer2	Answer3	수
Y	N	Y	2
Y	N	N	1

숫자 함수

sum_n, mean_n, sdev_n 함수를 사용하여 다중 필드 전반에서 통계를 확보할 수 있습니다.

```
sum_n(['card1bal' 'card2bal' 'card3bal'])
```

```
mean_n(['card1bal' 'card2bal' 'card3bal'])
```

자세한 정보는 [153 페이지의 『숫자 함수』](#)의 내용을 참조하십시오.

필드 목록 생성

필드 목록을 입력으로 허용하는 함수를 사용하는 경우 특수 함수 @FIELDS_BETWEEN(start, end) 및 @FIELDS_MATCHING(pattern)을 입력으로서 사용할 수 있습니다. 예를 들어, 필드 순서가 앞선 sum_n 예제에 표시된 것과 같다고 가정하면 다음과 같습니다.

```
sum_n(@FIELDS_BETWEEN(card1bal, card3bal))
```

또는 "card"로 시작하는 모든 필드에서 널 값 수를 계산하려면 다음을 사용하십시오.

```
count_nulls(@FIELDS_MATCHING('card*'))
```

자세한 정보는 [173 페이지의 『특수 필드』](#)의 내용을 참조하십시오.

다중반응 데이터에 대한 작업

다음과 같은 다수의 비교 함수를 사용하여 다중반응 데이터를 분석할 수 있습니다.

- value_at
- first_index / last_index
- first_non_null / last_non_null
- first_non_null_index / last_non_null_index
- min_index / max_index

예를 들어, 특정 구매 결정을 위해 첫 번째, 두 번째, 세 번째로 가장 중요한 이유(예: 가격, 개인적인 추천, 검토, 로컬 공급자 등)를 묻는 다중반응 질문을 가정하십시오. 이 경우, 처음 포함된 필드의 지수를 파생시켜 가격의 중요도를 결정합니다.

```
first_index("price", [Reason1 Reason2 Reason3])
```

마찬가지로, 고객에게 구매 우도가 있는 순서대로 자동차 3대를 순위화하도록 요청하고 다음과 같이 반응을 세 개의 개별 필드에 코드화했다고 가정하십시오.

표 9. 자동차 순위화 예제			
고객 id	car1	car2	car3
101	1	3	2
102	3	2	1
103	2	3	1

이 경우, `min_index` 함수를 사용하여 가장 좋아하는 자동차(#1로 순위가 매겨짐, 그렇지 않으면 최저 순위)에 대한 필드의 지수를 판별할 수 있습니다.

```
min_index(['car1' 'car2' 'car3'])
```

자세한 정보는 [151 페이지의 『비교 함수』](#)의 내용을 참조하십시오.

다중반응 세트 참조

특수 `@MULTI_RESPONSE_SET` 함수를 사용하여 다중반응 세트의 모든 필드를 참조할 수 있습니다. 예를 들어, 이전 예제에 있는 세 개의 `car` 필드가 `car_rankings`라는 다중반응 세트에 포함된 경우 다음은 동일한 결과를 리턴합니다.

```
max_index(@MULTI_RESPONSE_SET("car_rankings"))
```

표현식 작성기

CLEM 표현식을 수동으로 입력하거나 CLEM 함수와 연산자의 전체 목록뿐만 아니라 현재 스트림의 데이터 필드를 표시하는 표현식 작성기를 사용할 수 있습니다. 이를 통해 필드나 함수의 정확한 이름을 기억하지 않고도 표현식을 빠르게 작성할 수 있습니다. 또한 작성기는 필드와 값에 적합한 인용구를 자동으로 추가하여 통사적으로 정확한 표현식 작성을 보다 쉽게 해줍니다.

참고: 표현식 작성기는 스크립트나 모수 설정에서는 지원되지 않습니다.

참고: 데이터 소스를 변경하려는 경우에는 소스를 변경하기 전에 표현식 작성기가 사용자가 선택한 함수를 여전히 지원하는지를 확인해야 합니다. 모든 데이터베이스가 모든 함수를 지원하지는 않으므로 새 데이터 소스를 실행하는 경우 오류가 발생할 수도 있습니다.

표현식 작성기에 액세스

선택, 균형, 파생, 채움, 분석, 보고서 및 테이블 노드를 포함하여 CLEM 표현식이 사용하는 모든 노드에서 표현식 작성기를 사용할 수 있습니다. 수식 필드 오른쪽에 있는 계산기 단추를 클릭하여 열 수 있습니다.

표현식 작성

표현식 작성기는 필드, 함수 및 연산자 전체 목록은 물론 데이터가 인스턴스화된 경우 데이터 값에 대한 액세스도 제공합니다.

표현식 작성기를 사용하여 표현식을 작성하는 방법

1. 함수 및 필드 목록을 참조로 사용하여 표현식 필드를 입력하십시오.
2. 스크롤 목록에서 필수 필드 및 함수를 선택하십시오.
3. 노란색 화살표 단추를 두 번 클릭하거나 클릭하여 표현식 필드에 필드나 함수를 추가하십시오.
4. 대화 상자 중앙의 피연산자 단추를 사용하여 표현식에 연산을 삽입하십시오.

함수 선택

함수 목록은 사용 가능한 모든 CLEM 함수 및 연산자를 표시합니다. 스크롤하여 목록에서 함수를 선택하거나 보다 쉽게 검색하려면 드롭 다운 목록을 사용하여 함수 또는 연산자의 서브세트를 표시하십시오. 사용 가능한 함수는 쉽게 검색할 수 있도록 범주로 그룹화되어 있습니다.

이러한 범주의 대부분은 CLEM 언어 설명의 참조 섹션에 설명되어 있습니다. 추가 정보는 [148 페이지의 『함수 참조』](#)의 내용을 참조하십시오.

기타 범주는 다음과 같습니다.

- 일반 함수에는 가장 자주 사용하는 함수 중 일부의 선택사항이 포함됩니다.
- 최근에 사용에는 현재 세션 내에서 사용하는 CLEM 함수 목록이 포함됩니다.

- @ 함수에는 모든 특수 함수의 목록이 포함되며 각 이름 앞에 "@" 부호가 있습니다.

참고: @DIFF1(FIELD1, FIELD2) 및 @DIFF2(FIELD1, FIELD2) 함수의 경우 두 필드 유형이 동일해야 합니다(예: 둘 다 정수 또는 둘 다 Long 또는 Real).

- **데이터베이스 함수.** 스트림에 데이터베이스 연결이 포함되는 경우(데이터베이스 소스 노드 사용), 이 선택영역은 사용자 정의 함수(UDF)를 포함하여 해당 데이터베이스 내에서 사용 가능한 함수를 나열합니다. 추가 정보는 135 페이지의 『데이터베이스 함수』의 내용을 참조하십시오.
- **데이터베이스 통합.** 스트림에 데이터베이스 연결이 포함되는 경우(데이터베이스 소스 노드 사용), 이 선택영역은 해당 데이터베이스 내에서 사용 가능한 통합 옵션을 나열합니다. 이러한 옵션은 통합 노드의 표현식 작성기에서 사용 가능합니다.
- **데이터베이스 창 통합.** 스트림에 데이터베이스 연결이 포함되는 경우(데이터베이스 소스 노드 사용), 이 선택영역은 해당 데이터베이스 내에서 사용할 수 있는 창 통합 옵션을 나열합니다. 이러한 옵션은 필드 작업 팔레트에서만 노드 내의 표현식 작성기에서 사용 가능합니다.

참고: SPSS Modeler가 데이터베이스 시스템 보기에서 창 통합 함수를 획득하므로 사용 가능한 옵션은 데이터베이스 작동에 따라 다릅니다.

"통합"이라고 불리기는 하지만 이러한 옵션은 통합 노드 전용으로 지정되지 않았으므로 파생 또는 선택 등과 같은 노드에 보다 적합합니다. 이는 출력이 참 통합 대신에 스칼라이기 때문입니다. 즉, 통합 노드가 하는 것과 같은 방법으로 출력에 표시되는 데이터량을 줄이지 않습니다. 예를 들어, "현재 행과 모든 이전 행의 평균" 등과 같이 데이터 행을 아래쪽으로 이동 평균을 제공하기 위해 이 통합 정렬을 사용할 수 있습니다.

- **내장 통합.** 사용할 수 있는 가능한 통합 모드 목록을 포함합니다.
- **연산자는** 표현식을 작성할 때 사용할 수 있는 모든 연산자를 나열합니다. 연산자는 대화 상자의 중심에서 단추에서도 사용 가능합니다.
- **모든 함수에는** 사용 가능한 CLEM 함수의 전체 목록을 포함합니다.

함수 그룹을 선택한 후에는 두 번 클릭하여 함수를 커서 위치로 표시되는 지점의 표현식 필드에 삽입하십시오.

데이터베이스 함수

데이터베이스 함수는 서로 다른 여러 위치에 나열할 수 있습니다. 다음 테이블은 SPSS Modeler가 함수 세부사항을 찾을 때 검색하는 위치를 보여줍니다. 이 테이블은 데이터베이스 관리자가 필수 영역에 대한 액세스 권한이 있는 사용자가 서로 다른 함수를 사용할 수 있도록 확인하는 데 사용할 수 있습니다.

또한 테이블은 함수가 사용 가능할 때 데이터베이스와 함수 유형을 기반으로 필터링하는 데 사용하는 조건을 나열합니다.

참고: Amazon Redshift에서 데이터베이스 함수를 사용하는 경우 데이터베이스 관리자가 다음 6가지 데이터베이스 오브젝트에 대한 권한을 부여해야 합니다. 첫 번째 4개는 시스템 카탈로그 테이블이며 나머지 2개는 스키마입니다.

- pg_type
- pg_proc
- pg_namespace
- pg_aggregate
- information_schema
- pg_catalog

표 10. 표현식 작성기에서 데이터베이스 함수

데이터베이스	함수 유형	함수 찾는 위치	함수를 필터링하는 데 사용하는 조건
Db2 LUW	UDF	SYSCAT.ROUTINES SYSCAT.ROUTINEP RMS	ROUTINETYPE은 F이고 FUNCTIONTYPE은 S임

표 10. 표현식 작성기에서 데이터베이스 함수 (계속)

데이터베이스	함수 유형	함수 찾는 위치	함수를 필터링하는 데 사용하는 조건
Db2 LUW	UDA	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE은 F이고 FUNCTIONTYPE은 C임
Db2 iSeries	UDF	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE은 F이고 FUNCTION_TYPE은 S임
Db2 iSeries	UDA	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE은 F이고 FUNCTION_TYPE은 C임
Db2 z/OS	UDF	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE은 F이고 FUNCTIONTYPE은 S임
Db2 z/OS	UDA	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE은 F이고 FUNCTIONTYPE은 C임
SQL Server	UDF	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE은 FN 또는 FS임
SQL Server	UDA	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE은 AF임
Oracle	UDF	ALL_ARGUMENTS ALL_PROCEDURES	다음 모든 조건을 충족합니다. <ul style="list-style-type: none"> OBJECT_TYPE은 FUNCTION임 AGGREGATE는 NO임 PLS_TYPE은 NULL이 아님
Oracle	UDA	ALL_ARGUMENTS ALL_PROCEDURES	다음 모든 조건을 충족합니다. <ul style="list-style-type: none"> ARGUMENT_NAME은 NULL임 AGGREGATE는 YES임 PLS_TYPE은 NULL이 아님
Teradata	UDF	DBC.FUNCTIONS DBC.ALLRIGHTS	다음 모든 조건을 충족합니다. <ul style="list-style-type: none"> FUNCTIONTYPE은 F임 COLUMNNAME은 RETURN0임 SPPARAMETERTYPE은 0임 ACCESSRIGHT는 EF임
Teradata	UDA	DBC.FUNCTIONS DBC.ALLRIGHTS	다음 모든 조건을 충족합니다. <ul style="list-style-type: none"> FUNCTIONTYPE은 A임 COLUMNNAME은 RETURN0임 SPPARAMETERTYPE은 0임 ACCESSRIGHT는 EF임

표 10. 표현식 작성기에서 데이터베이스 함수 (계속)

데이터베이스	함수 유형	함수 찾는 위치	함수를 필터링하는 데 사용하는 조건
Netezza	UDF	####.._V_FUNCTION NZA.._V_FUNCTION INZA.._V_FUNCTION	####.._V_FUNCTION의 경우 다음 조건이 적용됩니다. <ul style="list-style-type: none"> RESULT에는 다음과 같은 값이 있는 문자열을 포함하지 않습니다. TABLE% FUNCTION에는 다음과 같은 값이 있는 문자열을 포함하지 않습니다. '/_%' escape '/' VARARGS는 FALSE임 NZA.._V_FUNCTION 및 INZA.._V_FUNCTION 둘 모두의 경우 다음 조건이 적용됩니다. <ul style="list-style-type: none"> RESULT에는 다음과 같은 값이 있는 문자열을 포함하지 않습니다. TABLE% FUNCTION에는 다음과 같은 값이 있는 문자열을 포함하지 않습니다. '/_%' escape '/' BUILTIN은 f임 VARARGS는 FALSE임
Netezza	UDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	다음 조건을 둘 모두 충족합니다. <ul style="list-style-type: none"> AGGTYPE은 ANY이거나 GROUPED임 VARARGS는 FALSE임
Netezza	WUDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	####.._V_AGGREGATE의 경우 다음 조건이 적용됩니다. <ul style="list-style-type: none"> AGGTYPE은 ANY이거나 ANALYTIC임 AGGREGATE은 MAX_LABEL이 아님 VARARGS는 FALSE임 NZA.._V_FUNCTION 및 INZA.._V_FUNCTION 둘 모두의 경우 다음 조건이 적용됩니다. <ul style="list-style-type: none"> AGGTYPE은 ANY이거나 ANALYTIC임 BUILTIN은 f임 VARARGS는 FALSE임

테이블에 사용하는 주요 용어

- UDF 사용자 정의 함수(User Defined Function)
- UDA 사용자 정의 집계 함수(User Defined Aggregate Function)
- WUDA 사용자 정의 창 집계 함수(User Defined Window Aggregate Function)
- #### 현재 연결된 데이터베이스

필드, 모수 및 글로벌 변수 선택

필드 목록은 데이터 스트림의 이 위치에서 사용 가능한 모든 필드를 표시합니다. 목록에서 필드를 선택하려면 스크롤하십시오. 표현식에 필드를 추가하려면 노란색 화살표 단추를 두 번 클릭하거나 클릭하십시오.

자세한 정보는 [130 페이지의 『스트림, 세션 및 수퍼노드 모수』](#)의 내용을 참조하십시오.

필드에 추가로 다음 항목에서 선택할 수도 있습니다.

다중반응 세트. 자세한 정보는 *IBM SPSS Modeler* 소스, 프로세스 및 출력 노드 안내서를 참조하십시오.

최근에 사용에는 필드 목록, 다중반응 세트, 모수 및 현재 세션 내에서 사용된 글로벌 값이 포함됩니다.

모수. 자세한 정보는 [130 페이지의 『스트림, 세션 및 수퍼노드 모수』](#)의 내용을 참조하십시오.

글로벌 값. 자세한 정보는 *IBM SPSS Modeler* 소스, 프로세스 및 출력 노드 안내서를 참조하십시오.

값 보기 또는 선택

필드 값은 표현식 작성기, 데이터 검토 보고서 및 시간 간격 노드에서 미래 값을 편집할 때를 포함하여 시스템 내의 여러 장소에서 볼 수 있습니다. 데이터가 이 기능을 사용하기 위해서는 소스 또는 유형 노드에서 완전히 인스턴스화되어야 저장 공간, 유형 및 값이 알려질 수 있음을 유의하십시오.

표현식 작성기 또는 시간 간격 노드에서 필드의 값을 보려면 필수 필드를 선택하고 값 선택도구 단추를 클릭하여 선택된 필드의 값을 나열하는 대화 상자를 여십시오. 그런 다음 값을 선택하고 **삽입**을 클릭하여 값을 현재 표현식 또는 목록에 붙여넣을 수 있습니다.



그림 16. 값 선택도구 단추

플래그 및 명목 필드의 경우 정의된 모든 값이 나열됩니다. 연속(숫자 범위) 필드의 경우 최소 및 최대 값이 표시됩니다.

CLEM 표현식 검사

표현식 작성기(오른쪽 하단 모서리)에서 **검사**를 클릭하여 표현식을 검증하십시오. 검사하지 않은 표현식은 빨간색으로 표시됩니다. 오류가 발견되면 원인을 나타내는 메시지가 표시됩니다.

다음 항목을 검사합니다.

- 값 및 필드 이름의 올바른 인용
- 모수 및 글로벌 변수의 올바른 사용
- 유효한 연산자 사용
- 참조된 필드의 존재
- 참조된 전역값의 존재 및 정의

구문에서 오류가 발생하면 표현식을 수동으로 입력하지 말고 목록과 연산자 단추를 사용하여 표현식을 작성하십시오. 필드 및 값에 대한 적절한 인용을 자동으로 추가합니다.

IBM Analytical Decision Management에서 표현식을 작성할 때 다음 제한사항에 유의하십시오. 표현식은 다음 항목을 포함할 수 없습니다.

- IBM SPSS Modeler 스트림 모수에 대한 참조
- IBM SPSS Modeler 스트림 글로벌에 대한 참조
- 데이터베이스 함수에 대한 참조
- 다음 특수 필드 또는 필드 값 @ 함수 중 하나에 대한 참조
 - @TARGET
 - @PREDICTED
 - @FIELD

- @PARTITION_FIELD
- @TRAINING_PARTITION
- @TESTING_PARTITION
- @VALIDATION_PARTITION

참고: 구분 문자가 포함된 필드 이름은 작은따옴표로 묶어야 합니다. 따옴표를 자동으로 추가하려면 표현식을 수동으로 입력하는 대신 목록 및 연산자 단추를 사용하여 표현식을 작성하면 됩니다. 필드 이름에 다음과 같은 문자가 있을 경우 오류가 발생할 수 있습니다. • ! "# \$% & ' () = ~ | - ^ ¥ @ " + * " < > ? . , / : ; → (화살표), □ △(도형 부호 등)

찾기 및 바꾸기

찾기/바꾸기 대화 상자는 스크립트 또는 표현식 텍스트를 편집하는 위치(스크립트 편집기, CLEM 표현식 작성기 포함)에서, 또는 보고서 노드에서 템플릿을 정의할 때 사용할 수 있습니다. 이러한 영역에서 편집하는 경우 Ctrl+F를 눌러 이 대화 상자에 액세스하고 커서가 텍스트 영역에서 초점을 갖고 있는지 확인하십시오. 예를 들어, 채움 노드에서 작업하는 경우 설정 템의 텍스트 영역에서 또는 표현식 작성기의 텍스트 필드에서 대화 상자에 액세스할 수 있습니다.

1. 텍스트 영역에 커서를 두고 Ctrl+F를 눌러서 찾기/바꾸기 대화 상자에 액세스하십시오.
2. 검색할 텍스트를 입력하거나 최근에 검색한 항목 드롭 다운 목록에서 선택하십시오.
3. 대체 텍스트(있는 경우)를 입력하십시오.
4. 다음 찾기를 클릭하여 검색을 시작하십시오.
5. 바꾸기를 클릭하여 현재 선택을 바꾸거나 모두 바꾸기를 클릭하여 선택된 또는 모든 인스턴스를 업데이트하십시오.
6. 각 작업 후에 대화 상자가 닫힙니다. 텍스트 영역에서 F3을 눌러 마지막 찾기 작업을 반복하거나 Ctrl+F를 눌러 대화 상자에 다시 액세스하십시오.

검색 옵션

대소문자 구분. 찾기 작업이 대소문자를 구분하는지 여부를 지정합니다(예: *myvar*이 *myVar*과 매치하는지 여부). 대체 텍스트는 이 설정에 관계없이 항상 입력된 대로 정확하게 삽입됩니다.

전체 단어만. 찾기 작업이 단어 안에 임베드된 텍스트와 매치하는지 여부를 지정합니다. 선택된 경우, 예를 들어 *spider*에 대한 검색은 *spiderman* 또는 *spider-man*과 매치하지 않습니다.

정규식. 정규식 구문 사용 여부를 지정합니다(다음 절 참조). 선택된 경우에는 전체 단어만 옵션을 사용할 수 없으며 해당 값을 무시합니다.

선택된 텍스트만. 모두 바꾸기 옵션 사용 시 검색 범위를 제어합니다.

정규식 구문

정규식을 사용하여 템이나 줄 바꾸기 문자와 같은 특수 문자, *a - d*와 같은 문자 클래스나 범위, 숫자나 비숫자, 행 시작이나 끝과 같은 경계에 대한 검색을 수행할 수 있습니다. 지원되는 표현식 유형은 다음과 같습니다.

표 11. 문자 매치	
문자	매치
x	x 문자
\\"	백슬래시 문자
\0n	8진 값 0n이 포함된 문자(0 <= n <= 7)
\0nn	8진 값 0nn이 포함된 문자(0 <= n <= 7)
\0mnn	8진 값 0mnn이 포함된 문자(0 <= m <= 3, 0 <= n <= 7)
\xhh	16진값 0xhh가 포함된 문자
\uhhhh	16진값 0uhhhh가 포함된 문자

표 11. 문자 매치 (계속)

문자	매치
\t	탭 문자('\'u0009')
\n	줄 바꾸기(줄 바꾸기) 문자('\'u000A')
\r	캐리지 리턴 문자('\'u000D')
\f	용지 넘김 문자('\'u000C')
\a	경고(벨) 문자('\'u0007')
\e	이스케이프 문자('\'u001B')
\cx	x에 해당되는 제어 문자

표 12. 문자 클래스 매치

문자 클래스	매치
[abc]	a, b 또는 c(단순 클래스)
[^abc]	a, b 또는 c를 제외한 모든 문자(뺄셈)
[a-zA-Z]	a - z 또는 A - Z(경계값 포함)
[a-d[m-p]]	a - d 또는 m - p(합집합). 또는 이를 [a-dm-p]로 지정할 수 있음
[a-z&&[def]]	a - z, d, e 또는 f(교집합)
[a-z&&[^bc]]	b와 c를 제외하고 a - z(뺄셈). 또는 이를 [ad-z]로 지정할 수 있음
[a-z&&[^m-p]]	m - p를 제외하고 a - z(뺄셈). 또는 이를 [a-lq-z]로 지정할 수 있음

표 13. 사전 정의된 문자 클래스

사전 정의된 문자 클래스	매치
.	모든 문자(행 종결자와 매치하거나 매치하지 않을 수 있음)
\d	모든 숫자: [0-9]
\D	비숫자: [^0-9]
\s	공백 문자: [\t\n\x0B\f\r]
\S	비공백 문자: [^ \s]
\w	단어 문자: [a-zA-Z_0-9]
\W	비단어 문자: [^ \w]

표 14. 경계 매치

경계 매치기	매치
^	행의 시작
\$	행의 끝
\b	단어 경계
\B	비단어 경계
\A	입력의 시작
\Z	입력의 끝이지만 최종 종결자로도 사용(해당하는 경우)

표 14. 경계 매치 (계속)

경계 매치기	매치
\z	입력의 끝

제 10 장 CLEM 언어 참조

CLEM 참조 개요

이 절에서는 IBM SPSS Modeler 스트림에서 사용하는 데이터를 분석하고 조작하는 데 사용하는 강력한 도구인 Control Language for Expression Manipulation(CLEM)에 대해 설명합니다. 노드 내에서 CLEM을 사용하여 조건 평가 또는 값 파생에서부터 보고서에 데이터 삽입에 이르기까지 다양한 작업을 수행할 수 있습니다.

CLEM 표현식은 값, 필드 이름, 연산자, 함수로 구성됩니다. 올바른 구문을 사용하면 강력한 데이터 연산을 다양하게 작성할 수 있습니다.

CLEM 데이터 유형

CLEM 데이터 유형은 다음으로 구성될 수 있습니다.

- 정수
- 실수
- 문자
- 문자열
- 목록
- 필드
- 날짜/시간

인용 규칙

CLEM 표현식에서 사용하는 필드, 값, 모수, 문자열을 판별할 때 IBM SPSS Modeler가 유연하긴 하지만 다음 일반 규칙은 표현식 작성 시 사용할 "우수 사례" 목록을 제공합니다.

- 문자열 - 문자열 작성 시 항상 큰따옴표를 사용하십시오(예: "Type 2"). 작은따옴표를 대신 사용할 수 있지만 인용된 필드와 혼동될 위험이 있습니다.
- 필드 - 공백 또는 기타 특수 문자를 묶는 데 필요한 경우에만 작은따옴표를 사용하십시오(예: 'Order Number'). 인용되었지만 데이터 세트에 정의되지 않은 필드는 문자열로 잘못 읽힙니다.
- 모수 - 모수 사용 시 항상 작은따옴표를 사용하십시오(예: '\$P-threshold').
- 문자 - 항상 단일 역따옴표(')를 사용하십시오(예: stripchar('d', "drugA")).

이러한 규칙에 대해서는 다음 주제에서 자세히 설명합니다.

정수

정수는 십진 숫자의 시퀀스로 표시됩니다. 또는 정수 앞에 음의 부호(−)를 두면 음수를 표시할 수 있습니다(예: 1234, 999, −77).

CLEM 언어는 임의의 정밀도의 정수를 처리합니다. 최대 정수 크기는 플랫폼에 따라 다릅니다. 값이 정수 필드에 표시하기에 너무 크면 일반적으로 필드 유형을 Real로 변경하면 값이 복원됩니다.

실수

실수는 부동 소수점을 가리킵니다. 실수는 하나 이상의 숫자 뒤에 소수점과 하나 이상의 숫자로 표시됩니다. CLEM 실수는 이중 정밀도를 가집니다.

또는 실수 앞에 음의 부호(−)를 두고 음수를 표시할 수도 있습니다(예: 1.234, 0.999, −77.001). $<number>$ e $<exponent>$ 양식을 사용하여 실수를 지수 표기법으로 표시하십시오(예: 1234.0e5, 1.7e−2). IBM SPSS Modeler 애플리케이션이 파일에서 숫자 문자열을 읽고 이를 숫자로 자동으로 변환하면 소수점 앞에 선행하는 숫자

자가 없는 숫자나 소수점 뒤에 숫자가 없는 숫자도 허용됩니다(예: 999. 또는 .11). 그러나 이러한 양식은 CLEM 표현식에서는 올바르지 않습니다.

참고: CLEM 표현식에서 실수를 참조할 때는 현재 스트림이나 로케일의 설정과 관계없이 마침표를 소수점 구분자로 사용해야 합니다. 예를 들어, 다음을 지정하십시오.

```
Na > 0.6
```

다음이 아닌 것을 유의하십시오.

```
Na > 0,6
```

이는 스트림 특성 대화 상자에서 쉼표가 소수점 기호로 선택되었더라도 적용되며 코드 구문이 특정 로케일이나 규칙으로부터 독립적이어야 하는 일반적인 지침과도 매치합니다.

문자

문자(대개 CHAR로 표시됨)는 일반적으로 CLEM 표현식 내에서 문자열에 대한 검정을 수행하는 데 사용합니다. 예를 들어, `isuppercode` 함수를 사용하여 문자열의 첫 번째 문자가 대문자인지 여부를 판별할 수 있습니다. 다음 CLEM 표현식은 문자를 사용하여 문자열의 첫 번째 문자에 대해 검정을 수행해야 함을 표시합니다.

```
isuppercode(subscrs(1, "MyString"))
```

CLEM 표현식에서 특정 문자의 코드(위치와 대조)를 표시하려면, '`<character>`' 양식의 단일 역따옴표(예: 'A', 'Z')를 사용하십시오.

참고: 필드의 경우 CHAR 저장 유형이 없으므로 필드가 CHAR을 초래하는 표현식으로 채워지거나 파생된 경우 해당 결과는 문자열로 변환됩니다.

문자열

일반적으로 문자열을 큰따옴표로 묶어야 합니다. 문자열의 예는 "c35product2" 및 "referrerID"입니다. 문자열에서 특수 문자를 표시하려면 백슬래시를 사용하십시오(예: "\\$65443"). (백슬래시 문자를 표시하려면 이중 백슬래시, \"를 사용하십시오.) 문자열을 작은따옴표로 묶을 수도 있지만 결과는 인용된 필드 ('referrerID')와 구별이 되지 않습니다. 자세한 정보는 [158 페이지의 『문자열 함수』](#) 주제를 참조하십시오.

목록

목록은 요소의 순서화된 시퀀스이며 이는 혼합 유형일 수도 있습니다. 목록은 꺽쇠 대괄호([])로 싸여 있습니다. 목록의 예제는 [1 2 4 16] 및 ["abc" "def"]입니다. 목록은 IBM SPSS Modeler 필드의 값으로 사용하지 않습니다. 이들은 `member` 및 `oneof` 등과 같이 함수에 인수를 제공하는 데 사용합니다.

참고: 목록은 함수 호출이 아니라 정적 오브젝트(예: 문자열, 번호 또는 필드 이름)로만 구성될 수 있습니다.

필드

함수의 이름이 아닌 CLEM 표현식에서의 이름은 필드 이름으로 가정됩니다. 이들을 단순하게 `Power`, `val27`, `state_flag` 등으로 작성할 수 있지만, 이름이 숫자로 시작하거나 공백 등과 같이 영문자가 아닌 문자를 포함하면(밑줄은 예외) 이름을 작은따옴표 안에 두십시오(예: 'Power Increase', '2nd answer', '#101', '\$P-NextField').

참고: 인용되지만 데이터 세트에 정의되어 있지 않은 필드는 문자열로 잘못 읽힙니다.

날짜

날짜 계산은 스트림 특성 대화 상자에 지정된 "기준" 날짜를 기반으로 합니다. 기본 기준 날짜는 1900년 1월 1일입니다.

CLEM 언어가 지원하는 날짜 형식은 다음과 같습니다.

표 15. CLEM 언어 날짜 형식

형식	예
DDMMYY	150163
MMDDYY	011563
YYMMDD	630115
YYYYMMDD	19630115
YYYYDDDD	연도의 일을 나타내는 3자리 숫자가 뒤에 오는 4자리 연도(예: 2000032는 2000년의 32번째 날 또는 2000년 2월 1일을 나타냄)
DAY	현재 로케일로 된 요일(예: 영어로 Monday, Tuesday, ...)
MONTH	현재 로케일로 된 월(예: January, February, ...)
DD/MM/YY	15/01/63
DD/MM/YYYY	15/01/1963
MM/DD/YY	01/15/63
MM/DD/YYYY	01/15/1963
DD-MM-YY	15-01-63
DD-MM-YYYY	15-01-1963
MM-DD-YY	01-15-63
MM-DD-YYYY	01-15-1963
DD.MM.YY	15.01.63
DD.MM.YYYY	15.01.1963
MM.DD.YY	01.15.63
MM.DD.YYYY	01.15.1963
DD-MON-YY	15-JAN-63, 15-jan-63, 15-Jan-63
DD/MON/YY	15/JAN/63, 15/jan/63, 15/Jan/63
DD.MON.YY	15.JAN.63, 15.jan.63, 15.Jan.63
DD-MON-YYYY	15-JAN-1963, 15-jan-1963, 15-Jan-1963
DD/MON/YYYY	15/JAN/1963, 15/jan/1963, 15/Jan/1963
DD.MON.YYYY	15.JAN.1963, 15.jan.1963, 15.Jan.1963
MON YYYY	Jan 2004
q Q YYYY	문자 Q와 4자리 연도가 뒤에 오는 분기를 나타내는 숫자(1-4)로 표시된 날짜(예: 2004년 12월 25일은 4 Q 2004로 표시됨)
ww WK YYYY	문자 WK와 4자리 연도가 뒤에 오는 연간 통산주를 나타내는 2자리 숫자. 연간 통산주는 첫 번째 요일이 월요일이고 첫 번째 주에 최소 1일이 있다고 가정하고 계산됩니다.

시간

CLEM 언어는 다음 시간 형식을 지원합니다.

표 16. CLEM 언어 시간 형식	
형식	예
HHMMSS	120112, 010101, 221212
HHMM	1223, 0745, 2207
MMSS	5558, 0100
HH:MM:SS	12:01:12, 01:01:01, 22:12:12
HH:MM	12:23, 07:45, 22:07
MM:SS	55:58, 01:00
(H)H:(M)M:(S)S	12:1:12, 1:1:1, 22:12:12
(H)H:(M)M	12:23, 7:45, 22:7
(M)M:(S)S	55:58, 1:0
HH.MM.SS	12.01.12, 01.01.01, 22.12.12
HH.MM	12.23, 07.45, 22.07
MM.SS	55.58, 01.00
(H)H.(M)M.(S)S	12.1.12, 1.1.1, 22.12.12
(H)H.(M)M	12.23, 7.45, 22.7
(M)M.(S)S	55.58, 1.0

CLEM 연산자

다음 연산자가 사용 가능합니다.

표 17. CLEM 언어 연산자		
작업	주석	우선순위(다음 섹션 참조)
or	두 개의 CLEM 표현식 사이에 사용합니다. 둘 중 하나라도 참이거나 둘 모두가 참이면 참 값을 리턴합니다.	10
and	두 개의 CLEM 표현식 사이에 사용합니다. 둘 모두 참이면 참 값을 리턴합니다.	9
=	두 개의 비교 가능한 항목 사이에 사용합니다. ITEM1이 ITEM2와 동일하면 참을 리턴합니다.	7
==	=과 동일합니다.	7
/=	두 개의 비교 가능한 항목 사이에 사용합니다. ITEM1이 ITEM2와 동일하지 않으면 참을 리턴합니다.	7
/==	/=과 동일합니다.	7
>	두 개의 비교 가능한 항목 사이에 사용합니다. ITEM1이 엄밀하게 ITEM2를 초과하면 참을 리턴합니다.	6
>=	두 개의 비교 가능한 항목 사이에 사용합니다. ITEM1이 ITEM2 이상이면 참을 리턴합니다.	6

표 17. CLEM 언어 연산자 (계속)

작업	주석	우선순위(다음 섹션 참조)
<	두 개의 비교 가능한 항목 사이에 사용합니다. ITEM1이 엄밀하게 ITEM2 미만이면 참을 리턴합니다.	6
<=	두 개의 비교 가능한 항목 사이에 사용합니다. ITEM1이 ITEM2 이하이면 참을 리턴합니다.	6
&&=_0	두 개의 정수 사이에 사용합니다. 부울 표현식 INT1 && INT2 = 0에 해당됩니다.	6
&&/=_0	두 개의 정수 사이에 사용합니다. 부울 표현식 INT1 && INT2 /= 0에 해당됩니다.	6
+	두 개의 수를 더합니다. NUM1 + NUM2	5
><	2개의 문자열을 연결합니다. 예를 들면, STRING1 >< STRING2.	5
-	하나의 수를 다른 하나에서 뺍니다. NUM1 - NUM2. 또한 숫자의 앞에 사용할 수도 있습니다. - NUM	5
*	2개의 수를 곱하는 데 사용합니다. NUM1 * NUM2	4
&&	두 개의 정수 사이에 사용합니다. 결과는 정수 INT1과 INT2의 bitwise 'and'입니다.	4
&&~~	두 개의 정수 사이에 사용합니다. 결과는 INT1의 bitwise 'and'와 INT2의 bitwise 보수입니다.	4
	두 개의 정수 사이에 사용합니다. 결과는 INT1과 INT2의 bitwise 'inclusive or'입니다.	4
~~	정수 앞에 사용합니다. INT의 bitwise 보수를 생성합니다.	4
/&	두 개의 정수 사이에 사용합니다. 결과는 INT1과 INT2의 bitwise 'exclusive or'입니다.	4
INT1 << N	두 개의 정수 사이에 사용합니다. N 위치만큼 왼쪽으로 이동된 INT의 비트 패턴을 생성합니다.	4
INT1 >> N	두 개의 정수 사이에 사용합니다. N 위치만큼 오른쪽으로 이동된 INT의 비트 패턴을 생성합니다.	4
/	한 수를 다른 수로 나누는 데 사용합니다. NUM1 / NUM2	4
**	두 개의 수 사이에 사용합니다. BASE ** POWER. POWER로 거듭제곱된 BASE를 리턴합니다.	3

표 17. CLEM 언어 연산자 (계속)

작업	주석	우선순위(다음 섹션 참조)
rem	두 개의 정수 사이에 사용합니다. INT1 rem INT2. 나머지를 리턴합니다. INT1 - (INT1 div INT2) * INT2	2
div	두 개의 정수 사이에 사용합니다. INT1 div INT2. 정수 나눗셈을 수행합니다.	2

연산자 우선순위

우선순위는 특히 둘 이상의 인픽스 연산자가 있는 비브래킷 표현식과 같은 복합 표현식의 구문 분석을 결정합니다. 예를 들어,

3 + 4 * 5

상대적인 우선순위에서 *가 +보다 먼저 구문 분석된다고 지시되므로 $(3 + 4) * 5$ 가 아니라 $3 + (4 * 5)$ 로 구문 분석됩니다. CLEM 언어의 모든 연산자에는 이와 연관된 우선순위 값이 있습니다. 이 값이 낮을 수록 구문 분석 목록에서 더 중요합니다. 즉, 우선순위 값이 높은 다른 연산자보다 먼저 처리된다는 의미입니다.

함수 참조

IBM SPSS Modeler에서 데이터에 대한 작업에 사용할 수 있는 CLEM 함수는 다음과 같습니다. 이러한 함수를 여러 대화 상자(예: 파생 및 플래그로 설정 노드)에서 코드로 입력하거나 표현식 작성기를 사용하여 함수 목록이나 필드 이름을 암기하지 않고 유효한 CLEM 표현식을 작성할 수 있습니다.

표 18. IBM SPSS Modeler 데이터에 사용할 CLEM 함수

함수 유형	설명
정보	필드 값을 간파하는 데 사용합니다. 예를 들어, <code>is_string</code> 함수는 유형이 문자열인 모든 레코드에 대해 참을 리턴합니다.
변환	새 필드를 구성하거나 저장 유형을 변환하는 데 사용합니다. 예를 들어, <code>to_timestamp</code> 함수는 선택된 필드를 시간소인으로 변환합니다.
비교	필드 값을 서로 비교하거나 지정된 문자열과 비교하는 데 사용합니다. 예를 들어, <code><=</code> 은 두 필드의 값이 작거나 같은지 여부를 비교하는 데 사용합니다.
논리	논리 연산(예: <code>if, then, else</code> 연산)을 수행하는 데 사용합니다.
숫자	수치 연산(예: 필드 값의 자연 로그)을 수행하는 데 사용합니다.
삼각	삼각 연산(예: 지정된 각의 아크코사인)을 수행하는 데 사용합니다.
확률	다양한 분포를 기반으로 하는 확률(예: 스튜던트의 t 분포 값이 지정된 값보다 작은 확률)을 리턴합니다.
공간	지리 공간 데이터에 대한 공간 연산을 수행하는 데 사용합니다.
Bitwise	정수를 비트 패턴으로 조작하는 데 사용합니다.
Random	임의로 항목을 선택하거나 숫자를 생성하는 데 사용합니다.
문자열	문자열에 대한 다양한 연산(예: 지정된 문자를 제거할 수 있게 하는 <code>stripchar</code>)을 수행하는 데 사용합니다.
SoundEx	정확한 맞춤법을 알 수 없는 경우 문자열을 찾는 데 사용되며, 특정 문자가 발음되는 방법에 대한 음성 가정을 기반으로 합니다.

표 18. IBM SPSS Modeler 데이터에 사용할 CLEM 함수 (계속)

함수 유형	설명
<u>날짜 및 시간</u>	날짜, 시간 및 시간소인 필드에 대한 다양한 연산을 수행하는 데 사용합니다.
<u>시퀀스</u>	데이터 세트의 레코드 시퀀스를 간파하거나 해당 시퀀스를 기반으로 하는 연산을 수행하는 데 사용합니다.
<u>글로벌</u>	전역값 설정 노드에 의해 작성된 글로벌 값에 액세스하는 데 사용합니다. 예를 들어, @MEAN은 전체 데이터 세트에서 필드의 모든 값의 평균을 참조하는 데 사용합니다.
<u>공백 및 널</u>	사용자 지정 공백 또는 시스템 결측값에 액세스하고 플래그를 지정하며 자주 채우는 데 사용합니다. 예를 들어, @BLANK(FIELD)는 공백이 있는 레코드에 true 플래그를 지정하는 데 사용합니다.
<u>특수 필드</u>	검사 중인 특정 필드를 표시하는 데 사용합니다. 예를 들어, @FIELD는 다중 필드 파생 시 사용합니다.

함수 설명 규칙

함수의 항목을 참조할 때 이 안내서 전반에서 사용하는 규칙은 다음과 같습니다.

표 19. 함수 설명 규칙

규칙	설명
<i>BOOL</i>	부울 또는 플래그(예: true 또는 false)입니다.
<i>NUM, NUM1, NUM2</i>	숫자입니다.
<i>REAL, REAL1, REAL2</i>	실수입니다(예: 1.234 또는 -77.01).
<i>INT, INT1, INT2</i>	정수입니다(예: 1 또는 -77).
<i>CHAR</i>	문자 코드입니다(예: 'A').
<i>STRING</i>	문자열입니다(예: "referrerID").
<i>LIST</i>	항목 목록입니다(예: ["abc" "def"]).
<i>ITEM</i>	필드입니다(예: Customer 또는 extract_concept).
<i>DATE</i>	start_date와 같은 날짜 필드입니다(여기서 값은 DD-MON-YYYY와 같은 형식임).
<i>TIME</i>	power_flux와 같은 시간 필드입니다(여기서 값은 HHMMSS와 같은 형식임).

이 안내서에서 함수는 한 열에는 함수, 다른 열에는 결과 유형(정수, 문자열 등), 세 번째 열에는 설명(사용 가능한 경우)과 함께 나열되어 있습니다. 예를 들어, *rem* 함수에 대한 설명은 다음과 같습니다.

표 20. *rem* 함수 설명

함수	결과	설명
<i>INT1 rem INT2</i>	숫자	<i>INT1</i> 을 <i>INT2</i> 로 나눈 나머지를 리턴합니다(예: <i>INT1</i> - (<i>INT1</i> div <i>INT2</i>) * <i>INT2</i>).

항목 나열 또는 함수의 문자 지정과 같은 사용 규칙에 대한 세부사항은 다른 곳에 설명되어 있습니다. 자세한 정보는 143 페이지의 『CLEM 데이터 유형』의 내용을 참조하십시오.

정보 함수

정보 함수는 특정 필드 값을 간파하는 데 사용합니다. 일반적으로 플래그 필드를 파생시키는 데 사용합니다. 예를 들어, @BLANK 함수를 사용하여 선택된 필드에 대해 값이 공백인 레코드를 표시하는 플래그 필드를 작성할 수 있습니다. 마찬가지로 is_string과 같은 저장 유형 함수를 사용하여 필드의 저장 유형을 확인할 수 있습니다.

표 21. CLEM 정보 함수

함수	결과	설명
@BLANK(FIELD)	부울	업스트림 유형 노드 또는 소스 노드(유형 탭)에 설정된 공백 처리 규칙에 따라 값이 공백인 모든 레코드에 대해 참을 리턴합니다.
@NULL(ITEM)	부울	값이 정의되지 않은 모든 레코드에 대해 참을 리턴합니다. 정의되지 않은 값은 시스템 널값입니다(IBM SPSS Modeler에 \$null\$로 표시됨).
is_date(ITEM)	부울	유형이 날짜인 모든 레코드에 대해 참을 리턴합니다.
is_datetime(ITEM)	부울	유형이 날짜, 시간 또는 시간소인인 모든 레코드에 대해 참을 리턴합니다.
is_integer(ITEM)	부울	유형이 정수인 모든 레코드에 대해 참을 리턴합니다.
is_number(ITEM)	부울	유형이 숫자인 모든 레코드에 대해 참을 리턴합니다.
is_real(ITEM)	부울	유형이 실수인 모든 레코드에 대해 참을 리턴합니다.
is_string(ITEM)	부울	유형이 문자열인 모든 레코드에 대해 참을 리턴합니다.
is_time(ITEM)	부울	유형이 시간인 모든 레코드에 대해 참을 리턴합니다.
is_timestamp(ITEM)	부울	유형이 시간소인인 모든 레코드에 대해 참을 리턴합니다.

변환 함수

변환 함수를 사용하여 새 필드를 구성하고 기존 파일의 저장 유형을 변환할 수 있습니다. 예를 들어, 문자열을 함께 결합하거나 문자열을 분해하여 새 문자열을 형성할 수 있습니다. 두 개의 문자열을 결합하려면 >< 연산자를 사용하십시오. 예를 들어, Site 필드 값이 "BRAMLEY"인 경우 "xx" >< Site는 "xxBRAMLEY"를 리턴합니다. 인수가 문자열이 아니더라도 ><의 결과는 항상 문자열입니다. 따라서 V1 필드가 3이고 V2 필드가 5인 경우 V1 >< V2는 "35"(숫자가 아닌 문자열)를 리턴합니다.

변환 함수(및 날짜 또는 시간 값과 같은 특정 유형의 입력을 필요로 하는 기타 함수)는 스트림 옵션 대화 상자에 지정된 현재 형식에 따라 다릅니다. 예를 들어, 값이 Jan 2003, Feb 2003 등인 문자열 필드를 변환하려면 매치하는 날짜 형식 **MON YYYY**를 스트림의 기본 날짜 형식으로 선택하십시오.

표 22. CLEM 변환 함수

기능	결과	설명
ITEM1 >< ITEM2	문자열	두 필드의 값을 연결하고 결과 문자열을 ITEM1ITEM2로 리턴합니다.
to_integer(ITEM)	정수	지정된 필드의 저장 공간을 정수로 변환합니다.
to_real(ITEM)	실수	지정된 필드의 저장 공간을 실수로 변환합니다.
to_number(ITEM)	숫자	지정된 필드의 저장 공간을 숫자로 변환합니다.
to_string(ITEM)	문자열	지정된 필드의 저장 공간을 문자열로 변환합니다. 이 함수를 사용하여 실수가 문자열로 변환되면 소수점 뒤 6자리의 값을 리턴합니다.
to_time(ITEM)	시간	지정된 필드의 저장 공간을 시간으로 변환합니다.

표 22. CLEM 변환 함수 (계속)

기능	결과	설명
to_date(ITEM)	날짜	지정된 필드의 저장 공간을 날짜로 변환합니다.
to_timestamp(ITEM)	시간소인	지정된 필드의 저장 공간을 시간소인으로 변환합니다.
to_datetime(ITEM)	날짜/시간	지정된 필드의 저장 공간을 날짜, 시간 또는 시간소인 값으로 변환합니다.
datetime_date(ITEM)	날짜	숫자, 문자열 또는 시간소인의 날짜 값을 리턴합니다. 이는 숫자(초)를 날짜로 다시 변환할 수 있게 하는 유일한 함수입니다. ITEM이 문자열인 경우 현재 날짜 형식으로 문자열을 구문 분석하여 날짜를 생성합니다. 이 함수가 성공하려면 스트림 특성 대화 상자에 지정된 날짜 형식이 올바라야 합니다. ITEM이 숫자인 경우 기준 날짜(또는 에포크) 이후의 초 단위로 해석됩니다. 1일의 분수는 잘립니다. ITEM이 시간소인인 경우 시간소인의 날짜 부분이 리턴됩니다. ITEM이 날짜인 경우 변경되지 않고 리턴됩니다.
stb_centroid_latitude(ITEM)	정수	Geohash 인수의 중심값에 해당하는 위도의 정수 값을 리턴합니다.
stb_centroid_longitude(ITEM)	정수	Geohash 인수의 중심값에 해당하는 경도의 정수 값을 리턴합니다.
to_geohash(ITEM)	문자열	<p>밀도에 대해 지정된 비트 수를 사용하여 위도 및 경도에 해당하는 Geohash 문자열을 리턴합니다.</p> <p>Geohash는 위도 및 경도 세부사항을 기반으로 지리 좌표 세트를 식별하는 데 사용하는 코드입니다. to_geohash에 대한 세 매개변수는 다음과 같습니다.</p> <ul style="list-style-type: none"> 위도: (-180, 180) 범위이고, 단위는 WGS84 좌표계에서 도입니다. 경도: (-90, 90) 범위이고, 단위는 WGS84 좌표계에서 도입니다. 비트: 해시 저장에 사용할 비트 수입니다. 범위는 [1, 75]입니다. 리턴된 문자열의 길이(5비트당 1자가 사용됨)와 해시 정확도에 모두 영향을 미칩니다. 예를 들어 5비트(1자)는 약 2500킬로미터를 나타내고, 45비트(9자)는 약 2.3미터를 나타냅니다.

비교 함수

비교 함수는 필드 값을 서로 비교하거나 지정된 문자열과 비교하는 데 사용합니다. 예를 들어, `=`을 사용하여 문자열이 같은지 확인할 수 있습니다. 문자열이 같은지 확인하는 예제는 `Class = "class 1"`입니다.

수치 비교를 위해 `greater`는 양의 무한대에 가까움을 의미하고 `lesser`는 음의 무한대에 가까움을 의미합니다. 즉, 모든 음수는 임의의 양수보다 작습니다.

표 23. CLEM 비교 함수

함수	결과	설명
count_equal(ITEM1, LIST)	정수	ITEM1과 같은 필드 목록의 값 수를 리턴하거나 ITEM1이 널인 경우 널을 리턴합니다.
count_greater_than(ITEM1, LIST)	정수	ITEM1보다 큰 필드 목록의 값 수를 리턴하거나 ITEM1이 널인 경우 널을 리턴합니다.

표 23. CLEM 비교 함수 (계속)

함수	결과	설명
count_less_than(ITEM1, LIST)	정수	ITEM1보다 작은 필드 목록의 값 수를 리턴하거나 ITEM1이 널인 경우 널을 리턴합니다.
count_not_equal(ITEM1, LIST)	정수	ITEM1과 같지 않은 필드 목록의 값 수를 리턴하거나 ITEM1이 널인 경우 널을 리턴합니다.
count_nulls(LIST)	정수	필드 목록의 널값 수를 리턴합니다.
count_non_nulls(LIST)	정수	필드 목록의 널이 아닌 값 수를 리턴합니다.
date_before(DATE1, DATE2)	부울	날짜 값 순서를 확인하는 데 사용합니다. DATE1이 DATE2 이전이면 참 값을 리턴합니다.
first_index(ITEM, LIST)	정수	필드 LIST에서 ITEM이 들어 있는 첫 번째 필드의 지수를 리턴하거나 값을 찾을 수 없으면 0을 리턴합니다. 문자열, 정수 및 실수 유형에만 지원됩니다.
first_non_null(LIST)	모두	제공된 필드 목록에서 널이 아닌 첫 번째 값을 리턴합니다. 모든 저장 유형이 지원됩니다.
first_non_null_index(LIST)	정수	지정된 LIST에서 널이 아닌 값이 들어 있는 첫 번째 필드의 지수를 리턴하거나 모든 값이 널인 경우 0을 리턴합니다. 모든 저장 유형이 지원됩니다.
ITEM1 = ITEM2	부울	ITEM1이 ITEM2와 같은 레코드에 대해 참을 리턴합니다.
ITEM1 /= ITEM2	부울	두 문자열이 동일하지 않으면 참을 리턴하거나 동일하면 0을 리턴합니다.
ITEM1 < ITEM2	부울	ITEM1이 ITEM2보다 작은 레코드에 대해 참을 리턴합니다.
ITEM1 <= ITEM2	부울	ITEM1이 ITEM2보다 작거나 같은 레코드에 대해 참을 리턴합니다.
ITEM1 > ITEM2	부울	ITEM1이 ITEM2보다 큰 레코드에 대해 참을 리턴합니다.
ITEM1 >= ITEM2	부울	ITEM1이 ITEM2보다 크거나 같은 레코드에 대해 참을 리턴합니다.
last_index(ITEM, LIST)	정수	필드 LIST에서 ITEM이 들어 있는 마지막 필드의 지수를 리턴하거나 값을 찾을 수 없으면 0을 리턴합니다. 문자열, 정수 및 실수 유형에만 지원됩니다.
last_non_null(LIST)	모두	제공된 필드 목록에서 널이 아닌 마지막 값을 리턴합니다. 모든 저장 유형이 지원됩니다.
last_non_null_index(LIST)	정수	지정된 LIST에서 널이 아닌 값이 들어 있는 마지막 필드의 지수를 리턴하거나 모든 값이 널인 경우 0을 리턴합니다. 모든 저장 유형이 지원됩니다.
max(ITEM1, ITEM2)	모두	두 개의 항목 중 더 큰 항목을 리턴합니다(ITEM1 또는 ITEM2).
max_index(LIST)	정수	숫자 필드 목록에서 최대값이 들어 있는 필드의 지수를 리턴하거나 모든 값이 널인 경우 0을 리턴합니다. 예를 들어, 나열된 세 번째 필드에 최대값이 포함된 경우 지수 값 3이 리턴됩니다. 여러 필드에 최대값이 포함된 경우 첫 번째로 나열된(맨 왼쪽) 필드가 리턴됩니다.
max_n(LIST)	숫자	숫자 필드 목록의 최대값을 리턴하거나 모든 필드 값이 널인 경우 널을 리턴합니다.

표 23. CLEM 비교 함수 (계속)

함수	결과	설명
member(ITEM, LIST)	부울	ITEM이 지정된 LIST의 멤버이면 참을 리턴합니다. 그렇지 않으면, 거짓 값이 리턴됩니다. 필드 이름 목록도 지정할 수 있습니다.
min(ITEM1, ITEM2)	모두	두 개의 항목 중 더 작은 항목을 리턴합니다(ITEM1 또는 ITEM2).
min_index(LIST)	정수	숫자 필드 목록에서 최소값이 들어 있는 필드의 지수를 리턴하거나 모든 값이 널인 경우 0을 리턴합니다. 예를 들어, 나열된 세 번째 필드에 최소값이 포함된 경우 지수 값 3이 리턴됩니다. 여러 필드에 최소값이 포함된 경우 첫 번째로 나열된(맨 왼쪽) 필드가 리턴됩니다.
min_n(LIST)	숫자	숫자 필드 목록의 최소값을 리턴하거나 모든 필드 값이 널인 경우 널을 리턴합니다.
time_before(TIME1, TIME2)	부울	시간 값 순서를 확인하는 데 사용합니다. TIME1이 TIME2 이전이면 참 값을 리턴합니다.
value_at(INT, LIST)		오프셋 INT에서 나열된 각 필드의 값을 리턴하거나 오프셋이 유효한 값 범위를 벗어난 경우(즉, 1보다 작거나 나열된 필드 수보다 큼) 널(NULL)을 리턴합니다. 모든 저장 유형이 지원됩니다.

논리 함수

CLEM 표현식을 사용하여 논리 연산을 수행할 수 있습니다.

표 24. CLEM 논리 함수

함수	결과	설명
COND1 and COND2	부울	이 연산은 논리곱이며 COND1과 COND2가 둘 다 참이면 참 값을 리턴합니다. COND1이 거짓이면 COND2는 평가되지 않습니다. 따라서 COND1이 COND2의 연산이 올바른지 처음으로 검정하는 경우 논리곱을 가질 수 있습니다(예: length(Label) >= 6 및 Label(6) = 'x').
COND1 or COND2	부울	이 연산은 논리합(포함적 분리)이며 COND1 또는 COND2가 참이거나 둘 다 참이면 참 값을 리턴합니다. COND1이 참이면 COND2가 평가되지 않습니다.
not(COND)	부울	이 연산은 논리 부정이며 COND가 거짓이면 참 값을 리턴합니다. 그렇지 않으면 이 연산은 0 값을 리턴합니다.
if COND then EXPR1 else EXPR2 endif	모두	이 연산은 조건부 평가입니다. COND가 참인 경우 이 연산은 EXPR1의 결과를 리턴합니다. 그렇지 않으면 EXPR2 평가 결과가 리턴됩니다.
if COND1 then EXPR1 elseif COND2 then EXPR2 else EXPR_N endif	모두	이 연산은 다중 분기 조건부 평가입니다. COND1이 참인 경우 이 연산은 EXPR1의 결과를 리턴합니다. 그렇지 않으면 COND2가 참인 경우 이 연산은 EXPR2 평가 결과를 리턴합니다. 그렇지 않으면 EXPR_N 평가 결과가 리턴됩니다.

숫자 함수

CLEM에는 일반적으로 사용하는 다수의 숫자 함수가 포함되어 있습니다.

표 25. CLEM 숫자 함수

함수	결과	설명
-NUM	숫자	NUM을 부정하는 데 사용합니다. 반대 부호를 가진 해당 숫자를 리턴합니다.
NUM1 + NUM2	숫자	NUM1과 NUM2의 합계를 리턴합니다.
NUM1 - NUM2	숫자	NUM1에서 NUM2를 뺀 값을 리턴합니다.
NUM1 * NUM2	숫자	NUM1에 NUM2를 곱한 값을 리턴합니다.
NUM1 / NUM2	숫자	NUM1을 NUM2로 나눈 값을 리턴합니다.
INT1 div INT2	숫자	정수 나눗셈을 수행하는 데 사용합니다. INT1을 INT2로 나눈 값을 리턴합니다.
INT1 rem INT2	숫자	INT1을 INT2로 나눈 나머지를 리턴합니다(예: INT1 - (INT1 div INT2) * INT2).
INT1 mod INT2	숫자	이 함수는 더 이상 사용되지 않습니다. rem 함수를 대신 사용하십시오.
BASE ** POWER	숫자	거듭제곱 POWER로 제곱된 BASE를 리턴합니다. 여기서 둘 중 하나는 임의의 숫자입니다(단, POWER가 정수 0이 아닌 다른 유형의 0이면 BASE가 0이 아니어야 함). POWER가 정수이면 연속적으로 BASE의 거듭제곱을 곱하여 계산이 수행됩니다. 따라서 BASE가 정수이면 결과는 정수입니다. POWER가 정수 0이면 결과는 항상 BASE와 동일한 유형의 1입니다. 그렇지 않으면 POWER가 정수가 아닌 경우 결과는 $\exp(POWER * \log(BASE))$ 로 계산됩니다.
abs(NUM)	숫자	항상 같은 유형의 숫자인 NUM의 절대값을 리턴합니다.
exp(NUM)	실수	거듭제곱 NUM으로 제곱된 e를 리턴합니다(여기서 e는 자연로그의 기수임).
fracof(NUM)	실수	NUM-intof(NUM)으로 정의된 NUM의 분수 부분을 리턴합니다.
intof(NUM)	정수	인수를 정수로 자릅니다. NUM과 동일 부호의 정수 및 abs(INT) <= abs(NUM)과 같은 가장 큰 크기의 정수를 리턴합니다.
log(NUM)	실수	0이 아니어야 하는 NUM의 자연(기수 e) 로그를 리턴합니다.
log10(NUM)	실수	0이 아니어야 하는 NUM의 기수 10 로그를 리턴합니다. 이 함수는 $\log(NUM) / \log(10)$ 으로 정의됩니다.
negate(NUM)	숫자	NUM을 부정하는 데 사용합니다. 반대 부호를 가진 해당 숫자를 리턴합니다.
round(NUM)	정수	NUM이 양수인 경우 intof(NUM+0.5)를 사용하고 NUM이 음수인 경우 intof(NUM-0.5)를 사용하여 NUM을 정수로 반올림하는 데 사용됩니다.
sign(NUM)	숫자	NUM의 부호를 판별하는 데 사용합니다. 이 연산은 NUM이 정수이면 -1, 0 또는 1을 리턴합니다. NUM이 실수이면 NUM이 음수인지, 0인지, 양수인지 여부에 따라 -1.0, 0.0 또는 1.0을 리턴합니다.
sqrt(NUM)	실수	NUM의 제곱근을 리턴합니다. NUM은 양수여야 합니다.
sum_n(LIST)	숫자	숫자 필드 목록의 합계를 리턴하거나 모든 필드 값이 널인 경우 널을 리턴합니다.

표 25. CLEM 숫자 함수 (계속)

함수	결과	설명
mean_n(LIST)	숫자	숫자 필드 목록의 평균 값을 리턴하거나 모든 필드 값이 널인 경우 널을 리턴합니다.
sdev_n(LIST)	숫자	숫자 필드 목록의 표준 편차를 리턴하거나 모든 필드 값이 널인 경우 널을 리턴합니다.

삼각 함수

이 절의 모든 함수는 각을 인수로 사용하거나 각을 결과로 리턴합니다. 두 경우 모두 관련 스트림 옵션 설정을 통해 각 단위(라디안 또는 도)를 제어합니다.

표 26. CLEM 삼각 함수

함수	결과	설명
arccos(NUM)	실수	지정된 각의 아크코사인을 계산합니다.
arccosh(NUM)	실수	지정된 각의 쌍곡선 아크코사인을 계산합니다.
arcsin(NUM)	실수	지정된 각의 아크사인을 계산합니다.
arcsinh(NUM)	실수	지정된 각의 쌍곡선 아크사인을 계산합니다.
arctan(NUM)	실수	지정된 각의 아크탄젠트를 계산합니다.
arctan2(NUM_Y, NUM_X)	실수	NUM_Y / NUM_X의 아크탄젠트를 계산하며 두 수의 부호를 사용하여 사분면 정보를 도출합니다. 결과는 - pi < ANGLE <= pi (radians) - 180 < ANGLE <= 180 (degrees) 범위의 실수입니다.
arctanh(NUM)	실수	지정된 각의 쌍곡아크탄젠트를 계산합니다.
cos(NUM)	실수	지정된 각의 코사인을 계산합니다.
cosh(NUM)	실수	지정된 각의 쌍곡선 코사인을 계산합니다.
pi	실수	이 상수는 파이(pi)에 가장 가까운 값입니다.
sin(NUM)	실수	지정된 각의 사인을 계산합니다.
sinh(NUM)	실수	지정된 각의 쌍곡선 사인을 계산합니다.
tan(NUM)	실수	지정된 각의 탄젠트를 계산합니다.
tanh(NUM)	실수	지정된 각의 쌍곡 탄젠트를 계산합니다.

확률 함수

확률 함수는 다양한 분포를 기반으로 하는 확률(예: 스튜던트의 t 분포 값이 특정 값보다 작은 확률)을 리턴합니다.

표 27. CLEM 확률 함수

함수	결과	설명
cdf_chisq(NUM, DF)	실수	지정된 자유도를 가진 카이제곱 분포 값이 지정된 숫자보다 작을 확률을 리턴합니다.
cdf_f(NUM, DF1, DF2)	실수	자유도가 DF1 및 DF2인 F 분포의 값이 지정된 숫자보다 작을 확률을 리턴합니다.

표 27. CLEM 확률 함수 (계속)		
함수	결과	설명
cdf_normal(NUM, MEAN, STDDEV)	실수	지정된 평균과 표준 편차를 가진 정규 분포 값이 지정된 숫자보다 작을 확률을 리턴합니다.
cdf_t(NUM, DF)	실수	지정된 자유도를 가진 스튜던트의 <i>t</i> 분포 값이 지정된 숫자보다 작을 확률을 리턴합니다.

공간 함수

지리 공간 데이터에 공간 함수를 사용할 수 있습니다. 예를 들어, 공간 함수를 사용하여 두 지점 간의 거리, 다각형의 영역 등을 계산할 수 있습니다. 공간 솔어(within, close to 등)를 기반으로 하는 다중 지리 공간 데이터 세트 병합이 필요한 상황도 있을 수 있습니다(병합 조건을 통해 이를 수행할 수 있음).

이러한 공간 함수는 도구 > 스트림 특성 > 옵션 > 지리 공간에 지정된 좌표계와 함께 작동합니다.

참고: 이러한 공간 함수는 3차원 데이터에는 적용되지 않습니다. 3차원 데이터를 시스템으로 가져오는 경우 이러한 함수에서는 처음 두 차원만 사용합니다. Z축 값은 무시합니다.

표 28. CLEM 공간 함수		
함수	결과	설명
close_to(SHAPE, SHAPE, NUM)	부울	두 개의 모양이 서로 일정한 거리만큼 떨어져 있는지 여부를 검정합니다. 투영된 좌표계를 사용하는 경우 거리는 미터 단위입니다. 좌표계를 사용하지 않으면 임의의 단위입니다.
crosses(SHAPE, SHAPE)	부울	두 개의 모양이 서로 교차하는지 여부를 검정합니다. 이 함수는 두 개의 행 문자열 모양 또는 하나의 행 문자열과 하나의 다각형에 적합합니다.
overlap(SHAPE, SHAPE)	부울	두 개의 다각형 사이에 교차점이 있고 교차점이 두 모양 내부에 있는지 여부를 검정합니다.
within(SHAPE, SHAPE)	부울	SHAPE1 전체가 POLYGON에 포함되는지 여부를 검정합니다.
area(SHAPE)	실수	지정된 POLYGON의 영역을 리턴합니다. 투영된 시스템을 사용하는 경우 이 함수는 제곱 미터를 리턴합니다. 좌표계를 사용하지 않으면 임의의 단위입니다. 모양은 POLYGON 또는 MULTIPOLYGON이어야 합니다.
num_points(SHAPE, LIST)	정수	POLYGON 범위 내에 포함된 포인트 필드의 포인트 수 (MULTIPOINT)를 리턴합니다. SHAPE1은 POLYGON 또는 MULTIPOLYGON이어야 합니다.
distance(SHAPE, SHAPE)	실수	SHAPE1과 SHAPE2 사이의 거리를 리턴합니다. 투영된 좌표계를 사용하는 경우 이 함수는 미터를 리턴합니다. 좌표계를 사용하지 않으면 임의의 단위입니다. SHAPE1과 SHAPE2는 임의의 지리 측정 유형입니다.

Bitwise 정수 연산

이러한 함수를 사용하여 정수를 2의 보수 값을 나타내는 비트 패턴으로 조작할 수 있습니다(여기서 비트 위치 *N*은 가중치 $2^{**}N$ 을 가짐). 비트는 0부터 위로 번호가 매겨집니다. 이러한 연산은 정수의 부호 비트가 왼쪽으로 무한정 확장되는 것처럼 작동합니다. 따라서 최상위 비트 위 어디서나 양의 정수는 0 비트를, 음의 정수는 1 비트를 갖습니다.

표 29. CLEM bitwise 정수 연산

함수	결과	설명
<code>~~ INT1</code>	정수	정수 $INT1$ 의 bitwise 보수를 생성합니다. 즉, $INT1$ 에 0이 있는 각 비트 위치에 대한 결과에 1이 있습니다. $~~ INT = -(INT + 1)$ 는 항상 참입니다.
<code>INT1 INT2</code>	정수	이 연산의 결과는 $INT1$ 및 $INT2$ 의 bitwise "inclusive or"입니다. 즉, $INT1$ 이나 $INT2$ 또는 둘 다에 1이 있는 각 비트 위치에 대한 결과에 1이 있습니다.
<code>INT1 /& INT2</code>	정수	이 연산의 결과는 $INT1$ 및 $INT2$ 의 bitwise "exclusive or"입니다. 즉, $INT1$ 이나 $INT2$ 에는 1이 있지만 둘 다에는 있지 않은 각 비트 위치에 대한 결과에 1이 있습니다.
<code>INT1 && INT2</code>	정수	정수 $INT1$ 및 $INT2$ 의 bitwise "and"를 생성합니다. 즉, $INT1$ 과 $INT2$ 둘 다에 1이 있는 각 비트 위치에 대한 결과에 1이 있습니다.
<code>INT1 &&~~ INT2</code>	정수	$INT1$ 의 bitwise "and" 및 $INT2$ 의 bitwise 보수를 생성합니다. 즉, $INT1$ 에는 1이 있고 $INT2$ 에는 0이 있는 각 비트 위치에 대한 결과에 1이 있습니다. 이는 $INT1 \&& (~INT2)$ 와 동일하며 $INT2$ 에 설정된 $INT1$ 의 비트를 지우는 데 유용합니다.
<code>INT << N</code>	정수	N 위치만큼 왼쪽으로 이동한 $INT1$ 의 비트 패턴을 생성합니다. N 에 대한 음수값은 오른쪽 이동을 생성합니다.
<code>INT >> N</code>	정수	N 위치만큼 오른쪽으로 이동한 $INT1$ 의 비트 패턴을 생성합니다. N 에 대한 음수값은 왼쪽 이동을 생성합니다.
<code>INT1 &&=_0 INT2</code>	부울	부울 표현식 $INT1 \&& INT2 /== 0$ 과 같지만 보다 효율적입니다.
<code>INT1 &&/_=0 INT2</code>	부울	부울 표현식 $INT1 \&& INT2 == 0$ 과 같지만 보다 효율적입니다.
<code>integer_bitcount(INT)</code>	정수	INT 에 대한 2의 보수 표시에서 1 또는 0 비트 수를 계산합니다. INT 가 음수가 아닌 경우 N 은 1 비트 수입니다. INT 가 음수이면 0 비트 수입니다. 부호 확장으로 인해 음수가 아닌 정수에 무제한 수의 0 비트가 있거나 음의 정수에 1 비트가 있습니다. 항상 $integer_bitcount(INT) = integer_bitcount(-(INT+1))$ 입니다.
<code>integer_leastbit(INT)</code>	정수	정수 INT 에 설정된 최하위 비트의 비트 위치 N 을 리턴합니다. N 은 INT 가 정확하게 나누는 2의 최대 거듭제곱입니다.
<code>integer_length(INT)</code>	정수	2의 보수 정수로 INT 의 비트 길이를 리턴합니다. 즉, N 은 $INT < (1 << N) \text{ if } INT \geq 0 \text{ INT } \geq (-1 << N) \text{ if } INT < 0$ 인 가장 작은 정수입니다. INT 가 음수가 아닌 경우, 부호 없는 정수로 INT 를 표시하려면 최소 N 비트의 필드가 필요합니다. 또는 부호에 관계없이 부호 있는 정수로 INT 를 표시하려면 최소 $N+1$ 비트가 필요합니다.
<code>testbit(INT, N)</code>	부울	정수 INT 의 N 위치에서 비트를 검정하고 1의 경우 참이고 0의 경우 거짓인 부울 값으로 비트 N 의 상태를 리턴합니다.

랜덤 함수

임의로 항목을 선택하거나 임의로 숫자를 생성하는 데 사용하는 함수는 다음과 같습니다.

표 30. CLEM 랜덤 함수		
함수	결과	설명
oneof(LIST)	모두	임의로 선택된 LIST의 요소를 리턴합니다. 목록 항목은 [ITEM1, ITEM2, ..., ITEM_N]으로 입력해야 합니다. 필드 이름 목록도 지정할 수 있습니다.
random(NUM)	숫자	1에서 시작하여 NUM까지 같은 유형(INT 또는 REAL)의 균등 분포된 난수를 리턴합니다. 정수를 사용하는 경우 정수만 리턴됩니다. 실수 (10진수)를 사용하는 경우 실수가 리턴됩니다(스트림 옵션이 판별한 10진수 정밀도). 이 함수가 리턴하는 최대 난수는 NUM과 같을 수 있습니다.
random0(NUM)	숫자	random(NUM)과 같은 특성을 갖지만 0에서 시작합니다. 이 함수가 리턴하는 최대 난수는 NUM과 같지 않습니다.

문자열 함수

CLEM에서 문자열에 대해 다음 작업을 수행할 수 있습니다.

- 문자열 비교
- 문자열 작성
- 문자 액세스

CLEM에서 문자열은 매치하는 큰따옴표("string quotes") 사이의 문자열 시퀀스입니다. 문자(CHAR)는 단일 영숫자입니다. '<character>' 양식으로(예: 'z', 'A' 또는 '2') 작은 역따옴표를 사용하여 CLEM 표현식에 선언됩니다. 범위를 벗어나거나 문자열에 음수 지수인 문자를 사용하면 정의되지 않은 동작이 발생합니다.

참고: SQL 푸시백을 사용하는 문자열과 사용하지 않는 문자열 간의 비교는 후미 공백이 존재하는 서로 다른 결과를 생성할 수 있습니다.

표 31. CLEM 문자열 함수		
함수	결과	설명
allbutfirst(N, STRING)	문자열	처음 N개의 문자가 제거된 STRING인 문자열을 리턴합니다.
allbutlast(N, STRING)	문자열	마지막 문자가 제거된 STRING인 문자열을 리턴합니다.
alphabefore(STRING1, STRING2)	부울	문자열의 알파벳 순서를 확인하는 데 사용합니다. STRING1이 STRING2에 선행하면 참을 리턴합니다.
endstring(LENGTH, STRING)	문자열	지정된 문자열에서 마지막 N개의 문자를 추출합니다. 문자열 길이가 지정된 길이보다 작거나 같으면 변경되지 않습니다.
hasendstring(STRING, SUBSTRING)	정수	이 함수는 isendstring(SUBSTRING, STRING)과 같습니다.
hasmidstring(STRING, SUBSTRING)	정수	이 함수는 ismidstring(SUBSTRING, STRING)(내장된 하위 문자열)과 같습니다.
hasstartstring(STRING, SUBSTRING)	정수	이 함수는 isstartstring(SUBSTRING, STRING)과 같습니다.
hassubstring(STRING, N, SUBSTRING)	정수	이 함수는 issubstring(SUBSTRING, N, STRING)과 같습니다(여기서 N 기본값은 1임).

표 31. CLEM 문자열 함수 (계속)

함수	결과	설명
count_substring(STRING, SUBSTRING)	정수	문자열 내에서 지정된 하위 문자열이 발생하는 횟수를 리턴합니다. 예를 들어, count_substring("foooo.txt", "oo")은 3을 리턴합니다.
hassubstring(STRING, SUBSTRING)	정수	이 함수는 issubstring(SUBSTRING, 1, STRING)과 같습니다(여기서 <i>N</i> 기본값은 1임).
isalphacode(CHAR)	부울	CHAR이 문자 코드가 문자인 지정된 문자열(대개 필드 이름)의 문자인 경우 참 값을 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다. (예: isalphacode(produce_num(1))).
isendstring(SUBSTRING, STRING)	정수	STRING 문자열이 SUBSTRING 하위 문자열로 끝나는 경우 이 함수는 STRING의 SUBSTRING의 정수 아래첨자를 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다.
islowercode(CHAR)	부울	CHAR이 지정된 문자열(대개 필드 이름)의 소문자인 경우 참 값을 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다. 예를 들어, islowercode('`')와 islowercode(country_name(2))는 둘 다 유효한 표현식입니다.
ismidstring(SUBSTRING, STRING)	정수	SUBSTRING이 STRING의 하위 문자열이지만 STRING의 첫 번째 문자에서 시작하거나 마지막 문자에서 끝나지 않는 경우 이 함수는 하위 문자열이 시작되는 아래첨자를 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다.
isnumbercode(CHAR)	부울	CHAR이 문자 코드가 숫자인 지정된 문자열(대개 필드 이름)의 문자인 경우 참 값을 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다. (예: isnumbercode(product_id(2))).
isstartstring(SUBSTRING, STRING)	정수	STRING 문자열이 SUBSTRING 하위 문자열로 시작하는 경우 이 함수는 아래첨자 1을 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다.
issubstring(SUBSTRING, N, STRING)	정수	<i>N</i> 번째 문자에서부터 시작하여 SUBSTRING 문자열에서 STRING 문자열과 같은 하위 문자열을 검색합니다. 찾으면 이 함수는 매치하는 하위 문자열이 시작되는 정수 아래첨자를 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다. <i>N</i> 이 주어지지 않으면 이 함수의 기본값은 1로 설정됩니다.
issubstring(SUBSTRING, STRING)	정수	<i>N</i> 번째 문자에서부터 시작하여 SUBSTRING 문자열에서 STRING 문자열과 같은 하위 문자열을 검색합니다. 찾으면 이 함수는 매치하는 하위 문자열이 시작되는 정수 아래첨자를 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다. <i>N</i> 이 주어지지 않으면 이 함수의 기본값은 1로 설정됩니다.

표 31. CLEM 문자열 함수 (계속)

함수	결과	설명
issubstring_count(SUBSTRING, N, STRING):	정수	지정된 STRING 내에서 SUBSTRING의 N번째 발생 지수를 리턴합니다. SUBSTRING의 N번 발생보다 적으면 0이 리턴됩니다.
issubstring_lim(SUBSTRING, N, STARTLIM, ENDLIM, STRING)	정수	이 함수는 issubstring과 같지만 매치는 아래첨자 STARTLIM이나 그 앞에서 시작하거나 아래첨자 ENDLIM이나 그 앞에서 끝나도록 제한됩니다. 둘 중 어느 한 인수에 거짓 값을 제공하여 STARTLIM 또는 ENDLIM 제한조건을 사용하지 않도록 설정할 수 있습니다. 예를 들어, issubstring_lim(SUBSTRING, N, false, false, STRING)은 issubstring과 같습니다.
isuppercode(CHAR)	부울	CHAR이 대문자인 경우 참 값을 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다. 예를 들어, isuppercode(``)와 isuppercode(country_name(2))는 둘 다 유효한 표현식입니다.
last(CHAR)	문자열	STRING(한 문자 이상이어야 함)의 마지막 문자 CHAR을 리턴합니다.
length(STRING)	정수	STRING 문자열의 길이(즉, 문자열 안의 문자 수)를 리턴합니다.
locchar(CHAR, N, STRING)	정수	<p>기호 필드에서 문자 위치를 식별하는 데 사용합니다. 이 함수는 STRING 문자열에서 CHAR 문자를 검색합니다(STRING의 N번째 문자에서 검색을 시작함). 이 함수는 문자를 찾은 위치(N에서 시작)를 표시하는 값을 리턴합니다. 문자를 찾을 수 없는 경우 이 함수는 0 값을 리턴합니다. 이 함수에 유효하지 않은 오프셋 (N)이 있는 경우(예: 문자열 길이를 벗어난 오프셋) 이 함수는 \$null\$을 리턴합니다.</p> <p>예를 들어, locchar(`n`, 2, web_page)는 web_page 필드에서 필드 값의 두 번째 문자에서 시작하여 'n' 문자를 검색합니다.</p> <p>참고: 지정된 문자를 캡슐화하려면 작은 역따옴표를 사용하십시오.</p>

표 31. CLEM 문자열 함수 (계속)

함수	결과	설명
locchar_back(CHAR, N, STRING)	정수	locchar과 비슷하나, <i>N</i> 번째 문자에서 시작하여 역방향으로 검색이 수행됩니다. 예를 들어, locchar_back(`n`, 9, web_page)는 9번째 문자에서 시작하여 문자열의 시작 방향으로 거꾸로 이동하여 web_page 필드를 검색합니다. 이 함수에 유효하지 않은 오프셋(예: 문자열 길이를 벗어난 오프셋)이 있는 경우 이 함수는 \$null\$을 리턴합니다. 이상적으로는 필드 현재 값의 길이를 동적으로 사용하려면 locchar_back을 length(<field>) 함수와 함께 사용해야 합니다(예: locchar_back(`n`, length(web_page)), web_page)).
lowertoupper(CHAR) lowertoupper (STRING)	CHAR 또는 문자열	입력은 이 함수에서 같은 유형의 새 항목을 리턴하는 데 사용하는 문자열 또는 문자입니다(소문자는 해당 대문자로 변환됨). 예를 들어, lowertoupper(`a`), lowertoupper("My string"), lowertoupper(field_name(2))는 모두 유효한 표현식입니다.
matches	부울	문자열이 지정된 패턴과 매치하면 참을 리턴합니다. 패턴은 문자열 리터럴이어야 합니다. 패턴을 포함하는 필드 이름이 아니어야 합니다. 물음표(?)를 패턴에 포함시켜 하나의 문자와 정확하게 매치할 수 있습니다. 별표(*)는 0개 이상의 문자와 매치됩니다. 리터럴 물음표나 별표와 매치하려면(이를 와일드카드로 사용하지 않음) 백슬래시(\)를 이스케이프 문자로 사용할 수 있습니다.
replace(SUBSTRING, NEWSUBSTRING, STRING)	문자열	지정된 STRING 내에서 SUBSTRING의 인스턴스를 NEWSUBSTRING으로 모두 바꿉니다.
replicate(COUNT, STRING)	문자열	지정된 횟수만큼 복사된 원래 문자열로 구성된 문자열을 리턴합니다.
stripchar(CHAR,STRING)	문자열	지정된 문자를 문자열 또는 필드에서 제거할 수 있게 합니다. 이 함수를 사용하여 통화 표기법과 같은 추가 기호를 데이터에서 제거하여 단순한 숫자나 이름을 만들 수 있습니다. 예를 들어, stripchar(`\$`, 'Cost') 구문을 사용하면 모든 값에서 달러 부호가 제거된 새 필드가 리턴됩니다. 참고: 지정된 문자를 캡슐화하려면 작은 역따옴표를 사용하십시오.

표 31. CLEM 문자열 함수 (계속)

함수	결과	설명
skipchar(CHAR, N, STRING)	정수	<p><i>N</i>번째 문자에서 시작하여 <i>STRING</i> 문자열에서 <i>CHAR</i> 이외의 다른 문자를 검색합니다. 이 함수는 문자를 찾은 위치를 표시하는 정수 하위 문자열을 리턴하거나 <i>N</i>번째 문자 이후의 모든 문자가 <i>CHAR</i>이면 0을 리턴합니다. 이 함수에 유효하지 않은 오프셋(예: 문자열 길이를 벗어난 오프셋)이 있는 경우 이 함수는 \$null\$을 리턴합니다.</p> <p><i>locchar</i>은 종종 <i>skipchar</i> 함수와 함께 사용되어 <i>N</i>(문자열 검색을 시작할 위치) 값을 판별합니다(예: <i>skipchar('s', (locchar('s', 1, "MyString")), "MyString"))</i>.</p>
skipchar_back(CHAR, N, STRING)	정수	<i>skipchar</i> 과 비슷하나, <i>N</i> 번째 문자에서 시작하여 역방향으로 검색이 수행됩니다.
startstring(LENGTH, STRING)	문자열	지정된 문자열에서 처음 <i>N</i> 개의 문자를 추출합니다. 문자열 길이가 지정된 길이보다 작거나 같으면 변경되지 않습니다.
strmember(CHAR, STRING)	정수	<i>locchar(CHAR, 1, STRING)</i> 과 같습니다. <i>CHAR</i> 이 처음 나타나는 위치를 표시하는 정수 하위 문자열이나 0을 리턴합니다. 이 함수에 유효하지 않은 오프셋(예: 문자열 길이를 벗어난 오프셋)이 있는 경우 이 함수는 \$null\$을 리턴합니다.
subscrs(N, STRING)	CHAR	입력 문자열 <i>STRING</i> 의 <i>N</i> 번째 문자 <i>CHAR</i> 을 리턴합니다. 이 함수를 간단한 양식으로 <i>STRING(N)</i> 으로 작성할 수도 있습니다. 예를 들어, <i>lowertoupper("name"(1))</i> 은 유효한 표현식입니다.
substring(N, LEN, STRING)	문자열	아래첨자 <i>N</i> 의 문자에서 시작하여 <i>STRING</i> 문자열의 <i>LEN</i> 문자로 구성되는 <i>SUBSTRING</i> 문자열을 리턴합니다.
substring_between(N1, N2, STRING)	문자열	아래첨자 <i>N1</i> 에서 시작하고 아래첨자 <i>N2</i> 에서 끝나는 <i>STRING</i> 의 하위 문자열을 리턴합니다.
trim(STRING)	문자열	지정된 문자열에서 선행 및 후행 공백 문자를 제거합니다.
trim_start(STRING)	문자열	지정된 문자열에서 선행 공백 문자를 제거합니다.
trimend(STRING)	문자열	지정된 문자열에서 후행 공백 문자를 제거합니다.
unicode_char(NUM)	CHAR	입력은 16진 값이 아닌 10진수여야 합니다. 유니코드 값 <i>NUM</i> 을 포함한 문자를 리턴합니다.
unicode_value(CHAR)	NUM	<i>CHAR</i> 의 유니코드 값을 리턴합니다.

표 31. CLEM 문자열 함수 (계속)		
함수	결과	설명
uppertolower(CHAR) uppertolower (STRING)	CHAR 또는 문자열	입력은 이 함수에서 같은 유형의 새 항목을 리턴하는 데 사용하는 문자열 또는 문자입니다(대문자는 해당 소문자로 변환됨). 참고: 큰따옴표로 문자열을 지정하고, 작은 역따옴표로 문자를 지정하십시오. 단순 필드 이름은 따옴표 없이 지정해야 합니다.

SoundEx 함수

SoundEx는 사운드는 알려져 있지만 정확한 맞춤법을 모를 경우 문자열을 찾는 데 사용하는 메소드입니다. 1918에 개발되었으며 특정 문자가 발음되는 방법에 대한 음성학적 가정을 기반으로 사운드가 유사한 단어를 검색합니다. 데이터베이스에서 이름을 검색하는 데 사용할 수 있습니다(예를 들어, 유사한 이름에 대해 맞춤법과 발음이 다른 경우). 기본 SoundEx 알고리즘은 알려진 제한사항(예를 들어, ph 및 f와 같은 선행 문자 조합은 동일한 것처럼 들리더라도 매치하지 않음)에도 불구하고 수많은 소스에 설명되어 있으며 대부분의 데이터베이스에서 일부 양식으로 지원됩니다.

표 32. CLEM soundex 함수		
함수	결과	설명
soundex(STRING)	정수	지정된 STRING의 4자 SoundEx 코드를 리턴합니다.
soundex_difference(STRING1, STRING2)	정수	두 개의 문자열에 대한 SoundEx 인코딩이 동일한 문자 수를 표시하는 0과 4 사이의 정수를 리턴합니다(여기서, 0은 유사성이 없음을 표시하고 4는 강력한 유사성 또는 동일한 문자열을 표시함).

날짜 및 시간 함수

CLEM은 날짜 및 시간을 나타내는 문자열 변수의 날짜/시간 저장 공간을 가진 필드를 처리하기 위한 함수 패밀리를 포함합니다. 사용하는 날짜 및 시간 형식은 각 스트림에 고유하며 스트림 특성 대화 상자에 지정됩니다. 날짜 및 시간 함수는 현재 선택된 형식에 따라 날짜 및 시간 문자열을 구문 분석합니다.

날짜에 두 자리만 사용하는 연도를 지정하는 경우(즉, 세기는 지정되지 않음) IBM SPSS Modeler는 스트림 특성 대화 상자에 지정되는 기본 세기를 사용합니다.

참고: 데이터 함수가 SQL 또는 IBM SPSS Analytic Server에 푸시백되는 경우, Analytic Server 데이터 소스를 뒤따르는 분기에서 해당 데이터 내 날짜 형식 문자열(to_date)은 SPSS Modeler 스트림에 지정된 날짜 형식과 매치해야 합니다.

표 33. CLEM 날짜 및 시간 함수		
함수	결과	설명
@TODAY	문자열	스트림 특성 대화 상자에서 룰오버 일/분(0)을 선택하는 경우 이 함수는 현재 날짜를 현재 날짜 형식의 문자열로 리턴합니다. 두 자리 숫자 날짜 형식을 사용하고 룰오버 일/분(0)을 선택하지 않는 경우 이 함수는 현재 서버에서 \$null\$을 리턴합니다.
to_time(ITEM)	시간	지정된 필드의 저장 공간을 시간으로 변환합니다.
to_date(ITEM)	날짜	지정된 필드의 저장 공간을 날짜로 변환합니다.
to_timestamp(ITEM)	시간소인	지정된 필드의 저장 공간을 시간소인으로 변환합니다.

표 33. CLEM 날짜 및 시간 함수 (계속)		
함수	결과	설명
to_datetime(ITEM)	날짜/시간	지정된 필드의 저장 공간을 날짜, 시간 또는 시간소인 값으로 변환합니다.
datetime_date(ITEM)	날짜	숫자, 문자열 또는 시간소인의 날짜 값을 리턴합니다. 이는 숫자(초)를 날짜로 다시 변환할 수 있게 하는 유일한 함수입니다. ITEM이 문자열인 경우 현재 날짜 형식으로 문자열을 구문 분석하여 날짜를 생성합니다. 이 함수가 성공 하려면 스트림 특성 대화 상자에 지정된 날짜 형식이 올발라야 합니다. ITEM이 숫자인 경우 기준 날짜(또는 에포크) 이후의 초 단위로 해석됩니다. 1일의 분수는 잘립니다. ITEM이 시간소인인 경우 시간소인의 날짜 부분이 리턴됩니다. ITEM이 날짜인 경우 변경되지 않고 리턴됩니다.
date_before(DATE1, DATE2)	부울	DATE1이 DATE2가 나타내는 날짜 또는 시간소인 이전의 날짜 또는 시간소인을 나타내는 경우 참 값을 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다.
date_days_difference(DATE1, DATE2)	정수	DATE1이 나타내는 날짜 또는 시간소인부터 DATE2가 나타내는 날짜 또는 시간소인까지의 시간(일)을 정수로 리턴합니다. DATE2가 DATE1 이전인 경우 이 함수는 음수를 리턴합니다.
date_in_days(DATE)	정수	기준 날짜부터 DATE가 나타내는 날짜 또는 시간소인까지의 시간(일)을 정수로 리턴합니다. DATE가 기준 날짜 이전인 경우 이 함수는 음수를 리턴합니다. 계산이 적절히 작동하려면 유효한 날짜를 포함시켜야 합니다. 예를 들어, 2001년 2월 29일을 날짜로 지정해서는 안 됩니다. 2001년은 윤년이 아니기 때문에 이 날짜가 존재하지 않습니다.
date_in_months(DATE)	실수	기준 날짜부터 DATE가 나타내는 날짜 또는 시간소인까지의 시간(월)을 실수로 리턴합니다. 이는 30.4375일을 1개 월로 한 대략적인 숫자입니다. DATE가 기준 날짜 이전인 경우 이 함수는 음수를 리턴합니다. 계산이 적절히 작동하려면 유효한 날짜를 포함시켜야 합니다. 예를 들어, 2001년 2월 29일을 날짜로 지정해서는 안 됩니다. 2001년은 윤년이 아니기 때문에 이 날짜가 존재하지 않습니다.
date_in_weeks(DATE)	실수	기준 날짜부터 DATE가 나타내는 날짜 또는 시간소인까지의 시간(주)을 실수로 리턴합니다. 이는 7.0일을 1주로 합니다. DATE가 기준 날짜 이전인 경우 이 함수는 음수를 리턴합니다. 계산이 적절히 작동하려면 유효한 날짜를 포함시켜야 합니다. 예를 들어, 2001년 2월 29일을 날짜로 지정해서는 안 됩니다. 2001년은 윤년이 아니기 때문에 이 날짜가 존재하지 않습니다.
date_in_years(DATE)	실수	기준 날짜부터 DATE가 나타내는 날짜 또는 시간소인까지의 시간(년)을 실수로 리턴합니다. 이는 365.25일을 1년으로 한 대략적인 숫자입니다. DATE가 기준 날짜 이전인 경우 이 함수는 음수를 리턴합니다. 계산이 적절히 작동하려면 유효한 날짜를 포함시켜야 합니다. 예를 들어, 2001년 2월 29일을 날짜로 지정해서는 안 됩니다. 2001년은 윤년이 아니기 때문에 이 날짜가 존재하지 않습니다.

표 33. CLEM 날짜 및 시간 함수 (계속)

함수	결과	설명
date_months_difference (DATE1, DATE2)	실수	DATE1이 나타내는 날짜 또는 시간소인부터 DATE2가 나타내는 날짜 또는 시간소인까지의 시간(월)을 실수로 리턴합니다. 이는 30.4375일을 1개월로 한 대략적인 숫자입니다. DATE2가 DATE1 이전인 경우 이 함수는 음수를 리턴합니다.
datetime_date(YEAR, MONTH, DAY)	날짜	주어진 YEAR, MONTH, DAY의 날짜 값을 생성합니다. 인수는 정수여야 합니다.
datetime_day(DATE)	정수	주어진 DATE 또는 시간소인의 날짜(일)를 리턴합니다. 결과는 1 - 31 범위의 정수입니다.
datetime_day_name(DAY)	문자열	주어진 DAY의 전체 이름을 리턴합니다. 인수는 1(일요일) - 7(토요일) 범위의 정수여야 합니다.
datetime_hour(TIME)	정수	TIME 또는 시간소인의 시간을 리턴합니다. 결과는 0 - 23 범위의 정수입니다.
datetime_in_seconds(TIME)	실수	TIME에 저장된 초 부분을 리턴합니다.
datetime_in_seconds(DATE), datetime_in_seconds(DATE TIME)	실수	현재 DATE 또는 DATETIME과 기준 날짜(1900-01-01) 간의 차이로부터 누적된 숫자(초로 변환됨)입니다.
datetime_minute(TIME)	정수	TIME 또는 시간소인의 분을 리턴합니다. 결과는 0 - 59 범위의 정수입니다.
datetime_month(DATE)	정수	DATE 또는 시간소인의 월을 리턴합니다. 결과는 1 - 12 범위의 정수입니다.
datetime_month_name (MONTH)	문자열	주어진 MONTH의 전체 이름을 리턴합니다. 인수는 1 - 12 범위의 정수여야 합니다.
datetime_now	시간소인	현재 시간을 시간소인으로 리턴합니다.
datetime_second(TIME)	정수	TIME 또는 시간소인의 초를 리턴합니다. 결과는 0 - 59 범위의 정수입니다.
datetime_day_short_name (DAY)	문자열	주어진 DAY의 축약 이름을 리턴합니다. 인수는 1(일요일) - 7(토요일) 범위의 정수여야 합니다.
datetime_month_short_nam e (MONTH)	문자열	주어진 MONTH의 축약 이름을 리턴합니다. 인수는 1 - 12 범위의 정수여야 합니다.
datetime_time(HOUR, MINUTE, SECOND)	시간	지정된 HOUR, MINUTE, SECOND의 시간 값을 리턴합니다. 인수는 정수여야 합니다.
datetime_time(ITEM)	시간	주어진 ITEM의 시간 값을 리턴합니다.
datetime_timestamp(YEAR, MONTH, DAY, HOUR, MINUTE, SECOND)	시간소인	주어진 YEAR, MONTH, DAY, HOUR, MINUTE, SECOND의 시간소인 값을 리턴합니다.
datetime_timestamp(DATE, TIME)	시간소인	주어진 DATE 및 TIME의 시간소인 값을 리턴합니다.
datetime_timestamp (NUMBER)	시간소인	주어진 초 단위의 시간소인 값을 리턴합니다.
datetime_weekday(DATE)	정수	주어진 DATE 또는 시간소인의 요일을 리턴합니다.

표 33. CLEM 날짜 및 시간 함수 (계속)		
함수	결과	설명
datetime_year(DATE)	정수	DATE 또는 시간소인의 연도를 리턴합니다. 결과는 2002와 같은 정수입니다.
date_weeks_difference(DATE1, DATE2)	실수	DATE1이 나타내는 날짜 또는 시간소인부터 DATE2가 나타내는 날짜 또는 시간소인까지의 시간(주)을 실수로 리턴합니다. 이는 7.0일을 1주로 합니다. DATE2가 DATE1 이전인 경우 이 함수는 음수를 리턴합니다.
date_years_difference(DATE1, DATE2)	실수	DATE1이 나타내는 날짜 또는 시간소인부터 DATE2가 나타내는 날짜 또는 시간소인까지의 시간(년)을 실수로 리턴합니다. 이는 365.25일을 1년으로 한 대략적인 숫자입니다. DATE2가 DATE1 이전인 경우 이 함수는 음수를 리턴합니다.
date_from_ywd(YEAR, WEEK, DAY)	정수	ISO 8601 표준을 사용하여 연도, 주(한 해 기준), 요일을 날짜로 변환합니다.
date_iso_day(DATE)	정수	ISO 8601 표준을 사용하여 날짜의 요일을 리턴합니다.
date_iso_week(DATE)	정수	ISO 8601 표준을 사용하여 날짜의 주(한 해 기준)를 리턴합니다.
date_iso_year(DATE)	정수	ISO 8601 표준을 사용하여 날짜의 연도를 리턴합니다.
time_before(TIME1, TIME2)	부울	TIME1이 TIME2가 나타내는 시간 또는 시간소인 이전의 시간 또는 시간소인을 나타내는 경우 참 값을 리턴합니다. 그렇지 않으면 이 함수는 0 값을 리턴합니다.
time_hours_difference(TIME1, TIME2)	실수	TIME1과 TIME2가 나타내는 시간 또는 시간소인 간의 시차(시간)를 실수로 리턴합니다. 스트림 특성 대화 상자에서 룰오버 일/분(O)을 선택하는 경우 이전 일을 참조하기 위해 상위 값 TIME1을 사용합니다. 룰오버 옵션을 선택하지 않으면 상위 값 TIME1으로 인해 리턴된 값은 음수가 됩니다.
time_in_hours(TIME)	실수	TIME이 나타내는 시간(시간)을 실수로 리턴합니다. 예를 들어, HHMM 시간 형식에서 time_in_hours('0130') 표현식은 1.5로 평가됩니다. TIME은 시간 또는 시간소인을 나타낼 수 있습니다.
time_in_mins(TIME)	실수	TIME이 나타내는 시간(분)을 실수로 리턴합니다. TIME은 시간 또는 시간소인을 나타낼 수 있습니다.
time_in_secs(TIME)	정수	TIME이 나타내는 시간(초)을 정수로 리턴합니다. TIME은 시간 또는 시간소인을 나타낼 수 있습니다.
time_mins_difference(TIME1, TIME2)	실수	TIME1과 TIME2가 나타내는 시간 또는 시간소인 간의 시차(분)를 실수로 리턴합니다. 스트림 특성 대화 상자에서 룰오버 일/분(O)을 선택하는 경우 이전 일(또는 현재 형식에 분과 초만 지정된 경우 이전 시간)을 참조하기 위해 상위 값 TIME1을 사용합니다. 룰오버 옵션을 선택하지 않으면 상위 값 TIME1으로 인해 리턴된 값은 음수가 됩니다.
time_secs_difference(TIME1, TIME2)	정수	TIME1과 TIME2가 나타내는 시간 또는 시간소인 간의 시차(초)를 정수로 리턴합니다. 스트림 특성 대화 상자에서 룰오버 일/분(O)을 선택하는 경우 이전 일(또는 현재 형식에 분과 초만 지정된 경우 이전 시간)을 참조하기 위해 상위 값 TIME1을 사용합니다. 룰오버 옵션을 선택하지 않으면 상위 값 TIME1으로 인해 리턴된 값은 음수가 됩니다.

날짜 및 시간 값 변환

변환 함수(및 날짜 또는 시간 값과 같은 특정 유형의 입력을 필요로 하는 기타 함수)는 스트림 옵션 대화 상자에 지정된 현재 형식에 따라 다릅니다. 예를 들어, 값이 Jan 2003, Feb 2003 등인 문자열로 저장된 DATE 필드가 있으면 다음과 같이 날짜 저장 공간으로 변환할 수 있습니다.

```
to_date(DATE)
```

이 변환이 작동하려면 매치하는 날짜 형식 **MON YYYY**를 스트림의 기본 날짜 형식으로 선택하십시오.

채움 노드를 사용하여 문자열 값을 날짜로 변환하는 예제는 |Demos 폴더의 *streams* 서브폴더에 설치된 *broadband_create_models.str*을 참조하십시오.

숫자로 저장된 날짜. 이전 예제에서 DATE는 필드 이름인 반면, to_date는 CLEM 함수입니다. 숫자로 저장된 날짜가 있으면 datetime_date 함수를 사용하여 변환할 수 있습니다(여기서 숫자는 기준 날짜(또는 에포크) 이후의 초 단위로 해석됨).

```
datetime_date(DATE)
```

날짜를 다수의 초로, 초를 날짜로 변환하여 현재 날짜 더하기 또는 빼기 고정 일 수와 같은 계산을 수행할 수 있습니다. 예를 들어, 다음과 같습니다.

```
datetime_date((date_in_days(DATE) - 7) * 60 * 60 * 24)
```

시퀀스 함수

일부 연산의 경우 이벤트 시퀀스가 중요합니다. 애플리케이션을 통해 다음 레코드 시퀀스에 대한 작업을 할 수 있습니다.

- 시퀀스 및 시계열
- 시퀀스 함수
- 레코드 색인화
- 값 평균, 합계 및 비교
- 변경 모니터링 - 차별
- @SINCE
- 오프셋 값
- 추가 시퀀스 기능

많은 애플리케이션의 경우, 스트림을 통과하는 각 레코드를 다른 모든 케이스와 독립된 개별 케이스로 간주할 수 있습니다. 이러한 상황에서는 일반적으로 레코드 순서가 중요하지 않습니다.

그러나 일부 문제점 클래스의 경우 레코드 시퀀스가 매우 중요합니다. 이는 일반적으로 레코드 시퀀스가 정렬된 이벤트 또는 발생 시퀀스를 나타내는 시계열 상황입니다. 각 레코드는 특정 순간의 스냅샷을 나타냅니다. 그러나 많은 양의 풍부한 정보가 순시값에 포함되지 않고 시간이 경과하면서 이러한 값이 변화하고 동작합니다.

물론 관련 모수가 시간이 아닌 다른 어떤 것일 수 있습니다. 예를 들어, 레코드가 선을 따라 거리를 두고 수행된 분석을 나타낼 수 있지만 동일한 원칙이 적용됩니다.

다음 특성을 통해 시퀀스 및 특수 함수를 즉시 인식할 수 있습니다.

- 모두 앞에 @가 붙습니다.
- 이름이 대문자입니다.

시퀀스 함수는 노드가 현재 처리 중인 레코드, 노드를 이미 통과한 레코드, 한 케이스에선 아직 노드를 통과해야 하는 레코드를 참조할 수 있습니다. 인수로 사용할 수 있는 것에 약간의 제한이 있긴 하지만 시퀀스 함수를 CLEM 표현식의 기타 구성요소와 자유롭게 혼합할 수 있습니다.

예

특정 이벤트가 발생했거나 조건이 참인 이후로 얼마나 경과했는지를 아는 것이 유용합니다. 이를 수행하려면 @SINCE 함수를 사용하십시오. 예를 들어, 다음과 같습니다.

```
@SINCE(Income > Outgoings)
```

이 함수는 이 조건이 참인 마지막 레코드의 오프셋을 리턴합니다(즉, 조건이 참이었던 이 레코드 이전의 레코드 수). 조건이 참이었던 적이 없으면 @SINCE가 @INDEX + 1을 리턴합니다.

@SINCE에서 사용하는 표현식에서 현재 레코드 값을 참조하려는 경우가 있습니다. 필드 이름이 항상 현재 레코드에 적용됨을 지정하는 @THIS 함수를 사용하여 이를 수행할 수 있습니다. 현재 레코드의 값보다 두 배가 넘는 Concentration 필드 값을 가진 마지막 레코드의 오프셋을 찾으려면 다음을 사용할 수 있습니다.

```
@SINCE(Concentration > 2 * @THIS(Concentration))
```

일부 경우 @SINCE에 지정된 조건이 정의상 현재 레코드에도 적용됩니다. 예를 들어, 다음과 같습니다.

```
@SINCE(ID == @THIS(ID))
```

이러한 이유로 @SINCE는 현재 레코드에 대해 조건을 평가하지 않습니다. 이전 레코드는 물론 현재 레코드에 대해서도 조건을 평가하려면 유사 함수인 @SINCE0을 사용하십시오. 조건이 현재 레코드에서 참이면 @SINCE0이 0을 리턴합니다.

표 34. CLEM 시퀀스 함수		
함수	결과	설명
MEAN(FIELD)	실수	지정된 FIELD 또는 FIELDS 값의 평균을 리턴합니다.
@MEAN(FIELD, EXPR)	실수	현재 레코드를 포함하여 현재 노드가 수신한 마지막 EXPR 레코드에 대한 FIELD 값의 평균을 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. EXPR은 0보다 큰 정수로 평가되는 표현식입니다. EXPR이 생략되었거나 지금까지 수신된 레코드 수를 초과하는 경우 지금까지 수신된 모든 레코드에 대한 평균이 리턴됩니다.
@MEAN(FIELD, EXPR, INT)	실수	현재 레코드를 포함하여 현재 노드가 수신한 마지막 EXPR 레코드에 대한 FIELD 값의 평균을 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. EXPR은 0보다 큰 정수로 평가되는 표현식입니다. EXPR이 생략되었거나 지금까지 수신된 레코드 수를 초과하는 경우 지금까지 수신된 모든 레코드에 대한 평균이 리턴됩니다. INT는 되돌아볼 값의 최대 수를 지정합니다. 이는 단지 두 개의 인수를 사용하는 것보다 훨씬 효율적입니다.
@DIFF1(FIELD)	실수	FIELD의 첫 번째 미분을 리턴합니다. 따라서 단일 인수 양식은 필드의 현재 값과 이전 값 간의 차이만 리턴합니다. 관련된 이전 레코드가 없으면 \$null\$을 리턴합니다.
@DIFF1(FIELD1, FIELD2)	실수	이 2인수 양식은 FIELD2와 관련하여 FIELD1의 첫 번째 미분을 제공합니다. 관련된 이전 레코드가 없으면 \$null\$을 리턴합니다. @DIFF1(FIELD1) / @DIFF1(FIELD2)로 계산됩니다.
@DIFF2(FIELD)	실수	FIELD의 두 번째 미분을 리턴합니다. 따라서 단일 인수 양식은 필드의 현재 값과 이전 값 간의 차이만 리턴합니다. 관련된 이전 레코드가 없으면 \$null\$을 리턴합니다. @DIFF2는 @DIFF(@DIFF(FIELD))로 계산됩니다.

표 34. CLEM 시퀀스 함수 (계속)

함수	결과	설명
@DIFF2(FIELD1, FIELD2)	실수	이 2인수 양식은 FIELD2와 관련하여 FIELD1의 두 번째 미분을 제공합니다. 관련된 이전 레코드가 없으면 \$null \$을 리턴합니다. 이는 복잡한 계산입니다 (@DIFF1(FIELD1)/@DIFF1(FIELD2) - @OFFSET(@DIFF1(FIELD1),1)/ @OFFSET(@DIFF1(FIELD2))) / @DIFF1(FIELD2)).
@INDEX	정수	현재 레코드의 지수를 리턴합니다. 지수는 현재 노드에 도착하면 레코드에 할당됩니다. 첫 번째 레코드에는 지수 1이 부여되며 지수는 각각의 후속 레코드에 대해 1씩 증분됩니다.
@LAST_NON_BLANK(FIELD)	모두	업스트림 소스 또는 유형 노드에 정의된 대로 공백이 아닌 FIELD의 마지막 값을 리턴합니다. 지금까지 읽은 레코드에 공백이 아닌 FIELD 값이 없으면 \$null \$이 리턴됩니다. 사용자 결측값이라고도 하는 공백값을 각 필드에 대해 개별적으로 정의할 수 있습니다.
@MAX(FIELD)	숫자	지정된 FIELD의 최대값을 리턴합니다.
@MAX(FIELD, EXPR)	숫자	현재 레코드를 포함하여 지금까지 수신된 마지막 EXPR 레코드에 대한 FIELD 최대값을 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. EXPR은 0보다 큰 정수로 평가되는 표현식입니다.
@MAX(FIELD, EXPR, INT)	숫자	현재 레코드를 포함하여 지금까지 수신된 마지막 EXPR 레코드에 대한 FIELD 최대값을 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. EXPR은 0보다 큰 정수로 평가되는 표현식입니다. EXPR이 생략되었거나 지금까지 수신된 레코드 수를 초과하는 경우 지금까지 수신된 모든 레코드에 대한 최대값이 리턴됩니다. INT는 되돌아볼 값의 최대 수를 지정합니다. 이는 단지 두 개의 인수를 사용하는 것보다 훨씬 효율적입니다.
@MIN(FIELD)	숫자	지정된 FIELD의 최소값을 리턴합니다.
@MIN(FIELD, EXPR)	숫자	현재 레코드를 포함하여 지금까지 수신된 마지막 EXPR 레코드에 대한 FIELD 최소값을 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. EXPR은 0보다 큰 정수로 평가되는 표현식입니다.
@MIN(FIELD, EXPR, INT)	숫자	현재 레코드를 포함하여 지금까지 수신된 마지막 EXPR 레코드에 대한 FIELD 최소값을 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. EXPR은 0보다 큰 정수로 평가되는 표현식입니다. EXPR이 생략되었거나 지금까지 수신된 레코드 수를 초과하는 경우 지금까지 수신된 모든 레코드에 대한 최소값이 리턴됩니다. INT는 되돌아볼 값의 최대 수를 지정합니다. 이는 단지 두 개의 인수를 사용하는 것보다 훨씬 효율적입니다.

표 34. CLEM 시퀀스 함수 (계속)

함수	결과	설명
@OFFSET(FIELD, EXPR)	모두	<p><i>EXPR</i> 값으로 현재 레코드의 레코드 오프셋에 있는 <i>FIELD</i> 값을 리턴합니다. 양수 오프셋은 이미 전달한("되돌아보기") 레코드를 참조하는 반면, 음수 오프셋은 도착해야 하는 레코드에 "미리보기"를 지정합니다. 예를 들어, @OFFSET(Status, 1)은 이전 레코드의 Status 필드 값을 리턴하는 반면, @OFFSET(Status, -4)는 값을 획득하기 위해 시퀀스에서 4개의 레코드(즉, 아직 이 노드를 통과하지 못한 레코드)를 "미리봅니다". 음수(미리보기) 오프셋을 상수로 지정해야 합니다. 양수 오프셋의 경우만 <i>EXPR</i>은 오프셋을 제공하기 위해 현재 레코드에 대해 평가되는 임의의 CLEM 표현식이기도 합니다. 이 경우, 이 함수의 3인수 버전은 성능을 향상시켜야 합니다(다음 함수 참조). 표현식이 음수가 아닌 정수 이외의 것을 리턴하는 경우 이로 인해 오류가 발생합니다. 즉, 계산된 미리보기 오프셋을 갖는 것이 올바르지 않습니다.</p> <p>참조: 자기 참조 @OFFSET 함수는 리터럴 미리보기를 사용할 수 없습니다. 예를 들어, 채움 노드에서는 @OFFSET(field1, -2)와 같은 표현식을 사용하여 field1의 값을 대체할 수 없습니다.</p> <p>참고: 채움 노드에서 필드를 채울 때 실질적으로 해당 필드의 두 개의 다른 값(즉, 사전에 채워진 값과 사후 채워진 값)이 있습니다. @OFFSET이 자체를 참조할 때는 사후 채워진 값을 참조합니다. 사후 채워진 이 값은 과거 행에 대해서만 존재하므로 자기 참조 @OFFSET은 과거 행만 참조할 수 있습니다. 자기 참조 @OFFSET은 미래를 참조할 수 없으므로 오프셋에 대한 다음 검사를 수행합니다.</p> <ul style="list-style-type: none"> • 오프셋이 리터럴이고 미래로 평가되는 경우 실행이 시작되기 전에 오류가 보고됩니다. • 오프셋이 표현식이고 런타임 시 미래로 평가되는 경우 @OFFSET은 \$null\$을 리턴합니다. <p>참고: 한 노드에서 "미리보기"와 "되돌아보기"를 둘 다 사용하는 것은 지원되지 않습니다.</p>

표 34. CLEM 시퀀스 함수 (계속)

함수	결과	설명
@OFFSET(FIELD, EXPR, INT)	모두	<p>되돌아볼 값의 최대 수를 지정하는 세 번째 인수인 <i>INT</i>를 추가하여 @OFFSET 함수와 동일한 작업을 수행합니다. 오프셋이 표현식에서 계산되는 경우 이 세 번째 인수는 성능을 향상시켜야 합니다.</p> <p>예를 들어, @OFFSET(Foo, Month, 12)와 같은 표현식에서 시스템은 Foo의 마지막 12개 값만 유지하도록 알고 있습니다. 그렇지 않으면 만약을 위해 모든 값을 저장해야 합니다. 오프셋 값이 음수의 "미리보기" 오프셋(상수여야 함)을 포함하여 상수인 경우 세 번째 인수는 무의미하며 이 함수의 2인수 버전을 사용해야 합니다. 앞서 설명한 2인수 버전의 자기 참조 함수에 대한 참고도 참조하십시오.</p> <p>참고: 한 노드에서 "미리보기"와 "되돌아보기"를 둘 다 사용하는 것은 지원되지 않습니다.</p>
@SDEV(FIELD)	실수	지정된 <i>FIELD</i> 또는 <i>FIELDS</i> 값의 표준 편차를 리턴합니다.
@SDEV(FIELD, EXPR)	실수	현재 레코드를 포함하여 현재 노드가 수신한 마지막 <i>EXPR</i> 레코드에 대한 <i>FIELD</i> 값의 표준 편차를 리턴합니다. <i>FIELD</i> 는 숫자 필드의 이름이어야 합니다. <i>EXPR</i> 은 0 보다 큰 정수로 평가되는 표현식입니다. <i>EXPR</i> 이 생략되었거나 지금까지 수신된 레코드 수를 초과하는 경우 지금까지 수신된 모든 레코드에 대한 표준 편차가 리턴됩니다.
@SDEV(FIELD, EXPR, INT)	실수	현재 레코드를 포함하여 현재 노드가 수신한 마지막 <i>EXPR</i> 레코드에 대한 <i>FIELD</i> 값의 표준 편차를 리턴합니다. <i>FIELD</i> 는 숫자 필드의 이름이어야 합니다. <i>EXPR</i> 은 0 보다 큰 정수로 평가되는 표현식입니다. <i>EXPR</i> 이 생략되었거나 지금까지 수신된 레코드 수를 초과하는 경우 지금까지 수신된 모든 레코드에 대한 표준 편차가 리턴됩니다. <i>INT</i> 는 되돌아볼 값의 최대 수를 지정합니다. 이는 단지 두 개의 인수를 사용하는 것보다 훨씬 효율적입니다.
@SINCE(EXPR)	모두	임의의 CLEM 표현식인 <i>EXPR</i> 이 참인 이후로 전달한 레코드 수를 리턴합니다.
@SINCE(EXPR, INT)	모두	두 번째 인수인 <i>INT</i> 를 추가하면 되돌아볼 레코드의 최대 수를 지정합니다. <i>EXPR</i> 이 참이었던 적이 없으면 <i>INT</i> 는 @INDEX+1입니다.
@SINCE0(EXPR)	모두	@SINCE는 고려하지 않는 반면 현재 레코드를 고려합니다. 따라서 <i>EXPR</i> 이 현재 레코드에 대해 참이면 @SINCE0이 0을 리턴합니다.
@SINCE0(EXPR, INT)	모두	두 번째 인수인 <i>INT</i> 를 추가하면 되돌아볼 레코드의 최대 수를 지정합니다.
@SUM(FIELD)	숫자	지정된 <i>FIELD</i> 또는 <i>FIELDS</i> 값의 합계를 리턴합니다.

표 34. CLEM 시퀀스 함수 (계속)

함수	결과	설명
@SUM(FIELD, EXPR)	숫자	현재 레코드를 포함하여 현재 노드가 수신한 마지막 EXPR 레코드에 대한 FIELD 값의 합계를 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. EXPR은 0보다 큰 정수로 평가되는 표현식입니다. EXPR이 생략되었거나 지금까지 수신된 레코드 수를 초과하는 경우 지금까지 수신된 모든 레코드에 대한 합계가 리턴됩니다.
@SUM(FIELD, EXPR, INT)	숫자	현재 레코드를 포함하여 현재 노드가 수신한 마지막 EXPR 레코드에 대한 FIELD 값의 합계를 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. EXPR은 0보다 큰 정수로 평가되는 표현식입니다. EXPR이 생략되었거나 지금까지 수신된 레코드 수를 초과하는 경우 지금까지 수신된 모든 레코드에 대한 합계가 리턴됩니다. INT는 되돌아볼 값의 최대 수를 지정합니다. 이는 단지 두 개의 인수를 사용하는 것보다 훨씬 효율적입니다.
@THIS(FIELD)	모두	현재 레코드에서 FIELD 필드의 값을 리턴합니다. @SINCE 표현식에서만 사용합니다.

글로벌 함수

@MEAN, @SUM, @MIN, @MAX, @SDEV 함수는 최대한으로 보면, 현재 레코드를 포함하여 지금까지 읽은 모든 레코드에서만 작동합니다. 그러나 일부 경우 현재 레코드의 값을 전체 데이터 세트에 표시된 값과 비교하는 방법을 이해하는 것이 유용합니다. 전역값 설정 노드를 사용하여 전체 데이터 세트에서 값을 생성하는 경우 글로벌 함수를 사용하여 CLEM 표현식에서 이러한 값에 액세스할 수 있습니다.

예를 들어, 다음 함수는

```
@GLOBAL_MAX(Age)
```

데이터 세트의 Age 최고값을 리턴하는 반면, 다음 표현식은

```
(Value - @GLOBAL_MEAN(Value)) / @GLOBAL_SDEV(Value)
```

이 레코드의 Value와 글로벌 평균 간의 차이를 다수의 표준 편차로 표현합니다. 전역값 설정 노드가 계산한 이후에만 글로벌 값을 사용할 수 있습니다. 스트림 특성 대화 상자의 전역값 탭에서 글로벌 값 지우기 단추를 클릭하여 현재 글로벌 값을 모두 취소할 수 있습니다.

표 35. CLEM 글로벌 함수

함수	결과	설명
@GLOBAL_MAX(FIELD)	숫자	전역값 설정 노드에 의해 이전에 생성된 전체 데이터 세트에 대한 FIELD의 최대값을 리턴합니다. FIELD는 숫자, 날짜/시간/날짜시간 또는 문자열 필드의 이름이어야 합니다. 해당 글로벌 값이 설정되지 않은 경우에는 오류가 발생합니다.
@GLOBAL_MIN(FIELD)	숫자	전역값 설정 노드에 의해 이전에 생성된 전체 데이터 세트에 대한 FIELD의 최소값을 리턴합니다. FIELD는 숫자, 날짜/시간/날짜시간 또는 문자열 필드의 이름이어야 합니다. 해당 글로벌 값이 설정되지 않은 경우에는 오류가 발생합니다.
@GLOBAL_SDEV(FIELD)	숫자	전역값 설정 노드에 의해 이전에 생성된 전체 데이터 세트에 대한 FIELD의 표준 편차 값을 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. 해당 글로벌 값이 설정되지 않은 경우에는 오류가 발생합니다.

표 35. CLEM 글로벌 함수 (계속)

함수	결과	설명
@GLOBAL_MEAN(FIELD)	숫자	전역값 설정 노드에 의해 이전에 생성된 전체 데이터 세트에 대한 FIELD 값의 평균을 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. 해당 글로벌 값이 설정되지 않은 경우에는 오류가 발생합니다.
@GLOBAL_SUM(FIELD)	숫자	전역값 설정 노드에 의해 이전에 생성된 전체 데이터 세트에 대한 FIELD의 합계를 리턴합니다. FIELD는 숫자 필드의 이름이어야 합니다. 해당 글로벌 값이 설정되지 않은 경우에는 오류가 발생합니다.

공백 및 널값 처리 함수

CLEM을 사용하여 필드의 특정 값이 "공백" 또는 결측값으로 간주되도록 지정할 수 있습니다. 공백에 대해 작동하는 함수는 다음과 같습니다.

표 36. CLEM 공백 및 널값 함수

함수	결과	설명
@BLANK(FIELD)	부울	업스트림 유형 노드 또는 소스 노드(유형 템)에 설정된 공백 처리 규칙에 따라 값이 공백인 모든 레코드에 대해 참을 리턴합니다.
@LAST_NON_BLANK(FIELD)	모두	업스트림 소스 또는 유형 노드에 정의된 대로 공백이 아닌 FIELD의 마지막 값을 리턴합니다. 지금까지 읽은 레코드에 공백이 아닌 FIELD 값이 없으면 \$null\$이 리턴됩니다. 사용자 결측값이라고도 하는 공백값을 각 필드에 대해 개별적으로 정의할 수 있습니다.
@NULL(FIELD)	부울	FIELD 값이 시스템 결측값 \$null\$이면 참을 리턴합니다. 사용자 정의 공백을 포함하여 기타 모든 값에 대해서는 거짓을 리턴합니다. 둘 다를 확인하려면 @BLANK(FIELD) 및 @NULL(FIELD)를 사용하십시오.
undef	모두	\$null\$ 값을 입력하기 위해(예: 공백값을 채움 노드의 널값으로 채우기 위해) CLEM에서 일반적으로 사용합니다.

공백 필드는 채움 노드로 "채워"집니다. 채움 노드와 파생 노드(다중 모드만) 둘 다에서 특수 CLEM 함수 @FIELD는 검사 중인 현재 필드를 참조합니다.

특수 필드

특수 함수는 검사 중인 특정 필드를 표시하거나 필드 목록을 입력으로 생성하는 데 사용합니다. 예를 들어, 동시에 여러 필드를 파생시키는 경우 @FIELD를 사용하여 "선택된 필드에 대해 이 파생 동작을 수행함"을 표시해야 합니다. `log(@FIELD)` 표현식을 사용하면 선택된 각 필드에 대해 새 로그 필드가 파생됩니다.

표 37. CLEM 특수 필드

함수	결과	설명
@FIELD	모두	표현식 컨텍스트에 지정된 모든 필드에 대해 동작을 수행합니다.

표 37. CLEM 특수 필드 (계속)

함수	결과	설명
@TARGET	모두	CLEM 표현식을 사용자 정의 분석 함수에서 사용하는 경우 @TARGET은 목표 필드 또는 분석 중인 목표/예측 쌍의 "올바른 값"을 나타냅니다. 이 함수는 일반적으로 분석 노드에서 사용합니다.
@PREDICTED	모두	CLEM 표현식을 사용자 정의 분석 함수에서 사용하는 경우 @PREDICTED는 분석 중인 목표/예측 쌍의 예측값을 나타냅니다. 이 함수는 일반적으로 분석 노드에서 사용합니다.
@PARTITION_FIELD	모두	현재 파티션 필드의 이름을 대체합니다.
@TRAINING_PARTITION	모두	현재 훈련 파티션 값을 리턴합니다. 예를 들어, 선택 노드를 사용하여 훈련 레코드를 선택하려면 CLEM 표현식(@PARTITION_FIELD = @TRAINING_PARTITION)을 사용하십시오. 이렇게 하면 데이터의 각 파티션을 나타내는 데 사용하는 값에 관계없이 선택 노드가 항상 작동합니다.
@TESTING_PARTITION	모두	현재 검정 파티션 값을 리턴합니다.
@VALIDATION_PARTITION	모두	현재 검증 파티션 값을 리턴합니다.
@FIELDS_BETWEEN(start, end)	모두	데이터 필드의 자연(즉, 삽입) 순서를 기반으로 지정된 시작 필드와 종료 필드(포함) 사이의 필드 이름 목록을 리턴합니다.
@FIELDS_MATCHING(pattern)	모두	지정된 패턴과 매치하는 필드 이름 목록을 리턴합니다. 물음표(?)를 패턴에 포함시켜 하나의 문자와 정확하게 매치할 수 있습니다. 별표(*)는 0개 이상의 문자와 매치됩니다. 리터럴 물음표나 별표와 매치하려면 (이를 와일드카드로 사용하지 않음) 백슬래시(\)를 이스케이프 문자로 사용할 수 있습니다. 참고: 이 경우 문자열 리터럴이 인수로 필요합니다(중첩 표현식을 사용하여 인수를 생성할 수 없음).
@MULTI_RESPONSE_SET	모두	이름 지정된 다중 응답 세트의 필드 목록을 리턴합니다.

제 11 장 리포지토리에서 IBM SPSS Modeler 사용

IBM SPSS Collaboration and Deployment Services Repository 정보

SPSS Modeler는 IBM SPSS Collaboration and Deployment Services 리포지토리와 함께 사용될 수 있으므로, 사용자는 데이터 마이닝 모델 및 관련 예측 개체의 라이프 사이클을 관리하고 이러한 개체가 엔터프라이즈 애플리케이션, 도구 및 솔루션에 의해 사용되도록 할 수 있습니다. 이러한 방식으로 공유할 수 있는 IBM SPSS Modeler 개체에는 스트림, 노드, 스트림 결과, 프로젝트 및 모델이 포함됩니다. 개체가 중앙 리포지토리에 저장되고, 이것은 여기서 다른 애플리케이션과 공유할 수 있으며 확장된 버저닝, 메타데이터 및 검색 기능을 사용하여 추적될 수 있습니다.

리포지토리와 함께 SPSS Modeler를 사용하기 전에, 리포지토리 호스트에 어댑터를 설치해야 합니다. 이 어댑터가 없으면 특정 SPSS Modeler 노드나 모델에서 리포지토리 개체에 액세스를 시도할 때 다음과 같은 메시지가 표시될 수 있습니다.

새 노드, 모델 및 출력 유형을 지원하도록 리포지토리를 업데이트해야 할 수 있습니다.

어댑터 설치 지시사항은 제품 다운로드에서 PDF 파일로 제공되는 *SPSS Modeler 배포 설치 안내서*를 참조하십시오. IBM SPSS Deployment Manager에서의 IBM SPSS Modeler 리포지토리 오브젝트 액세스 방법에 대한 세부사항은 *SPSS Modeler 배포 안내서*를 참조하십시오.

다음 절에서는 SPSS Modeler 내에서의 리포지토리 액세스에 대한 정보를 제공합니다.

광범위한 버저닝 및 검색 지원

리포지토리는 포괄적인 개체 버저닝 및 검색 기능을 제공합니다. 예를 들어, 스트림을 작성하여 리포지토리에 저장한다고 가정합시다. 여기서 스트림은 다른 부서의 연구원들과 공유될 수 있습니다. 나중에 스트림을 SPSS Modeler에서 업데이트하면, 이전 버전을 덮어쓰지 않고도 업데이트된 버전을 리포지토리에 추가할 수 있습니다. 모든 버전은 액세스 가능한 상태로 남아 있으며 이름, 레이블, 사용된 필드 또는 다른 속성별로 검색될 수 있습니다. 예를 들면, 순수익을 입력으로 사용하는 모든 모델 버전 또는 특정 작성자가 작성한 모든 모델을 검색할 수 있습니다. (기존의 파일 시스템으로 이를 수행하려면 각 버전을 서로 다른 파일 이름으로 저장해야 하고 버전 간 관계가 소프트웨어에 알려지지 않았습니다.)

싱글 사인온

싱글 사인온 기능을 통해 사용자는 매번 사용자 이름 및 비밀번호 세부사항을 입력하지 않고도 리포지토리에 연결할 수 있습니다. 사용자의 기존 로컬 네트워크 로그인 세부사항은 필수 인증을 IBM SPSS Collaboration and Deployment Services에 제공합니다. 이 기능은 다음 사항에 따라 달라집니다.

- IBM SPSS Collaboration and Deployment Services는 싱글 사인온 제공자를 사용하도록 구성되어야 합니다.
- 사용자는 제공자와 호환 가능한 호스트에 로그인되어야 합니다.

추가 정보는 176 페이지의 [『리포지토리에 연결』](#)의 내용을 참조하십시오.

리포지토리 오브젝트 저장 및 배포

IBM SPSS Modeler에서 작성된 스트림을 확장자가 .str인 파일로 있는 그대로 리포지토리에 저장할 수 있습니다. 이런 방식으로 엔터프라이즈 전반에서 여러 사용자가 단일 스트림에 액세스할 수 있습니다. 자세한 정보는 177 페이지의 [『리포지토리에 오브젝트 저장』](#)의 내용을 참조하십시오.

리포지토리에 스트림을 배포하는 것도 가능합니다. 배포된 스트림은 추가 메타데이터를 가진 파일로 저장됩니다. 배포된 스트림은 IBM SPSS Collaboration and Deployment Services의 엔터프라이즈 레벨 기능(예: 자동화된 스코어링 및 모델 새로 고침)을 최대한 활용할 수 있습니다. 예를 들어, 새 데이터를 사용할 수 있게 되면 정규 스케줄 간격으로 모델을 자동으로 업데이트할 수 있습니다. 또는 Champion Challenger 분석을 위해 스트림 세트를 배포할 수 있는데, 스트림을 비교하여 가장 효과적인 예측 모델이 포함된 스트림을 판별합니다.

참고: 연관 규칙, STP 및 TCM 모델링 노드는 IBM SPSS Collaboration and Deployment Services에서 모델 평가 또는 Champion Challenger 단계를 지원하지 않습니다.

스트림(.str 확장자)을 배포할 수 있습니다. 스트림으로 배포하면 씬 클라이언트 애플리케이션 IBM SPSS Modeler Advantage가 스트림을 사용할 수 있습니다. 자세한 정보는 [193 페이지의 『스트림을 IBM SPSS Modeler Advantage에서 열기』](#)의 내용을 참조하십시오.

추가 정보는 [186 페이지의 『스트림 배포 옵션』](#)의 내용을 참조하십시오.

기타 배포 옵션

IBM SPSS Collaboration and Deployment Services가 엔터프라이즈 컨텐츠 관리를 위한 가장 광범위한 기능을 제공하긴 하지만 다음을 포함하여 스트림 배포 또는 내보내기를 위한 다수의 기타 메커니즘도 사용할 수 있습니다.

- IBM SPSS Modeler Solution Publisher Runtime에서 나중에 사용할 수 있도록 스트림 및 모델 내보내기
- 모델 정보 인코딩을 위한 XML 기반 형식인 PMML로 하나 이상의 모델 내보내기. 자세한 정보는 [193 페이지의 『모델을 PMML로 가져오기 및 내보내기』](#)의 내용을 참조하십시오.

리포지토리에 연결

1. 리포지토리에 연결하려면 IBM SPSS Modeler 기본 메뉴에서 다음을 클릭하십시오.

도구 > 리포지토리 > 옵션...

2. **RepositoryURL**. 필드에서 액세스하려는 리포지토리 설치의 디렉토리 경로 또는 URL을 입력하거나 선택하십시오. 한 번에 단 하나의 리포지토리에만 연결할 수 있습니다.

설정은 각 사이트 또는 설치에 특정적입니다. 특정 로그인 세부사항은 로컬 시스템 관리자에게 문의하십시오.

자격 설정. 로컬 컴퓨터 사용자 이름과 비밀번호 세부사항을 사용하여 로그인하려고 시도하는 싱글 사인온 기능을 사용하려면 이 상자를 선택 해제 상태로 두십시오. 싱글 사인온이 가능하지 않거나 싱글 사인온을 사용 안함으로 설정하기 위해 이 상자를 선택한 경우에는(예: 관리자 계정에 로그인하는 경우), 신임 정보를 입력하라는 화면이 표시됩니다.

리포지토리의 신임 정보 입력

사용자의 설정에 따라서 리포지토리: 신임 정보 대화 상자에 다음 필드가 필수일 수 있습니다.

사용자 ID와 비밀번호. 로그온할 유효한 사용자 이름과 비밀번호를 지정하십시오. 필요한 경우 자세한 정보는 로컬 관리자에게 문의하십시오.

제공자. 인증을 위해 보안 제공자를 선택하십시오. 리포지토리는 다른 보안 제공자를 사용하도록 구성될 수 있습니다. 필요한 경우 자세한 정보는 로컬 관리자에게 문의하십시오.

리포지토리 및 사용자 ID 기억. 연결하려고 할 때마다 기억할 필요가 없도록 현재 설정을 기본값으로 저장하십시오.

리포지토리 신임 정보 찾기

Analytic Server, Cognos, ODBC 또는 TM1 소스 노드에서 리포지토리에 연결할 때는 이전에 레코드된 신임 정보를 선택하여 리포지토리에 연결할 수 있습니다. 이러한 신임 정보는 리포지토리 신임 정보 대화 상자에 나열되어 있습니다. 이 대화 상자를 선택하려면 신임 정보 필드 옆의 찾아보기를 클릭하십시오.

리포지토리 신임 정보 선택 대화 상자에서 제공된 목록에 있는 신임 정보를 강조 표시하고 확인을 클릭하십시오. 목록이 너무 크면 필터 필드를 사용하여 이름 또는 이름의 일부를 입력하여 필요한 신임 정보를 찾으십시오.

리포지토리 내용 찾아보기

리포지토리를 사용하면 Windows 탐색기와 유사한 방법으로 저장된 내용을 찾아볼 수 있습니다. 저장된 각 오브젝트의 버전을 찾아볼 수도 있습니다.

1. IBM SPSS Collaboration and Deployment Services Repository 창을 열려면 SPSS Modeler 메뉴에서 다음을 클릭하십시오.

도구 > 리포지토리 > 탐색...

1. 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 [176 페이지의 『리포지토리에 연결』](#)의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.

탐색기 창은 처음에는 폴더 계층의 트리 보기를 표시합니다. 내용을 표시하려는 폴더 이름을 클릭하십시오.

현재 선택 또는 검색 기준과 매치하는 오브젝트가 오른쪽 분할창에 나열되어 있고 선택된 버전에 대한 자세한 정보는 오른쪽 하단 분할창에 표시되어 있습니다. 표시된 속성은 가장 최신 버전에 적용됩니다.

리포지토리에 오브젝트 저장

리포지토리에 다른 사용자와 애플리케이션이 액세스할 수 있는 위치에 스트림, 노드, 모델, 모델 팔레트, 프로젝트 및 출력 오브젝트를 저장할 수 있습니다.

다른 사용자가 IBM SPSS Collaboration and Deployment Services Deployment Portal을 사용하여 인터넷을 통해 볼 수 있는 형식으로 스트림 출력을 리포지토리에 공개할 수도 있습니다.

오브젝트 특성 설정

오브젝트를 저장할 때, 리포지토리: 저장 대화 상자가 표시되어 오브젝트의 여러 특성 값을 설정할 수 있습니다. 다음을 수행할 수 있습니다.

- 오브젝트가 저장될 이름과 리포지토리 폴더를 선택
- 버전 레이블과 기타 검색 가능한 특성과 같은 오브젝트에 대한 정보 추가
- 하나 이상의 분류 주제를 오브젝트에 지정
- 오브젝트의 보안 옵션 설정

다음 섹션은 설정할 수 있는 특성을 설명합니다.

오브젝트 저장을 위한 위치 선택

리포지토리: 저장 대화 상자에 다음을 입력하십시오.

저장 위치. 오브젝트가 저장될 현재 폴더--위치를 표시합니다. 목록에서 폴더 이름을 두 번 클릭하여 폴더를 현재 폴더로 설정하십시오. 폴더 위로 이동 단추를 사용하여 상위 폴더로 이동하십시오. 새 폴더 단추를 사용하여 현재 수준에 폴더를 작성하십시오.

파일 이름. 오브젝트가 저장될 이름입니다.

저장. 오브젝트를 현재 위치에 저장합니다.

저장된 오브젝트에 대한 정보 추가

리포지토리: 저장 대화 상자의 정보 탭에 있는 모든 필드는 선택사항입니다.

작성자. 리포지토리에서 오브젝트를 작성하는 사용자의 사용자 이름. 기본적으로 이는 리포지토리 연결에 사용된 사용자 이름을 표시하지만 여기에서 이 이름을 변경할 수 있습니다.

버전 레이블. 목록에서 레이블을 선택하여 오브젝트 레이블을 표시하거나 추가를 클릭하여 새 레이블을 작성하십시오. 레이블에서 "[" 문자를 피하십시오. 이 오브젝트 버전에 레이블을 지정하지 않으려는 경우에는 어떤 상자도 선택되어 있지 않은지 확인하십시오. 자세한 정보는 [184 페이지의 『오브젝트 특성 보기 및 편집』](#)의 내용을 참조하십시오.

설명 오브젝트의 설명. 사용자는 설명으로 오브젝트를 검색할 수 있습니다(참고 참조).

키워드. 오브젝트와 관련된 하나 이상의 키워드이며 검색 용도로 사용할 수 있습니다(참고 참조).

만료. 오브젝트가 더 이상 일반 사용자에게 표시되지 않는 날짜입니다. 그렇지만 소유자나 리포지토리 관리자에게는 여전히 표시됩니다. 만료 날짜를 설정하려면 날짜 옵션을 선택하고 날짜를 입력하거나 달력 단추를 사용하여 이를 선택하십시오.

저장. 오브젝트를 현재 위치에 저장합니다.

참고: 설명 및 키워드 필드는 SPSS Modeler에서 오브젝트의 주석(Annotation) 탭에 입력한 것과는 별개로 처리됩니다. 설명이나 키워드로 리포지토리를 검색하면 주석(Annotation) 탭에서 정보를 리턴하지 않습니다. 자세한 정보는 [181 페이지의 『리포지토리에서 오브젝트 검색』](#)의 내용을 참조하십시오.

저장된 오브젝트에 주제 지정

주제는 리포지토리에 저장된 내용의 계층 구조 분류 시스템입니다. 오브젝트를 저장할 때 사용 가능한 주제에서 선택할 수 있고 사용자는 주제별로 오브젝트를 검색할 수도 있습니다. 사용 가능한 주제 목록은 적합한 권한이 있는 리포지토리 사용자에 의해 설정됩니다. 자세한 정보는 배포 관리자 사용자 안내서를 참조하십시오.

리포지토리: 저장 대화 상자의 주제 탭에서 오브젝트에 주제 지정:

1. 추가 단추를 클릭하십시오.
2. 사용 가능한 주제 목록에서 주제 이름을 클릭하십시오.
3. 확인을 클릭하십시오.

주제 지정 제거:

4. 지정된 주제 목록에서 주제를 선택하십시오.
5. 삭제를 클릭하십시오.

저장된 오브젝트에 보안 옵션 설정

리포지토리: 저장 대화 상자의 보안 탭에서 저장된 오브젝트에 여러 보안 옵션을 설정하거나 변경할 수 있습니다. 하나 이상의 주성분(즉, 사용자 또는 사용자 그룹)에 대해 다음을 수행할 수 있습니다.

- 오브젝트에 액세스 권한 지정
- 오브젝트에 대한 액세스 권한 수정
- 오브젝트에 대한 액세스 권한 제거

주성분. 이 오브젝트에 대한 액세스 권한을 가지고 있는 사용자 또는 그룹의 리포지토리 사용자 이름.

권한. 이 사용자 또는 그룹이 오브젝트에 대해 가지고 있는 액세스 권한.

추가. 하나 이상의 사용자나 그룹을 이 오브젝트에 대한 액세스 권한이 있는 목록에 추가할 수 있습니다. 자세한 정보는 [178 페이지의 『사용자를 권한 목록에 추가』](#)의 내용을 참조하십시오.

수정. 이 오브젝트에 대한 선택한 사용자나 그룹의 액세스 권한을 수정할 수 있습니다. 읽기 액세스는 기본적으로 부여됩니다. 이 옵션을 사용하면 추가 액세스 권한, 즉 소유자, 쓰기, 삭제 및 수정 권한을 부여할 수 있습니다.

삭제. 이 오브젝트의 권한 목록에서 선택된 사용자 또는 그룹을 삭제합니다.

사용자를 권한 목록에 추가

다음 필드는 리포지토리: 저장 대화 상자의 보안 탭에서 추가를 선택하면 사용 가능합니다.

제공자 선택. 인증을 위해 보안 제공자를 선택하십시오. 리포지토리는 다른 보안 제공자를 사용하도록 구성될 수 있습니다. 필요한 경우 자세한 정보는 로컬 관리자에게 문의하십시오.

찾기. 추가할 사용자나 그룹의 리포지토리 사용자 이름을 입력하고 검색을 클릭하여 사용자 목록에 해당 이름을 표시하십시오. 둘 이상의 사용자 이름을 동시에 추가하려면 이 필드를 공백으로 두고 검색을 클릭하여 모든 리포지토리 사용자 이름의 목록을 표시하십시오.

사용자 목록. 목록에서 하나 이상의 사용자 이름을 선택하고 확인을 클릭하여 이를 권한 목록에 추가하십시오.

오브젝트의 액세스 권한 수정

다음 필드는 리포지토리: 저장 대화 상자의 보안 탭에서 수정을 선택하면 사용 가능합니다.

소유자. 이 옵션을 선택하여 이 사용자나 그룹 소유자에게 오브젝트에 대한 액세스 권한을 제공하십시오. 소유자는 삭제 및 수정 액세스 권한을 포함하여 오브젝트에 대한 완전한 제어를 가지고 있습니다.

읽기. 기본적으로 오브젝트 소유자가 아닌 사용자나 그룹에는 오브젝트에 대한 읽기 액세스 권한만 가지고 있습니다. 적합한 확인 상자를 선택하여 이 사용자나 그룹에 쓰기, 삭제 및 수정 권한 액세스 권한을 추가하십시오.

스트림 저장

스트림을 다른 사용자가 액세스할 수 있는 리포지토리에 *.str* 파일로서 저장할 수 있습니다.

참고: 추가 리포지토리 기능을 이용하기 위해 스트림 배포에 대한 정보는 [185 페이지의 『스트림 배포』](#)의 내용을 참조하십시오.

현재 스트림 저장:

1. 기본 메뉴에서 다음을 클릭하십시오.

파일 > 저장 > 스트림으로 저장...

2. 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 [176 페이지의 『리포지토리에 연결』](#)의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.

3. 리포지토리: 저장 대화 상자에서 오브젝트를 저장하려는 폴더를 선택하고, 기록할 다른 정보를 지정하고, 저장 단추를 클릭하십시오. 자세한 정보는 [177 페이지의 『오브젝트 특성 설정』](#)의 내용을 참조하십시오.

프로젝트 저장

전체 IBM SPSS Modeler 프로젝트를 다른 사용자가 액세스할 수 있도록 리포지토리에 *.cpj* 파일로서 저장할 수 있습니다.

프로젝트 파일은 다른 IBM SPSS Modeler 오브젝트의 컨테이너이므로 IBM SPSS Modeler에 프로젝트의 오브젝트를 리포지토리에 저장하도록 알려야 합니다. 프로젝트 특성 대화 상자에서 설정을 사용하여 이를 수행할 수 있습니다. 자세한 정보는 [199 페이지의 『프로젝트 특성 설정』](#)의 내용을 참조하십시오.

오브젝트를 리포지토리에 저장하도록 프로젝트를 구성한 후에는 프로젝트에 새 오브젝트를 추가할 때마다 IBM SPSS Modeler는 자동으로 오브젝트를 저장하라는 프롬프트를 표시합니다.

IBM SPSS Modeler 세션을 완료한 후에는 추가사항을 기억할 수 있도록 프로젝트 파일의 새 버전을 저장해야 합니다. 프로젝트 파일은 자동으로 해당 오브젝트의 최신 버전을 포함 및 검색합니다. IBM SPSS Modeler 세션 중에 오브젝트를 프로젝트에 추가하지 않은 경우에는 프로젝트 파일을 다시 저장할 필요가 없습니다. 그러나 변경한 프로젝트 오브젝트(스트림, 출력 등)의 새 버전을 저장해야 합니다.

프로젝트 저장

1. IBM SPSS Modeler의 관리자 분할창에서 CRISP-DM 또는 클래스 탭에서 프로젝트를 선택하고 기본 메뉴에서 다음을 클릭하십시오.

파일 > 프로젝트 > 프로젝트 저장...

2. 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 [176 페이지의 『리포지토리에 연결』](#)의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.

3. 리포지토리: 저장 대화 상자에서 오브젝트를 저장하려는 폴더를 선택하고, 기록할 다른 정보를 지정하고, 저장 단추를 클릭하십시오. 자세한 정보는 [177 페이지의 『오브젝트 특성 설정』](#)의 내용을 참조하십시오.

노드 저장

현재 스트림의 개별 노드 정의를 다른 사용자가 액세스할 수 있는 리포지토리에 *.nod* 파일로서 저장할 수 있습니다.

노드 저장:

1. 스트림 캔버스에서 노드를 마우스 오른쪽 단추로 클릭하고 노드 저장을 클릭하십시오.

- 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 [176 페이지의 『리포지토리에 연결』](#)의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.
- 리포지토리: 저장 대화 상자에서 오브젝트를 저장하려는 폴더를 선택하고, 기록할 다른 정보를 지정하고, 저장 단추를 클릭하십시오. 자세한 정보는 [177 페이지의 『오브젝트 특성 설정』](#)의 내용을 참조하십시오.

출력 오브젝트 저장

현재 스트림의 출력 오브젝트를 다른 사용자가 액세스할 수 있는 리포지토리에 .cou 파일로서 저장할 수 있습니다.

출력 오브젝트 저장:

- SPSS Modeler에서 관리자 분할창의 출력 탭에서 오브젝트를 클릭하고 기본 메뉴에서 다음을 클릭하십시오.
파일 > 출력 > 출력 저장...
- 또는 출력 탭에서 오브젝트를 마우스 오른쪽 단추로 클릭하고 **저장**을 클릭하십시오.
- 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 [176 페이지의 『리포지토리에 연결』](#)의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.
- 리포지토리: 저장 대화 상자에서 오브젝트를 저장하려는 폴더를 선택하고, 기록할 다른 정보를 지정하고, 저장 단추를 클릭하십시오. 자세한 정보는 [177 페이지의 『오브젝트 특성 설정』](#)의 내용을 참조하십시오.

모델 및 모델 팔레트 저장

개별 모델을 다른 사용자가 액세스할 수 있는 리포지토리에 .gm 파일로서 저장할 수 있습니다. 리포지토리에 모델 팔레트의 전체 내용을 .gen 파일로서 저장할 수도 있습니다.

모델 저장:

- SPSS Modeler의 모델 팔레트에서 오브젝트를 클릭하고 기본 메뉴에서 다음을 클릭하십시오.
파일 > 모델 > 모델 저장...
- 또는 모델 팔레트에서 오브젝트를 마우스 오른쪽 단추로 클릭하고 **모델 저장**을 클릭하십시오.
- "저장 공간 프로시저 완료"에서 계속하십시오.

모델 팔레트 저장:

- 모델 팔레트의 배경을 마우스 오른쪽 단추로 클릭하십시오.
- 팝업 메뉴에서 **팔레트 저장**을 클릭하십시오.
- "저장 공간 프로시저 완료"에서 계속하십시오.

저장 공간 프로시저 완료:

- 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 [176 페이지의 『리포지토리에 연결』](#)의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.
- 리포지토리: 저장 대화 상자에서 오브젝트를 저장하려는 폴더를 선택하고, 기록할 다른 정보를 지정하고, 저장 단추를 클릭하십시오. 자세한 정보는 [177 페이지의 『오브젝트 특성 설정』](#)의 내용을 참조하십시오.

리포지토리에서 오브젝트 검색

리포지토리에 저장된 스트림, 모델, 모델 팔레트, 노드, 프로젝트 및 출력 오브젝트를 검색할 수 있습니다.

참고: 여기에 설명된 메뉴 옵션을 사용하는 것 외에도 SPSS Modeler 창의 오른쪽 상단에서 관리자 분할창의 적합한 탭을 마우스 오른쪽 단추로 클릭하여 스트림, 출력 오브젝트, 모델 및 모델 팔레트를 검색할 수도 있습니다.

- 스트림을 검색하려면 IBM SPSS Modeler 기본 메뉴에서 다음을 클릭하십시오.

파일 > 스트림 검색...

2. IBM SPSS Modeler 기본 메뉴에서 모델, 모델 팔레트, 프로젝트 또는 출력 오브젝트를 검색하려면 다음을 클릭하십시오.

파일 > 모델 > 모델 검색...

or

파일 > 모델 > 모델 팔레트 검색...

or

파일 > 프로젝트 > 프로젝트 검색...

or

파일 > 출력 > 출력 검색...

3. 또는 관리자나 프로젝트 분할창을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 검색을 클릭하십시오.

4. 노드를 검색하려면 IBM SPSS Modeler 기본 메뉴에서 다음을 클릭하십시오.

삽입 > 리포지토리의 노드 또는 수퍼노드...

a. 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 [176 페이지의 『리포지토리에 연결』](#)의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.

5. 리포지토리: 검색 대화 상자에서 오브젝트로 이동하고 이를 선택하고 검색 단추를 클릭하십시오. 자세한 정보는 [의 내용을 참조하십시오.](#)

검색할 오브젝트 선택

다음 필드는 리포지토리: 검색/검색 대화 상자에서 사용할 수 있습니다.

찾을 위치. 현재 폴더의 폴더 계층 구조를 보여줍니다. 다른 폴더로 이동하려면 이 목록에서 하나를 선택하여 직접 이동하거나 이 필드 아래에 있는 오브젝트 목록을 사용하여 이동하십시오.

폴더 위로 단추. 계층 구조에서 현재 폴더의 한 수준 위로 이동합니다.

새 폴더 단추. 계층 구조에서 현재 수준에 새 폴더를 작성합니다.

파일 이름. 선택한 오브젝트의 리포지토리 파일 이름입니다. 해당 오브젝트를 검색하려면 **검색**을 클릭하십시오.

파일 유형. 검색하려고 선택한 오브젝트의 유형입니다. 이 유형의 오브젝트만이 폴더와 함께 오브젝트 목록에 표시됩니다. 검색할 다른 유형의 오브젝트를 표시하려면 목록에서 오브젝트 유형을 선택하십시오.

잠긴 상태로 열기. 기본적으로 오브젝트가 검색되면 다른 사용자가 이를 업데이트할 수 없도록 리포지토리에 잠겨 있습니다. 검색 시에 오브젝트가 잠겨 있는 것을 원하지 않으면 이 상자를 선택 취소하십시오.

설명, 키워드. 오브젝트가 저장될 때 오브젝트에 대한 추가 세부사항이 정의된 경우에는 이러한 세부사항이 여기에 표시됩니다. 자세한 정보는 [177 페이지의 『저장된 오브젝트에 대한 정보 추가』](#)의 내용을 참조하십시오.

버전. 오브젝트의 최신이 아닌 버전을 검색하려면 이 단추를 클릭하십시오. 모든 버전의 자세한 정보가 표시되므로 원하는 버전을 선택할 수 있습니다.

오브젝트 버전 선택하기

리포지토리 오브젝트의 특정 버전을 선택하려면 리포지토리: 버전 선택 대화 상자에서 다음을 수행하십시오.

- (선택적) 해당 열의 헤더를 두 번 클릭하여 목록을 버전, 레이블, 크기, 작성 날짜 또는 작성 사용자별로 정렬합니다.
- 작업하려는 오브젝트 버전을 선택하십시오.
- 계속을 클릭하십시오.

리포지토리에서 오브젝트 검색

이름, 폴더, 유형, 레이블, 날짜 또는 기타 기준별로 오브젝트를 검색할 수 있습니다.

오브젝트를 이름으로 검색

1. IBM SPSS Modeler 기본 메뉴에서 다음을 클릭하십시오.

도구 > 리포지토리 > 탐색...

a. 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 [176 페이지의 『리포지토리에 연결』](#)의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.

2. 검색 탭을 클릭하십시오.

3. 이름 지정된 오브젝트 검색 필드에서 찾으려는 오브젝트 이름을 지정하십시오.

이름으로 오브젝트를 검색할 때 별표(*)를 임의의 문자열과 매치시키기 위해 와일드카드 문자를 사용할 수 있고 물음표(?)는 단일 문자와 매치합니다. 예를 들어, *cluster*는 이름 어딘가에 cluster 문자열을 포함하는 모든 오브젝트와 매치합니다. m0?_* 검색 문자열은 M01_cluster.str 및 M02_cluster.str과 매치하지만 M01a_cluster.str과는 매치하지 않습니다. 검색은 대소문자를 구분하지 않습니다(cluster는 Cluster 및 CLUSTER와 매치함).

참고: 오브젝트 수가 많으면 검색에 약간의 시간이 소요될 수 있습니다.

기타 기준으로 검색

제목, 레이블, 날짜, 작성자, 키워드, 색인 컨텐츠 또는 설명을 기반으로 검색을 수행할 수 있습니다. 모든 지정된 검색 기준과 매치하는 오브젝트만이 발견됩니다. 예를 들어, 특정 레이블이 적용되어 있고 특정 날짜 이후에 수정된 하나 이상의 군집 모델을 포함하는 모든 스트림을 찾을 수 있습니다.

오브젝트 유형 검색을 모델, 스트림, 출력, 노드, 수퍼노드, 프로젝트, 모델 팔레트 또는 기타 오브젝트 유형으로 제한할 수 있습니다.

- 모델 범주별로 모델을 검색하거나(분류, 근사, 군집 등). 코호넨 등과 같이 특정 모델 알고리즘으로 검색할 수 있습니다.

사용한 필드로 검색할 수도 있습니다. 예를 들어, 이름이 *income*인 필드를 입력이나 출력(목표) 필드로 사용하는 모든 모델을 검색할 수 있습니다.

- 스트림 스트림의 경우 사용한 필드 또는 스트림에 포함된 모델 유형(범주 또는 알고리즘)으로 검색을 제한할 수 있습니다.

주제 적합한 권한이 있는 리포지토리 사용자가 설정한 목록에서 특정 주제와 연관된 모델을 검색할 수 있습니다. 자세한 정보는 배포 관리자 사용자 안내서를 참조하십시오. 목록을 획득하려면 이 상자를 선택한 다음 표시되는 주제 추가 단추를 클릭하고, 목록에서 하나 이상의 주제를 선택하고 확인을 클릭하십시오.

레이블 검색을 특정 오브젝트 베전 레이블로 제한합니다.

날짜 작성 또는 수정 날짜를 지정하고 지정된 날짜 범위 이전, 이후 또는 사이에서 오브젝트에 대해 검색할 수 있습니다.

작성자 검색을 특정 사용자에 의해 작성된 오브젝트로 제한합니다.

키워드 특정 키워드에 대해 검색합니다. IBM SPSS Modeler에서, 키워드는 스트림, 모델 또는 출력 오브젝트에 대한 주석(Annotation) 탭에 지정됩니다.

설명 설명 필드에서 특정 용어를 검색합니다. IBM SPSS Modeler에서 설명은 스트림, 모델 또는 출력 오브젝트에 대한 주석(Annotation) 탭에 지정되어 있습니다. 다중 검색 절은 세미콜론으로 구분할 수 있습니다(예: *income; crop type; claim value*). (검색 절 내에서는 공백이 중요함을 유의하십시오. 예를 들어, 하나의 공백이 있는 *crop type*과 두 개의 공백이 있는 *crop type*은 서로 다릅니다.)

리포지토리 오브젝트 수정

SPSS Modeler에서 직접 리포지토리의 기존 오브젝트를 수정할 수 있습니다. 다음을 수행할 수 있습니다.

- 폴더 작성, 이름 변경 또는 삭제
- 오브젝트 잠금 또는 잠금 해제

- 오브젝트 삭제

폴더 생성, 이름 변경 및 삭제

1. 리포지토리에서 폴더에 대해 작업을 수행하려면 SPSS Modeler 기본 메뉴에서 다음을 클릭하십시오.

도구 > 리포지토리 > 탐색...

a. 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 176 페이지의 『리포지토리에 연결』의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.

2. 폴더 탭이 활성인지 확인하십시오.

3. 새 폴더를 작성하려면 상위 폴더를 마우스 오른쪽 단추로 클릭하고 새 폴더를 클릭하십시오.

4. 폴더의 이름을 변경하려면 이를 마우스 오른쪽 단추로 클릭하고 폴더 이름 변경을 클릭하십시오.

5. 폴더를 삭제하려면 마우스 오른쪽 단추로 클릭하고 폴더 삭제를 클릭하십시오.

리포지토리 오브젝트 잠금 및 잠금 해제

다른 사용자가 기존 버전을 업데이트하거나 새 버전을 작성하지 못하도록 오브젝트를 잠글 수 있습니다. 잠긴 오브젝트는 오브젝트 아이콘 위에 자물쇠 기호로 표시됩니다.



그림 17. 잠긴 오브젝트

오브젝트 잠그기

1. 리포지토리 탐색기 창에서 필요한 오브젝트를 마우스 오른쪽 단추로 클릭하십시오.

2. 잠금을 클릭하십시오.

오브젝트 잠금 해제

1. 리포지토리 탐색기 창에서 필요한 오브젝트를 마우스 오른쪽 단추로 클릭하십시오.

2. 잠금 해제를 클릭하십시오.

리포지토리 오브젝트 삭제

리포지토리에서 오브젝트를 삭제하기 전에 오브젝트의 모든 버전을 삭제하려는지 또는 특정 버전만을 삭제하려는지를 결정해야 합니다.

오브젝트의 모든 버전 삭제

1. 리포지토리 탐색기 창에서 필요한 오브젝트를 마우스 오른쪽 단추로 클릭하십시오.

2. 오브젝트 삭제를 클릭하십시오.

오브젝트의 가장 최신 버전 삭제

1. 리포지토리 탐색기 창에서 필요한 오브젝트를 마우스 오른쪽 단추로 클릭하십시오.

2. 삭제를 클릭하십시오.

오브젝트의 이전 버전 삭제

1. 리포지토리 탐색기 창에서 필요한 오브젝트를 마우스 오른쪽 단추로 클릭하십시오.

2. 버전 삭제를 클릭하십시오.

3. 삭제할 버전을 선택하고 확인을 클릭하십시오.

리포지토리 오브젝트의 특성 관리

SPSS Modeler에서 다양한 오브젝트 특성을 제어할 수 있습니다. 다음을 수행할 수 있습니다.

- 폴더의 특성 보기

- 오브젝트의 특성 보고 편집하기
- 오브젝트의 버전 레이블을 작성, 적용 및 삭제

폴더 특성 보기

리포지토리 창에서 폴더의 특성을 보려면 필수 폴더를 마우스 오른쪽 단추로 클릭하십시오. 폴더 특성을 클릭하십시오.

일반 탭

이 탭은 폴더 이름, 작성 및 수정 날짜를 표시합니다.

권한 탭

이 탭에서 폴더의 읽기 및 쓰기 권한을 지정합니다. 상위 폴더에 대한 액세스가 있는 모든 사용자와 그룹이 나열됩니다. 권한은 계층 구조를 따릅니다. 예를 들어, 읽기 권한이 없는 경우에는 쓰기 권한을 가질 수 없습니다. 쓰기 권한이 없는 경우에는 삭제 권한을 가질 수 없습니다.

사용자 및 그룹 이 폴더에 대해 읽기 액세스 이상이 있는 리포지토리 사용자 및 그룹을 나열합니다. 쓰기 및 삭제 확인 상자를 선택하여 이 폴더의 액세스 권한을 특정 사용자나 그룹에 추가하십시오. 권한 탭의 오른쪽에 있는 사용자/그룹 추가 아이콘을 클릭하여 추가 사용자 및 그룹에 액세스를 지정하십시오. 사용 가능한 사용자와 그룹 목록은 관리자에 의해 제어됩니다.

계단식 권한. 현재 폴더에 작성된 변경사항이 해당 하위 폴더(있는 경우)에 적용되는 방법을 제어하는 옵션을 선택하십시오.

- **모든 권한 계단식 설정**. 현재 폴더에서 모든 하위 및 자손 폴더까지 권한 설정을 계단식으로 설정합니다. 이는 여러 폴더에 동시에 권한을 설정하는 빠른 방법입니다. 상위 폴더에 필요한 권한을 설정한 다음에 필요에 따라 계단식으로 설정하십시오.
- **변경사항만 계단식 설정**. 마지막 변경이 적용된 이후에 작성된 변경사항만을 계단식으로 설정합니다. 예를 들어, 새 그룹이 추가되었고 여기에 판매 분기 아래에 있는 모든 폴더에 대한 액세스를 제공하려면 루트 판매 폴더에 그룹 액세스를 제공하고 변경사항을 모든 하위 폴더에 계단식으로 설정하면 됩니다. 기존 하위 폴더에 대한 다른 모든 권한은 이전과 같이 남아 있습니다.
- **계단식 설정 안함**. 모든 변경사항은 현재 폴더에만 적용되고 하위 폴더에 계단식으로 적용하지 않습니다.

오브젝트 특성 보기 및 편집

오브젝트 특성 대화 상자에서 특성을 보고 편집할 수 있습니다. 몇몇 특성은 변경할 수 없지만 항상 새 버전을 추가하여 오브젝트를 업데이트할 수 있습니다.

1. 리포지토리 창에서 필수 오브젝트를 마우스 오른쪽 단추로 클릭하십시오.
2. 오브젝트 특성을 클릭하십시오.

일반 탭

이름. 리포지토리에 표시되는 오브젝트 이름.

작성 날짜. 오브젝트(버전이 아님)가 작성된 날짜입니다.

마지막 수정 날짜. 가장 최신 버전이 수정된 날짜입니다.

작성자. 사용자의 로그인 이름입니다.

설명. 기본적으로 여기에는 SPSS Modeler에서 오브젝트의 주석(Annotation) 탭에 지정된 설명이 포함됩니다.

링크된 주제. 리포지토리는 필요한 경우 모델과 관련 오브젝트가 주제별로 구성될 수 있게 해줍니다. 사용 가능한 주제 목록은 적합한 권한이 있는 리포지토리 사용자에 의해 설정됩니다. 자세한 정보는 배포 관리자 사용자 안내서를 참조하십시오.

키워드. 스트림, 모델 또는 출력 오브젝트의 주석(Annotation) 탭에 키워드를 지정합니다. 여러 키워드는 최대 255문자까지 공백으로 분리되어야 합니다. (키워드에 공백이 포함되는 경우에는 따옴표를 사용하여 이를 분리하십시오.)

버전 탭

리포지토리에 저장된 오브젝트에는 여러 버전이 있을 수 있습니다. 버전 탭은 각 버전에 대한 정보를 표시합니다. 저장된 오브젝트의 특정 버전에 대해 다음 특성을 지정하거나 수정할 수 있습니다.

버전. 버전이 저장될 때 시간을 기반으로 생성된 버전의 고유한 ID입니다.

레이블. 버전의 현재 레이블(있는 경우). 버전 ID와는 달리, 레이블은 오브젝트의 한 버전에서 다른 버전으로 이동할 수 있습니다.

파일 크기, 작성 날짜 및 작성자 또한 각 버전마다 표시됩니다.

레이블 편집. 버전 탭 오른쪽 상단의 레이블 편집 아이콘을 클릭하여 저장된 오브젝트의 레이블을 정의, 적용 또는 제거하십시오. 자세한 정보는 [185 페이지의 『오브젝트 버전 레이블 관리』](#)의 내용을 참조하십시오.

권한 탭

권한 탭에서 오브젝트의 읽기 및 쓰기 권한을 설정할 수 있습니다. 현재 오브젝트에 대한 액세스가 있는 모든 사용자와 그룹이 나열됩니다. 권한은 계층 구조를 따릅니다. 예를 들어, 읽기 권한이 없는 경우에는 쓰기 권한을 가질 수 없습니다. 쓰기 권한이 없는 경우에는 삭제 권한을 가질 수 없습니다.

사용자 및 그룹 이 오브젝트에 대해 읽기 액세스 이상이 있는 리포지토리 사용자 및 그룹을 나열합니다. 쓰기 및 삭제 확인 상자를 선택하여 이 오브젝트의 액세스 권한을 특정 사용자나 그룹에 추가하십시오. 권한 탭의 오른쪽에 있는 **사용자/그룹 추가** 아이콘을 클릭하여 추가 사용자 및 그룹에 액세스를 지정하십시오. 사용 가능한 사용자와 그룹 목록은 관리자에 의해 제어됩니다.

오브젝트 버전 레이블 관리

버전 레이블 편집 대화 상자를 사용하면 다음을 수행할 수 있습니다.

- 선택한 오브젝트에 레이블 적용
- 선택한 오브젝트에서 레이블 제거
- 새 레이블을 정의하고 이를 오브젝트에 적용

레이블을 오브젝트에 적용

- 사용 가능한 레이블 목록에서 하나 이상의 레이블을 선택하십시오.
- 오른쪽 화살표 단추를 클릭하여 선택한 레이블을 적용된 레이블 목록으로 이동하십시오.
- 확인을 클릭하십시오.

오브젝트에서 레이블 제거

- 적용된 레이블 목록에서 하나 이상의 레이블을 선택하십시오.
- 왼쪽 화살표 단추를 클릭하여 선택한 레이블을 사용 가능한 레이블 목록으로 이동하십시오.
- 확인을 클릭하십시오.

새 레이블을 정의하고 이를 오브젝트에 적용

- 새 레이블 필드에 레이블 이름을 입력하십시오.
- 오른쪽 화살표 단추를 클릭하여 새 레이블을 적용된 레이블 목록으로 이동하십시오.
- 확인을 클릭하십시오.

스트림 배포

스트림을 씬 클라이언트 애플리케이션 IBM SPSS Modeler Advantage와 함께 사용할 수 있게 하려면 이는 리포지토리에서 스트림(.str 파일)으로 배포되어야 합니다.

참고: 스코어링 분기에 둘 이상의 소스 노드가 있는 스트림을 배포할 수 없습니다.

스트림은 IBM SPSS Collaboration and Deployment Services의 엔터프라이즈 레벨 기능을 최대한 활용할 수 있습니다. 자세한 정보는 [175 페이지의 『리포지토리 오브젝트 저장 및 배포』](#)의 내용을 참조하십시오.

현재 스트림을 배포하려면(파일 메뉴 방법)

1. 기본 메뉴에서 다음을 클릭하십시오.

파일 > 저장 > 배포

2. 배포 유형을 선택하고 필요에 따라 대화 상자의 나머지 부분을 완료하십시오.
3. 스트림으로 배포를 클릭하여 스트림을 IBM SPSS Modeler Advantage 또는 IBM SPSS Collaboration and Deployment Services와 함께 사용할 수 있도록 배포하십시오.
4. 저장을 클릭하십시오. 자세한 정보는 [도움말](#)을 클릭하십시오.
5. "배포 프로세스 완료"에서 계속하십시오.

현재 스트림 배포하기(도구 메뉴 방법)

1. 기본 메뉴에서 다음을 클릭하십시오.

도구 > 스트림 특성 > 배포

2. 배포 유형을 선택하고 필요에 따라 배포 탭의 나머지를 완료하고 저장을 클릭하십시오. 자세한 정보는 [186 페이지의 『스트림 배포 옵션』](#)의 내용을 참조하십시오.

배포 프로세스 완료하기

1. 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 [176 페이지의 『리포지토리에 연결』](#)의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.
2. 리포지토리: 저장 대화 상자에서 오브젝트를 저장하려는 폴더를 선택하고 기록할 다른 정보를 지정한 후 저장을 클릭하십시오. 자세한 정보는 [177 페이지의 『오브젝트 특성 설정』](#)의 내용을 참조하십시오.

스트림 배포 옵션

스트림 옵션 대화 상자에서 배포 탭을 사용하면 스트림 배포를 위한 옵션을 지정할 수 있습니다.

스트림을 배포할 때 확장자가 .str인 파일로 리포지토리에 저장됩니다.

스트림을 배포하면 다중 사용자 액세스, 자동화된 스코어링, 모델 새로 고침 및 Champion Challenger 분석 등과 같이 IBM SPSS Collaboration and Deployment Services에서 사용 가능한 추가 기능을 사용할 수 있습니다.

참고: 연관 규칙, STP 및 TCM 모델링 노드는 IBM SPSS Collaboration and Deployment Services에서 모델 평가 또는 Champion Challenger 단계를 지원하지 않습니다.

배포 탭에서 스트림에 대해 IBM SPSS Modeler가 작성하는 스트림 설명을 미리볼 수도 있습니다. 자세한 정보는 [44 페이지의 『스트림 설명』](#)의 내용을 참조하십시오.

배포 유형. 스트림을 배포할 방법을 선택하십시오. 모든 스트림은 배포하려면 먼저 전용 스코어링 노드가 있어야 합니다. 추가 요구사항과 옵션은 배포 유형에 따라 다릅니다.

참고: 연관 규칙, STP 및 TCM 모델링 노드는 IBM SPSS Collaboration and Deployment Services에서 모델 평가 또는 Champion Challenger 단계를 지원하지 않습니다.

- <없음>. 스트림은 리포지토리에 배포되지 않습니다. 스트림 설명 미리보기를 제외하고 모든 옵션은 사용 안함으로 설정됩니다.
- **스코어링만.** 저장 단추를 클릭하면 스트림이 리포지토리에 배포됩니다. 데이터는 **스코어링 노드** 필드에 지정하는 노드를 사용하여 스코어될 수 있습니다.
- **모델 새로 고침.** 스코어링만과 마찬가지이지만 추가로 모델을 **모델링 노드**와 **모델 너깃** 필드에 지정하는 오브젝트를 사용하여 리포지토리에서 업데이트할 수 있습니다. 자동 모델 새로 고침은 기본적으로 IBM SPSS Collaboration and Deployment Services에서 지원되지 않으므로, 리포지토리에서 스트림을 실행할 때 이 기능을 사용하려면 이 배포 유형을 선택해야 합니다. 자세한 정보는 [188 페이지의 『모델 새로 고침』](#)을 참조하십시오.

스코어링 노드 데이터를 스코어링하기 위해 사용할 스트림 분기를 식별하기 위해 그레프, 출력 또는 내보내기 노드를 선택하십시오. 스트림에는 실제로 원하는 개수만큼의 유효 분기, 모델 및 터미널 노드를 포함할 수 있지만

단 하나의 스코어링 분기를 배포 용도로 지정해야 합니다. 이는 스트림을 배포하기 위한 가장 기본적인 요구사항입니다.

스코어링 모수 스코어링 분기가 실행될 때 수정할 수 있는 모수를 지정할 수 있습니다. 자세한 정보는 [187 페이지의 『스코어링 및 모델링 모수』](#)을 참조하십시오.

모델링 노드 모델 새로 고침을 위해 리포지토리에서 모델을 재생성하거나 업데이트하는 데 사용된 모델링 노드를 지정합니다. 모델 너깃에 지정된 것과 같은 유형의 모델링 노드여야 합니다.

모델 작성 모수 모델링 노드가 실행될 때 수정할 수 있는 모수를 지정할 수 있습니다. 자세한 정보는 [187 페이지의 『스코어링 및 모델링 모수』](#)을 참조하십시오.

모델 너깃 모델 새로 고치기를 위해 스트림이 리포지토리에서 업데이트될 때마다(일반적으로 스케줄된 작업의 일부로서) 업데이트되거나 재생성될 모델 너깃을 지정합니다. 모델은 스코어링 분기에 있어야 합니다. 스코어링 분기에는 여러 모델이 존재할 수 있지만 단 하나만 지정할 수 있습니다. 스트림이 처음에 작성될 때는 이는 새 데이터가 사용 가능할 때 업데이트되거나 재생성되는 효과적인 플레이스홀더 모델일 수 있음을 참고하십시오.

확인. 배포에 유효한 스트림인지 여부를 확인하려면 이 단추를 클릭하십시오. 모든 스트림은 배포하려면 전용 스코어링 노드가 있어야 합니다. 이러한 조건이 충족되지 않으면 오류 메시지가 표시됩니다.

저장. 유효한 경우에는 스트림을 배포합니다. 유효하지 않은 경우에는 오류 메시지가 표시됩니다. 수정 단추를 클릭하고 오류를 정정한 후 다시 시도하십시오.

미리보기 스트림 설명. IBM SPSS Modeler가 스트림을 위해 작성하는 스트림 설명의 내용을 볼 수 있습니다. 자세한 정보는 [44 페이지의 『스트림 설명』](#)을 참조하십시오.

참고: 연관 규칙, STP 및 TCM 모델링 노드는 IBM SPSS Collaboration and Deployment Services에서 모델 평가 또는 Champion Challenger 단계를 지원하지 않습니다.

스코어링 및 모델링 모수

스트림을 IBM SPSS Collaboration and Deployment Services에 배포할 때 모델이 업데이트되거나 스코어링될 때 어떤 모수를 보거나 편집할지를 선택할 수 있습니다. 예를 들어, 최대 및 최소값을 지정하거나 작업이 실행될 때마다 변경되는 다른 값을 지정할 수도 있습니다.

1. 스트림이 배포된 후에 모수를 보거나 편집할 수 있도록 표시하게 하려면 스코어링 모수 대화 상자의 목록에서 이를 선택하십시오.

사용 가능한 모수 목록은 스트림 특성 대화 상자의 모수 탭에 정의됩니다. 자세한 정보는 [41 페이지의 『스트림과 세션 모수 설정』](#)의 내용을 참조하십시오.

스코어링 분기

스트림을 배포 중이라면 스트림의 한 분기는 **스코어링 분기**(즉, 스코어링 노드를 포함하는 분기)로 지정되어야 합니다. 분기를 스코어링 분기로 지정하면 해당 분기는 스코어링 분기에서 너깃에 대한 모델 링크이므로 스트림 캔버스에서 강조 표시됩니다. 이 시각적 표시는 특히 스코어링 분기가 즉시 명백하지 않을 수 있는 여러 개의 분기가 있는 복합 스트림에서 유용합니다.

참고: 단 하나의 스트림 분기만을 스코어링 분기로 지정할 수 있습니다.

스트림에 이미 스코어링 분기가 정의되어 있으면 새로 지정된 분기는 이를 스코어링 분기로서 대체합니다. 사용자 정의 색상 옵션을 사용하여 스코어링 분기 표시의 색상을 설정할 수 있습니다. 자세한 정보는 [208 페이지의 『표시 옵션 설정』](#)의 내용을 참조하십시오.

표시/숨김 스트림 마크업 도구 모음 단추를 사용하여 스코어링 분기 표시를 표시하거나 숨길 수 있습니다.



그림 18. 표시/숨김 스트림 마크업 도구 모음 단추

배포를 위한 스코어링 분기 식별

터미널 노드의 팝업 메뉴 또는 도구 메뉴에서 스코어링 분기를 지정할 수 있습니다. 팝업 메뉴를 사용하는 경우 스코어링 노드는 스트림 특성의 배포 탭에 자동으로 설정됩니다.

분기를 스코어링 분기로 지정하기(팝업 메뉴)

1. 모델 너깃을 터미널 노드에 연결합니다(너깃으로부터 처리 중 또는 출력 노드 다운스트림).
2. 터미널 노드를 마우스 오른쪽 단추로 클릭하십시오.
3. 메뉴에서 **스코어링 분기로 사용**을 클릭하십시오.

분기를 스코어링 분기로 지정하기(도구 메뉴)

1. 모델 너깃을 터미널 노드에 연결합니다(너깃으로부터 처리 중 또는 출력 노드 다운스트림).
2. 기본 메뉴에서 다음을 클릭하십시오.

도구 > 스트림 특성 > 배포

3. 배포 유형 목록에서 필요에 따라 스코어링 전용 또는 모델 새로 고침을 클릭하십시오. 자세한 정보는 [186 페이지의 『스트림 배포 옵션』](#)의 내용을 참조하십시오.
4. 스코어링 노드 필드를 클릭하고 목록에서 터미널 노드를 선택하십시오.
5. 확인을 클릭하십시오.

모델 새로 고침

모델 새로 고침은 새로운 데이터를 사용하는 스트림에서 기존 모델을 다시 작성하는 프로세스입니다. 스트림 자체는 리포지토리에서 변경되지 않습니다. 예를 들어, 알고리즘 유형과 스트림 특정 설정은 동일하게 남아 있지만 모델은 새 데이터에 대해 다시 교육을 받고, 모델의 새 버전이 이전 버전보다 잘 작동하는 경우 업데이트됩니다.

스트림에서 단 하나의 모델 너깃만을 새로 고침으로 설정할 수 있습니다. 이는 **새로 고침 모델**로 알려져 있습니다. 스트림 특성의 배포 탭에서 모델 새로 고침 옵션을 클릭하면([186 페이지의 『스트림 배포 옵션』](#) 참조) 그 당시에 지정한 모델 너깃이 새로 고침 모델이 됩니다. 모델 너깃의 팝업 메뉴에서 모델을 새로 고침 모델로 지정할 수도 있습니다. 이를 가능하게 하려면 너깃이 이미 스코어링 분기에 있어야 합니다.

너깃의 "새로 고침 모델" 상태를 끄면 이는 스트림의 배포 유형을 스코어링 전용으로 설정하는 것과 동등하게 되고 스트림 특성 대화 상자의 배포 탭이 이에 따라 업데이트됩니다. 현재 스코어링 분기에서 너깃의 팝업 메뉴에서 **새로 고침 모델로 사용** 옵션을 사용하여 이 상태를 켜고 끌 수 있습니다.

스코어링 분기에서 너깃의 모델 링크를 제거하면 너깃의 "새로 고침 모델" 상태 또한 제거됩니다. 편집 메뉴나 도구 모음을 사용하여 모델 링크 제거를 실행 취소할 수 있습니다. 이렇게 하면 너깃의 "새로 고침 모델" 상태가 다시 회복됩니다.

새로 고침 모델 선택 방법

스코어링 분기와 마찬가지로 새로 고침 모델에 대한 링크는 스트림에서도 강조 표시됩니다. 모델 너깃은 새로 고침 모델로 선택되었으므로 강조 표시된 링크는 스트림에 얼마나 많은 너깃이 있는지에 따라 다릅니다.

스트림에서 단일 모델

식별된 하나의 링크된 모델 너깃이 스코어링 분기에 있으면 해당 너깃은 스트림의 새로 고침 모델이 됩니다.

스트림의 다중 모델

스트림에 둘 이상의 링크된 너깃이 있다면 새로 고침 모델이 다음과 같이 선택됩니다.

모델 너깃이 스트림 특성 대화 상자의 배포 탭에 정의되었고 스트림에도 정의되었으면 해당 너깃은 새로 고침 모델이 됩니다.

배포 탭에 너깃이 정의되어 있지 않거나 정의되었지만 스코어링 분기에 없는 경우에는 터미널 노드와 가장 가까운 너깃이 새로 고침 모델이 됩니다.

이후에 새로 고침 링크 같은 모든 모델 링크를 선택 취소하면 링크가 아니라 스코어링 분기만이 강조 표시됩니다. 배포 유형은 스코어링 전용으로 설정됩니다.

참고: 링크 중 하나를 바꾸기 상태로 설정하고 나머지 하나는 설정하지 않기로 선택할 수 있습니다. 이 경우 새로 고침 모델로 선택된 모델 너깃은 새로 고침 링크가 있는 모델 너깃이고 스코어링 분기가 지정된 경우 터미널 노드에 가장 가까운 모델 너깃입니다.

스트림에 모델 없음

스트림에 모델이 없거나 모델 링크가 없는 모델만이 있는 경우에는 배포 유형이 스코어링 전용으로 설정됩니다.

스코어링 분기의 오류 확인

스코어링 분기를 지정할 때 오류가 있는지 확인합니다.

오류가 발견되면 스코어링 분기는 스코어링 분기 오류 색상으로 강조 표시되고 오류 메시지가 표시됩니다. 사용자 정의 색상 옵션을 사용하여 오류 색상을 설정할 수 있습니다. 자세한 정보는 [208 페이지의 『표시 옵션 설정』](#)의 내용을 참조하십시오.

오류가 발견되면 다음과 같이 진행하십시오.

1. 오류 메시지의 내용에 따라 오류를 수정하십시오.
2. 기본 메뉴에서 다음을 클릭하십시오.

도구 > 스트림 특성 > 배포

확인을 클릭하십시오.

3. 필요하면 오류가 발견되지 않을 때까지 이 프로세스를 반복하십시오.

제 12 장 IBM Cloud Pak for Data에 스트림 저장

스트림을 IBM Cloud Pak for Data Server에 저장하고 여기서 스트림을 플로우로 열어 Cloud Pak for Data에서 사용 가능한 여러 기능을 활용할 수 있습니다.

Cloud Pak for Data는 데이터 수집, 구성 및 분석 방법을 단순화하고 비즈니스 전반에 걸쳐 AI를 주입합니다. 광범위한 데이터 및 AI 소프트웨어, 선도적인 거버넌스 및 보안, 협업을 위한 통합 환경을 원활하게 통합합니다. 추가 정보는 <https://www.ibm.com/products/cloud-pak-for-data>의 내용을 참조하십시오. Cloud Pak for Data에 관심이 있는 경우 IBM 담당자에게 문의하십시오.

서버 연결을 구성하는 방법

이 지시사항은 Firefox를 기반으로 합니다. 사용하는 브라우저에 따라 단계가 달라질 수 있습니다.

1. Firefox를 열고 주소 표시줄에 Cloud Pak for Data URL을 붙여넣은 후 Enter 키를 누르십시오.

URL을 확인하려면 Cloud Pak for Data 웹 클라이언트를 열고 브라우저 주소 표시줄의 URL(/zen 앞까지의 모든 항목)을 검토하십시오. 기본적으로 URL 형식은 `https://<namespace>-cpd-<namespace>.apps.<cluster-subdomain>`입니다.

각 항목에 대한 설명은 다음과 같습니다.

- <namespace>는 사용자의 네임스페이스입니다(예: ab123).
- <cluster-subdomain>은 클러스터의 하위 도메인 이름입니다(예: ocp454x86-aqt-mycompany.com).

이 예의 경우 URL은 `https://ab123-cpd-ab123.apps.ocp454x86-aqt-mycompany.com`입니다.

Cloud Pak for Data URL에 대한 자세한 정보는 [Cloud Pak for Data 문서](#)를 참조하십시오.

2. 주소 표시줄 왼쪽에 있는 노란색 자물쇠 아이콘을 클릭한 후 **안전하지 않은 연결** 옆의 화살표를 클릭하십시오.
3. **자세한 정보**를 클릭하십시오. 페이지 정보 대화 상자가 열립니다.
4. 페이지 정보 대화 상자에서 **보안**으로 이동한 후 **인증서 보기**를 클릭하십시오.
5. 기타까지 아래로 스크롤하여 PEM 인증 파일을 다운로드하십시오.
6. 명령행에서 다음과 같이 인증 파일을 IBM SPSS Modeler 클라이언트 디렉토리로 가져오십시오.

Windows:

```
“%JAVA_HOME%\bin\keytool.exe -importcert -trustcacerts -file <Certification file absolute path> -alias <alias> -keystore "<Modeler client installation path>"\jre\lib\security\cacerts
```

macOS:

```
“%JAVA_HOME%/bin/keytool -importcert -trustcacerts -file <Certification file absolute path> -alias <alias> -keystore "<Modeler client installation path>"/jre/lib/security/cacerts
```

프롬프트가 표시되면 키 저장소 비밀번호로 `changeit`을 입력한 후 `yes`를 입력하여 인증을 신뢰하십시오.

7. IBM SPSS Modeler의 기본 메뉴에서 다음을 클릭하십시오.

도구 > IBM Cloud Pak for Data Server > IBM Cloud Pak for Data Server 연결...

8. 서버에 대한 연결 설정을 지정하십시오. 1단계에서 사용한 것과 동일한 URL을 사용하십시오. 사용할 사용자 이름 또는 비밀번호를 모를 경우 로컬 시스템 관리자에게 문의하십시오. API 키 로그인을 선택한 경우, 아래 방법으로 URL을 얻으십시오.

- Cloud Pak for Data 클러스터에 로그온하고 다음 명령을 실행하십시오. `oc get routes -n ibm-common-services`. 그런 다음, cp-console에 HOST/PORT 값을 복사하십시오. 다음 예의 경우 URL은 `https://cp-console.apps.wm146x04.cp.myserver.com`입니다.

```
[root@wm146x04-inf ocp-scripts]# oc get routes -n ibm-common-services
NAME           HOST/PORT          PATH      SERVICES
PORT   TERMINATION   WILDCARD
cp-console   cp-console.apps.wm146x04.cp.myserver.com      icp-management-ingress
https  reencrypt/Redirect  None
cp-proxy     cp-proxy.apps.wm146x04.cp.myserver.com      nginx-ingress-controller
https  passthrough/Redirect  None
```

- 연결을 클릭하여 서버에 연결하십시오.

Cloud Pak for Data에 스트림을 저장하는 방법

- 기본 메뉴에서 다음을 클릭하십시오.

파일 > 플로우로 보관.... 스트림은 Cloud Pak for Data로 마이그레이션된 후 플로우라고 합니다.

또는 스트림 관리자 분할창을 마우스 오른쪽 단추로 클릭한 후 **플로우로 보관...**을 선택할 수도 있습니다.

- 앞에서 구성한 Cloud Pak for Data Server를 선택하십시오. 그런 다음 플로우를 저장할 프로젝트를 선택하십시오.

팁: 저장할 스트림을 한 번에 여러 개 선택할 수도 있습니다. 기본 메뉴에서 **도구 > IBM Cloud Pak for Data Server... > IBM Cloud Pak for Data Server에 스트림 대량 업로드**를 클릭하십시오. 그러면 로컬 파일 시스템에서 스트림을 선택하고 서버를 선택한 후 프로젝트를 선택하라는 프롬프트가 표시됩니다.

제 13 장 외부 애플리케이션으로 내보내기

외부 애플리케이션으로 내보내기 정보

IBM SPSS Modeler에서는 전체 데이터 마이닝 프로세스를 외부 애플리케이션으로 내보내기 위한 많은 메커니즘을 제공하므로 데이터를 준비하고 모델을 작성하기 위해 수행하는 작업을 IBM SPSS Modeler 외부에서도 사용자의 편의를 위해 사용할 수 있습니다.

이전 섹션에서는 다중 사용자 액세스, 작업 스케줄링 및 기타 기능을 활용하여 IBM SPSS Collaboration and Deployment Services 리포지토리에 스트림을 배포하는 방법을 보여주었습니다. 이와 비슷한 방법으로 IBM SPSS Modeler 스트림은 다음과 결합하여 사용할 수도 있습니다.

- IBM SPSS Modeler Advantage
- PMML 형식으로 파일을 가져오고 내보낼 수 있는 애플리케이션

IBM SPSS Modeler Advantage와 함께 스트림 사용에 대한 자세한 정보는 [193 페이지의 『스트림을 IBM SPSS Modeler Advantage에서 열기』](#)의 내용을 참조하십시오.

모델을 PMML 파일로 내보내고 가져오고, 모델을 이 형식을 지원하는 다른 애플리케이션과 공유 가능하게 만드는 방법에 대한 자세한 정보는 [193 페이지의 『모델을 PMML로 가져오기 및 내보내기』](#)의 내용을 참조하십시오.

스트림을 IBM SPSS Modeler Advantage에서 열기

IBM SPSS Modeler 스트림은 씬 클라이언트 애플리케이션 IBM SPSS Modeler Advantage와 함께 사용할 수 있습니다. 사용자 정의된 애플리케이션을 IBM SPSS Modeler Advantage 내에서 전적으로 작성하는 것이 가능하긴 하지만 IBM SPSS Modeler에서 이미 작성된 스트림을 애플리케이션 작업흐름의 기초로 사용할 수도 있습니다.

IBM SPSS Modeler Advantage에서 스트림 열기

1. 스트림을 IBM SPSS Collaboration and Deployment Services 리포지토리에 배포하고, [스트림으로 배포](#) 읍션을 반드시 클릭하십시오. 자세한 정보는 [185 페이지의 『스트림 배포』](#)의 내용을 참조하십시오.
2. IBM SPSS Modeler Advantage 도구 모음 단추에서 열기를 클릭하거나 주 메뉴에서 다음을 클릭하십시오.

파일 > **IBM SPSS Modeler Advantage**에서 열기

1. 필요한 경우 리포지토리에 연결 설정을 지정하십시오. 자세한 정보는 [176 페이지의 『리포지토리에 연결』](#)의 내용을 참조하십시오. 특정 포트, 비밀번호 및 기타 연결 세부사항에 대해서는 로컬 시스템 관리자에게 문의하십시오.

참고: 리포지토리 서버에는 또한 IBM SPSS Modeler Advantage 소프트웨어가 설치되어 있어야 합니다.

1. 리포지토리: 저장 대화 상자에서 오브젝트를 저장하려는 폴더를 선택하고 기록할 다른 정보를 지정한 다음 저장 단추를 클릭하십시오. 자세한 정보는 [177 페이지의 『오브젝트 특성 설정』](#)의 내용을 참조하십시오.

그리면 스트림이 이미 열린 상태로 IBM SPSS Modeler Advantage가 실행됩니다. 스트림은 IBM SPSS Modeler에서 닫힙니다.

모델을 PMML로 가져오기 및 내보내기

PMML 또는 예측 모델 마크업 언어는 모델에 대한 입력, 데이터를 데이터 마이닝을 위해 준비하는 데 사용하는 변환 및 모델 자체를 정의하는 모수를 포함하여 데이터 마이닝과 통계 모델을 설명하기 위한 XML 형식입니다. IBM SPSS Modeler에서는 PMML을 가져오고 내보내고, IBM SPSS Statistics 등과 같이 이 형식을 지원하는 다른 애플리케이션과 모델을 공유할 수 있게 만들 수 있습니다.

PMML에 대한 자세한 정보는 데이터 마이닝 그룹 웹 사이트(<http://www.dmg.org>)를 참조하십시오.

모델 내보내기

PMML 내보내기는 IBM SPSS Modeler에서 생성되는 대부분의 모델 유형에 지원됩니다. 자세한 정보는 [194 페이지의 『PMML을 지원하는 모델 유형』](#)의 내용을 참조하십시오.

1. 모델 팔레트에서 모델 너깃을 마우스 오른쪽 단추로 클릭하십시오. (또는 캔버스에서 모델 너깃을 두 번 클릭하고 파일 메뉴를 선택하십시오.)
2. 메뉴에서 **PMML 내보내기**를 클릭하십시오.
3. 내보내기(또는 저장) 대화 상자에서 대상 디렉토리와 모델의 고유 이름을 지정하십시오.

참고:

PMML 내보내기 옵션은 사용자 옵션 대화 상자에서 변경할 수 있습니다. 기본 메뉴에서 다음을 클릭하십시오.

도구 > 옵션 > 사용자 옵션

PMML 탭을 클릭하십시오.

자세한 정보는 [208 페이지의 『PMML 내보내기 옵션 설정』](#)의 내용을 참조하십시오.

PMML로 저장된 모델 가져오기

IBM SPSS Modeler 또는 또 다른 애플리케이션에서 PMML로서 내보낸 모델은 모델 팔레트로 가져올 수 있습니다. 자세한 정보는 [194 페이지의 『PMML을 지원하는 모델 유형』](#)의 내용을 참조하십시오.

1. 모델 팔레트에서 팔레트를 마우스 오른쪽 단추로 클릭하고 메뉴에서 **PMML 가져오기**를 선택하십시오.
2. 가져올 파일을 선택하고 필요에 따라 변수 레이블의 옵션을 지정하십시오.
3. 열기를 클릭하십시오.

모델에 있는 경우 변수 레이블을 사용하십시오. PMML은 데이터 사전에서 변수에 변수 이름과 변수 레이블(예: *RefID*의 경우 Referrer ID) 둘 모두를 지정할 수도 있습니다. 변수 레이블이 원래 내보낸 PMML에 있는 경우 이를 사용하려면 이 옵션을 선택하십시오.

변수 레이블 옵션을 선택했지만 PMML에 변수 레이블이 없는 경우에는 변수 이름을 통상적으로 사용합니다.

PMML을 지원하는 모델 유형

PMML 내보내기

IBM SPSS Modeler 모델. IBM SPSS Modeler에서 작성된 다음 모델은 PMML 4.0으로서 내보낼 수 있습니다.

- C&R 트리
- QUEST
- CHAID
- 신경망
- C5.0
- 로지스틱 회귀분석
- Genlin
- SVM
- Apriori
- Carma
- K-평균
- 코호넨
- TwoStep
- TwoStep-AS
- GLMM(PMML은 모든 GLMM 모델용으로 내보낼 수 있지만, 고정 효과만 있습니다.)

- 의사결정 목록
- Cox
- 시퀀스(시퀀스 PMML 모델에 대한 스코어링은 지원되지 않음)
- 랜덤 트리
- Tree-AS
- 선형
- Linear-AS
- 회귀분석
- 로지스틱
- GLE
- LSVM
- 이상 항목 발견
- KNN
- 연관 규칙

데이터베이스 원시 모델. 데이터베이스 원시 알고리즘을 사용하여 생성된 모델의 경우 PMML 내보내기를 사용할 수 없음. Microsoft 또는 Oracle Data Miner에서 분석 서비스를 사용하여 작성된 모델은 내보낼 수 없습니다.

PMML 가져오기

IBM SPSS Modeler는 IBM SPSS Modeler에서 내보낸 모델뿐만 아니라 IBM SPSS Statistics 17.0 이상에 의해 생성된 모델 또는 변환 PMML을 포함하여 모든 IBM SPSS Statistics 제품의 현재 버전에 의해 생성된 PMML 모델을 가져오고 스코어링할 수 있습니다. 본질적으로, 이는 스코어링 엔진이 스코어링할 수 있는 모든 PMML을 의미하며 다음과 같은 예외가 있습니다.

- Apriori, CARMA, 이상 항목 발견, 시퀀스 및 연관 규칙 모델은 가져올 수 없습니다.
- PMML 모델은 스코어링에 사용할 수 있더라도 IBM SPSS Modeler에 가져온 후에는 찾아볼 수 없을 수 있습니다. (여기에는 우선 IBM SPSS Modeler에서 내보낸 모델이 포함됨을 유의하십시오. 이 제한을 피하려면 모델을 PMML이 아니라 생성된 모델 파일[*.gm]로서 내보내십시오.)
- 가져올 때는 제한된 검증이 발생하지만 모델을 스코어링하려고 시도할 때 전체 검증이 수행됩니다. 따라서 가져오기가 성공할 수는 있지만 스코어링은 실패하거나 부정확한 결과를 낼 수 있습니다.

참고: IBM SPSS Modeler에 가져온 타사 PMML의 경우, IBM SPSS Modeler는 인지되고 스코어된 유효한 PMML을 스코어링하려고 시도합니다. 모든 PMML이 스코어할지 또는 이를 생성한 애플리케이션과 같은 방식으로 스코어할지는 보장이 되지 않습니다.

제 14 장 프로젝트 및 보고서

프로젝트에 대한 소개

프로젝트는 데이터 마이닝 태스크에 관련된 파일 그룹입니다. 프로젝트에는 데이터 스트림, 그래프, 모형 생성, 보고서 및 그 밖에 IBM SPSS Modeler에서 작성된 것이 포함됩니다. 처음에는 IBM SPSS Modeler 프로젝트가 단순히 결과를 구성하는 방법이지만, 실제로는 더 많은 것을 할 수 있습니다. 프로젝트를 사용하여 할 수 있는 것은 다음과 같습니다.

- 프로젝트 파일의 각 개체에 주석(Annotation)을 달 수 있습니다.
- 데이터 마이닝에 대한 노력을 이끌어 주는 CRISP-DM 방법론을 사용할 수 있습니다. 또한 프로젝트에는 CRISP-DM을 사용한 데이터 마이닝에 대한 세부사항 및 실제 예를 제공하는 CRISP-DM 도움말 시스템이 포함되어 있습니다.
- 비IBM SPSS Modeler 개체를 프로젝트(예: 데이터 마이닝 목표를 나타내는 데 사용되는 PowerPoint 슬라이드 쇼 또는 사용하려는 알고리즘에 대한 백서)에 추가할 수 있습니다.
- 주석(Annotation)을 기반으로 한 포괄적이고 단순한 업데이트 보고서를 만들 수 있습니다. 이 보고서는 조직의 인트라넷에 쉽게 게재할 수 있도록 HTML로 생성될 수 있습니다.

참고: 프로젝트 분할창이 IBM SPSS Modeler 창에 표시되지 않는 경우, 보기 메뉴에서 **프로젝트**를 클릭하십시오.

프로젝트에 추가하는 개체는 두 가지 방식, **클래스 보기**와 **CRISP-DM 보기**로 표시될 수 있습니다. 사용자가 프로젝트에 추가하는 것은 두 개의 보기로 표시되며, 보기 간에 전환이 가능하므로 가장 잘 작동하는 구성을 작성할 수 있습니다.

CRISP-DM 보기

CRISP-DM(Cross-Industry Standard Process for Data Mining)을 지원하면 IBM SPSS Modeler 프로젝트는 데이터 마이닝 노력의 일부를 구성하는 업계에서 증명된 비독점 방법을 제공합니다. CRISP-DM은 6개의 단계를 사용하여 시작(비즈니스 요구사항 수집)부터 마침(결과 배포)까지의 프로세스를 설명합니다. 몇몇 단계는 일반적으로 IBM SPSS Modeler에서의 작업과 관련되지 않지만 프로젝트 분할창에는 프로젝트와 연관된 모든 자료를 저장하고 추적하기 위한 중앙 위치를 가질 수 있도록 6개의 단계가 모두 포함됩니다. 예를 들어, 비즈니스 이해 단계는 일반적으로 요구사항 수집과 IBM SPSS Modeler에서 데이터에 대해 작업하는 대신에 목표를 결정하기 위해 동료와 미팅하는 것이 관련됩니다. 프로젝트 분할창을 사용하면 이러한 미팅에서 나온 사용자의 노트를 나중에 참조하거나 보고서에 포함하기 위해 비즈니스 이해 폴더에 저장할 수 있습니다.

프로젝트 분할창에서 CRISP-DM 보기는 또한 데이터 마이닝 라이프 사이클을 안내하기 위해 자체 도움말 시스템을 갖추고 있습니다. IBM SPSS Modeler에서 이 도움말은 도움말 메뉴에서 **CRISP-DM 도움말**을 클릭하여 액세스할 수 있습니다.

참고: 프로젝트 분할창이 창에 표시되지 않으면 보기 메뉴에서 **프로젝트**를 클릭하십시오.

기본 프로젝트 단계 설정

프로젝트에 추가된 오브젝트는 CRISP-DM의 기본 단계에 추가됩니다. 즉 오브젝트를 사용하는 데이터 마이닝 단계에 따라 수동으로 오브젝트를 구성해야 함을 의미합니다. 현재 작업 중인 단계에 기본 폴더를 설정하는 것이 바람직합니다.

기본값으로 사용할 단계 선택

1. CRISP-DM 보기에서 기본값으로 설정할 단계의 폴더를 마우스 오른쪽 단추로 클릭하십시오.
2. 메뉴에서 **기본값으로 설정**을 클릭하십시오.

기본 폴더는 굵은 글자로 표시됩니다.

클래스 보기

프로젝트 분할창에서 클래스 보기는 사용자의 작업을 IBM SPSS Modeler에서 작성된 오브젝트의 유형별로 범주별로 구성합니다. 저장된 오브젝트는 다음 범주에 추가할 수 있습니다.

- 스트림
- 노드
- 모델
- 테이블, 그래프, 보고서
- 기타(슬라이드 쇼 또는 데이터 마이닝 작업과 관련된 백서 등과 같이 IBM SPSS Modeler가 아닌 파일)

클래스 보기에서 오브젝트를 추가하면 CRISP-DM 보기에서 기본 단계 폴더에도 추가됩니다.

참고: 프로젝트 분할창이 창에 표시되지 않으면 보기 메뉴에서 **프로젝트**를 클릭하십시오.

프로젝트 작성

프로젝트는 본질적으로 프로젝트와 연관시키는 모든 파일에 대한 참조를 포함하는 파일입니다. 즉 프로젝트 항목이 개별적으로 저장되고 프로젝트 파일(.cpj)에 참조로서도 저장된다는 의미입니다. 이 참조 구조 때문에 다음을 참조하십시오.

- 프로젝트 항목은 먼저 프로젝트에 추가되기 전에 개별적으로 저장되어야 합니다. 항목이 저장되지 않은 경우에는 현재 프로젝트에 추가하기 전에 먼저 저장하라는 메시지가 표시됩니다.
- 스트림 등과 같이 개별적으로 업데이트된 오브젝트는 또한 프로젝트 파일에서도 업데이트됩니다.
- 파일 시스템에서 스트림, 노드 및 출력 오브젝트 등과 같은 오브젝트를 수동으로 이동하거나 삭제하면 프로젝트 파일에 있는 링크가 무효한 것으로 렌더링됩니다.

새 프로젝트 작성

새 프로젝트는 IBM SPSS Modeler 창에서 쉽게 작성할 수 있습니다. 열려 있는 프로젝트가 없으면 새로 작성하거나, 기존 프로젝트를 닫고 처음부터 다시 시작할 수 있습니다.

기본 메뉴에서 다음을 클릭하십시오.

파일 > 프로젝트 > 새 프로젝트...

프로젝트에 추가

프로젝트를 작성하거나 연 후에는 여러 방법을 사용하여 데이터 스트림, 노드 및 보고서 등과 같은 오브젝트를 추가할 수 있습니다.

관리자에서 오브젝트 추가

IBM SPSS Modeler 창의 오른쪽 상단 구석에서 관리자를 사용하여 스트림이나 출력을 추가할 수 있습니다.

1. 관리자 탭 중 하나에서 테이블 또는 스트림 등과 같은 오브젝트를 선택할 수 있습니다.
2. 마우스 오른쪽 단추로 클릭하고 **프로젝트에 추가**를 클릭하십시오.

오브젝트가 이미 저장된 경우에는 적합한 오브젝트 폴더(클래스 보기에서)나 기본 단계 폴더(CRISP-DM 보기)로 자동으로 추가됩니다.

3. 또는 관리자에서 프로젝트 분할창까지 오브젝트를 끌어다 놓을 수 있습니다.

참고: 먼저 오브젝트를 저장하도록 요청받을 수 있습니다. 저장할 때 저장 대화 상자에서 **프로젝트에 파일 추가**가 선택되었는지 확인하십시오. 그러면 저장 후에 프로젝트에 자동으로 오브젝트가 추가됩니다.

캔버스에서 노드 추가

저장 대화 상자를 사용하여 스트림 캔버스에서 개별 노드를 추가할 수 있습니다.

1. 캔버스에서 노드를 선택하십시오.
2. 마우스 오른쪽 단추를 클릭하고 **노드 저장**을 클릭하십시오. 또는 기본 메뉴에서 다음을 클릭하십시오.

편집 > 노드 > 노드 저장...

3. 저장 대화 상자에서 **프로젝트에 파일 추가**를 선택하십시오.
4. 노드의 이름을 작성하고 **저장**을 클릭하십시오.

그러면 파일이 저장되고 프로젝트에 추가됩니다. 노드는 클래스 보기에서는 노드 폴더에 추가되고 CRISP-DM 보기에서는 기본 단계 폴더에 추가됩니다.

외부 파일 추가

다양한 비IBM SPSS Modeler 오브젝트를 프로젝트에 추가할 수 있습니다. 이는 IBM SPSS Modeler 내에서 전체 데이터 마이닝 프로세스를 관리할 때 유용합니다. 예를 들어, 데이터, 노트, 프리젠테이션 및 그래픽에 대한 링크를 프로젝트에 저장할 수 있습니다. CRISP-DM 보기에서 외부 파일은 사용자가 선택하는 폴더에 추가할 수 있습니다. 클래스 보기에서 외부 파일은 기타 폴더에만 저장할 수 있습니다.

외부 파일을 프로젝트에 추가

1. 파일을 데스크탑에서 프로젝트까지 끌십시오.

or

2. CRISP-DM 또는 클래스 보기에서 대상 폴더를 마우스 오른쪽 단추로 클릭하십시오.
3. 메뉴에서 **폴더에 추가**를 클릭하십시오.
4. 대화 상자에서 파일을 선택하고 **열기**를 클릭하십시오.

그러면 IBM SPSS Modeler 프로젝트 내에 선택한 오브젝트에 대한 참조가 추가됩니다.

프로젝트를 IBM SPSS Collaboration and Deployment Services Repository에 전송

모든 구성요소 파일을 포함하여 전체 프로젝트를 한 단계에서 IBM SPSS Collaboration and Deployment Services Repository로 전송할 수 있습니다. 대상 위치에 이미 있는 모든 오브젝트는 이동되지 않습니다. 이 기능은 반대로도 작동합니다. 전체 프로젝트를 IBM SPSS Collaboration and Deployment Services Repository에서 로컬 파일 시스템으로 전송할 수도 있습니다.

프로젝트 전송

전송하려는 프로젝트가 프로젝트 분할창에 열려 있는지 확인하십시오.

프로젝트 전송

1. 루트 프로젝트 폴더를 마우스 오른쪽 단추로 클릭하고 **프로젝트 전송**을 클릭하십시오.
2. 프롬프트가 표시되면 IBM SPSS Collaboration and Deployment Services Repository에 로그인하십시오.
3. 프로젝트의 새 위치를 지정하고 **확인**을 클릭하십시오.

프로젝트 특성 설정

프로젝트 특성 대화 상자를 사용하여 프로젝트의 내용과 문서를 사용자 정의할 수 있습니다. 프로젝트 특성에 액세스

1. 프로젝트 분할창에서 오브젝트나 폴더를 마우스 오른쪽 단추로 클릭하고 **프로젝트 특성**을 클릭하십시오.
2. **프로젝트** 탭을 클릭하여 기본 프로젝트 정보를 지정하십시오.

작성 날짜. 프로젝트의 작성 날짜를 보여줍니다(편집 불가능).

요약. 프로젝트 보고서에 표시될 데이터 마이닝 프로젝트의 요약을 입력할 수 있습니다.

내용. 프로젝트 파일이 참조하는 구성요소 유형과 수를 나열합니다(편집 불가능).

저장되지 않은 오브젝트를 다른 이름으로 저장. 저장되지 않은 오브젝트를 로컬 파일 시스템에 저장하는지 리포지토리에 저장하는지 여부를 지정합니다. 자세한 정보는 [175 페이지의 『IBM SPSS Collaboration and Deployment Services Repository 정보』](#)의 내용을 참조하십시오.

프로젝트를 로드할 때 오브젝트 참조를 업데이트. 프로젝트의 참조를 해당 구성요소로 업데이트하려면 이 옵션을 선택하십시오. 참고: 프로젝트에 추가된 파일은 프로젝트 파일 자체에 저장되지 않습니다. 대신 파일에 대한 참조가 프로젝트에 저장됩니다. 즉 파일을 이동하거나 삭제하면 프로젝트에서 오브젝트가 제거됩니다.

프로젝트에 주석달기

프로젝트 분할창은 데이터 마이닝 노력에 주석(Annotation)을 달기 위한 수많은 방법을 제공합니다. 프로젝트 수준 주석(Annotation)은 종종 "큰 그림" 목표와 의사결정을 추적하는 데 사용하는 반면에, 폴더나 노드 주석(Annotation)은 추가 세부사항을 제공합니다. 주석(Annotation) 탭은 검색할 수 없는 결측 데이터가 있는 데이터 제외 또는 데이터 탐색 동안 형성된 유망한 가설 등과 같은 프로젝트 수준 세부사항을 문서화하기에 충분한 공간을 제공합니다.

프로젝트에 주석(Annotation) 달기:

1. CRISP-DM 또는 클래스 보기에서 프로젝트 폴더를 선택하십시오.
2. 폴더를 마우스 오른쪽 단추로 클릭하고 **프로젝트 특성**을 클릭하십시오.
3. **주석(Annotation)** 탭을 클릭하십시오.
4. 프로젝트를 설명하는 키워드와 텍스트를 입력하십시오.

폴더 특성 및 주석(Annotation)

개별 프로젝트 폴더(CRISP-DM 및 클래스 보기 둘 모두에서)에 주석(Annotation)이 작성됩니다. CRISP-DM 보기에서 이는 데이터 마이닝의 각 단계별 조직의 목표를 문서화하는 매우 효과적인 방법일 수 있습니다. 예를 들어, 비즈니스 이해 폴더에 주석(Annotation) 도구를 사용하면 "이 연구의 비즈니스 목표는 VIP 고객 중의 이탈을 줄이기 위한 것임" 등과 같은 문서를 포함할 수 있습니다. 이 텍스트는 그런 다음 **보고서에 포함** 옵션을 선택하여 프로젝트 보고서에 자동으로 포함될 수 있습니다.

폴더에 주석(Annotation) 달기:

1. 프로젝트 분할창에서 폴더를 선택하십시오.
2. 폴더를 마우스 오른쪽 단추로 클릭하고 **폴더 특성**을 클릭하십시오.

CRISP-DM 보기에서 폴더에는 각 단계의 목적의 요약뿐만 아니라 관련 데이터 마이닝 작업 완료에 대한 안내가 주석(Annotation)으로 작성됩니다. 이러한 주석(Annotation)을 제거하거나 편집할 수 있습니다.

이름. 이 영역은 선택된 필드의 이름을 표시합니다.

도구 템 텍스트. 마우스 포인터를 프로젝트 폴더 위에서 움직일 때 표시되는 사용자 정의 도구 템을 작성하십시오. 이는 CRISP-DM 보기에서 유용합니다. 예를 들어, 각 단계의 목표의 빠른 개요를 제공하거나 단계의 상태를 표시하려고 할 때입니다(예: "진행 중" 또는 "완료").

주석(Annotation) 필드. 프로젝트 보고서에서 대조할 수 있는 보다 길이가 긴 주석(Annotation)에는 이 필드를 사용하십시오. CRISP-DM 보기에는 주석(Annotation)에 각 데이터 마이닝 단계의 설명이 포함되지만 사용자 고유의 프로젝트에 맞게 이를 자유롭게 사용자 정의할 수 있습니다.

보고서에 포함. 보고서에 주석(Annotation)을 포함하려면 **보고서에 포함**을 선택하십시오.

오브젝트 특성

오브젝트 특성을 보고 개별 오브젝트를 프로젝트 보고서에 포함할지 여부를 선택할 수 있습니다. 오브젝트 특성에 액세스

1. 프로젝트 분할창에서 오브젝트를 마우스 오른쪽 단추로 클릭합니다.
2. 메뉴에서 **오브젝트 특성**을 클릭합니다.

이름. 이 영역은 저장된 오브젝트의 이름을 나열합니다.

경로. 이 영역은 저장된 오브젝트의 위치를 나열합니다.

보고서에 포함. 생성된 보고서에 오브젝트 세부사항을 포함하려면 이 옵션을 선택하십시오.

프로젝트 닫기

IBM SPSS Modeler를 종료하거나 새 프로젝트를 열 때 기존 프로젝트 파일(.cpj)이 닫힙니다.

프로젝트와 연관된 몇몇 파일(예: 스트림, 노드 또는 그래프)은 여전히 열려 있을 수 있습니다. 이러한 파일을 열어 두려면 ... 이러한 파일을 저장하고 닫으시겠습니까? 메시지에 아니오로 응답하십시오.

프로젝트를 닫은 후에 연관된 파일을 수정하고 저장하는 경우에는 이러한 업데이트된 버전은 다음 번에 프로젝트를 열 때 프로젝트에 포함됩니다. 이러한 작동을 막기 위해서는 프로젝트에서 파일을 제거하거나 이를 다른 파일 이름으로 저장하십시오.

보고서 생성

프로젝트의 가장 유용한 기능 중 하나는 프로젝트 항목과 주석(Annotation)을 기반으로 보고서를 생성하는 기능입니다. CRISP-DM 방법론에서 논의된 것처럼 이는 효과적인 데이터 마이닝을 위한 중요한 구성요소입니다. 보고서를 몇몇 파일 유형 중 하나로 직접 생성하거나 즉각적으로 볼 수 있도록 화면의 출력 창으로 생성할 수 있습니다. 여기에서 보고서를 웹 브라우저에서 인쇄, 저장 또는 볼 수 있습니다. 사용자의 조직에서 다른 사람에게 저장된 보고서를 배포할 수 있습니다.

보고서는 종종 프로젝트에 관련된 사람에게 배포하기 위해 데이터 마이닝 프로세스 중에 프로젝트 파일에서부터 몇 차례 생성되기도 합니다. 보고서는 프로젝트 파일뿐만 아니라 작성된 주석(Annotation)에서 참조된 오브젝트에 대한 정보를 모읍니다. 클래스 보기 또는 CRISP-DM 보기 기반으로 보고서를 작성할 수 있습니다.

보고서 생성:

1. CRISP-DM 또는 클래스 보기에서 프로젝트 폴더를 선택하십시오.
2. 폴더를 마우스 오른쪽 단추로 클릭하고 **프로젝트 보고서**를 클릭하십시오.
3. 보고서 옵션을 지정하고 **보고서 생성**을 클릭하십시오.

보고서 대화 상자에 있는 옵션은 필요한 보고서 유형을 생성하는 몇 가지 방법을 제공합니다.

출력 이름. 보고서의 출력을 화면에 전송하기로 선택한 경우에는 출력 창의 이름을 지정하십시오. 사용자 정의 이름을 지정하거나 IBM SPSS Modeler가 자동으로 창에 이름을 지정하도록 두십시오.

화면으로 출력. 보고서를 생성하고 출력 창에 표시하려면 이 옵션을 선택하십시오. 출력 창으로부터 보고서를 다양한 파일 유형으로 내보내는 옵션이 있음을 유의하십시오.

파일로 출력. 보고서를 파일 유형 목록에 지정된 파일 유형으로 생성하고 저장하려면 이 옵션을 선택하십시오.

파일 이름. 생성된 보고서의 파일 이름을 지정하십시오. 파일은 기본적으로 IBM SPSS Modeler |bin 디렉토리에 저장됩니다. 생략 기호(...)를 사용하여 다른 위치를 지정하십시오.

파일 유형. 사용 가능한 파일 유형은 다음과 같습니다.

- **HTML 문서.** 보고서는 단일 HTML 파일로 저장됩니다. 보고서에 그래프가 포함된 경우에는 이들은 PNG 파일로 저장되고 HTML 파일에 의해 참조됩니다. 인터넷에서 보고서를 공개할 때는 HTML 파일과 함께 이 파일이 참조하는 이미지도 업로드하십시오.
- **텍스트 문서.** 보고서는 단일 텍스트 파일로 저장됩니다. 보고서에 그래프가 포함된 경우에는 파일 이름과 경로 참조만이 보고서에 포함됩니다.
- **Microsoft Word 문서.** 보고서는 단일 문서로 저장되고 그래프가 문서에 직접 임베드됩니다.
- **Microsoft Excel 문서.** 보고서가 단일 스프레드시트로 저장되고 그래프가 스프레드시트에 직접 임베드됩니다.
- **Microsoft PowerPoint 문서.** 각 단계가 새 슬라이드에 표시됩니다. 모든 그래프는 PowerPoint 슬라이드에 직접 임베드됩니다.
- **출력 오브젝트.** IBM SPSS Modeler에서 열리면 이 파일(.cou)은 **보고서 형식 그룹**의 화면으로 출력 옵션과 같습니다.

참고: Microsoft Office 파일로 내보내려면 해당하는 애플리케이션이 설치되어 있어야 합니다.

제목. 보고서의 제목을 지정하십시오.

보고서 구조. CRISP-DM 또는 클래스를 선택하십시오. CRISP-DM 보기는 "큰 그림" 시놉시스뿐만 아니라 데이터 마이닝의 각 단계에 대한 세부사항이 있는 상태 보고서를 제공합니다. 클래스 보기는 데이터와 스트림의 내부 추적에 보다 적합한 오브젝트 기반 보기입니다.

작성자. 기본 사용자 이름이 표시되지만 이를 변경할 수 있습니다.

보고서 포함. 보고서에 오브젝트를 포함하기 위한 방법을 선택하십시오. 프로젝트 파일에 추가된 모든 항목을 포함하려면 모든 폴더 및 오브젝트를 선택하십시오. 보고서에 포함이 오브젝트 특성에 선택되어 있는지 여부를 기반으로 항목을 포함할 수도 있습니다. 또는 보고되지 않은 항목을 확인하려면 제외 표시된 항목만을 포함하도록 선택할 수도 있습니다(여기서 보고서에 포함은 선택되어 있지 않음).

선택. 이 옵션을 사용하면 보고서에서 최근 항목만을 선택하여 프로젝트 업데이트를 제공할 수 있습니다. 또는 이전 항목의 모수를 설정하여 오래되었거나 아마도 해결되지 않은 문제점을 추적할 수도 있습니다. 시간을 보고서의 모수로서 제거하려면 모든 항목을 선택하십시오.

출력순서. 다음 오브젝트 특성의 조합을 선택하여 이들을 폴더 내에 순서화할 수 있습니다.

- **유형.** 오브젝트를 유형별로 그룹화합니다.
- **이름.** 오브젝트를 문자순으로 구성합니다.
- **추가된 날짜.** 프로젝트에 추가된 날짜를 사용하여 오브젝트를 정렬합니다.

생성된 보고서 저장 및 내보내기

화면에 생성된 보고서가 새 출력 창에 표시됩니다. 보고서에 포함된 모든 그래프가 인라인 이미지로 표시됩니다.

보고서 용어

각 스트림에서 총 노드 수는 보고서 내에 나열됩니다. 숫자는 CRISP-DM 용어가 아니라 IBM SPSS Modeler 용어를 사용하는 다음 머리말 아래에 표시됩니다.

- **데이터 리더기.** 소스 노드.
- **데이터 기록기.** 내보내기 노드.
- **모델 작성기.** 작성 또는 모델링, 노드
- **모델 적용자.** 너깃으로도 알려진 생성된 모델.
- **출력 작성기.** 그래프 또는 출력 노드.
- **기타.** 프로젝트와 관련된 다른 노드. 예를 들어, 노드 팔레트의 필드 작업 탭 또는 레코드 작업 탭에서 사용 가능한 노드입니다.

보고서 저장

1. 파일 메뉴에서 저장을 클릭하십시오.
2. 파일 이름을 지정하십시오.

보고서가 출력 오브젝트로 저장됩니다.

보고서 내보내기

3. 파일 메뉴에서 내보내기와 내보낼 파일 유형을 클릭하십시오.
4. 파일 이름을 지정하십시오.

보고서는 선택한 형식으로 저장됩니다.

다음 파일 유형으로 내보낼 수 있습니다.

- HTML
- 텍스트
- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint

참고: Microsoft Office 파일로 내보내려면 해당하는 애플리케이션이 설치되어 있어야 합니다.

창 위쪽에 있는 단추를 사용하여 다음을 수행하십시오.

- 보고서를 인쇄합니다.
- 보고서를 외부 웹 브라우저에서 HTML로 봅니다.

제 15 장 IBM SPSS Modeler 사용자 정의

IBM SPSS Modeler 옵션 사용자 정의

필요에 맞게 IBM SPSS Modeler를 사용자 정의하기 위해 수행할 수 있는 많은 작업이 있습니다. 주로 이 사용자 정의는 메모리 할당, 기본 디렉토리, 사운드 및 색상 사용과 같은 특정 사용자 옵션 설정으로 이루어집니다. IBM SPSS Modeler 창 맨 아래에 있는 노드 팔레트도 사용자 정의할 수 있습니다.

IBM SPSS Modeler 옵션 설정

IBM SPSS Modeler 옵션을 사용자 정의하고 설정하는 몇 가지 방법이 있습니다.

- 도구 > 옵션 메뉴에서 시스템 옵션을 클릭하여 메모리 사용, 로케일과 같은 시스템 옵션을 설정하십시오.
- 도구 > 옵션 메뉴에서 사용자 옵션을 클릭하여 디스플레이 글꼴, 색상과 같은 사용자 옵션을 설정하십시오.
- 도구 > 옵션 메뉴에서 헬퍼 애플리케이션을 클릭하여 IBM SPSS Modeler에 대해 작업하는 애플리케이션의 위치를 지정하십시오.
- 파일 메뉴에서 디렉토리 설정 또는 서버 디렉토리 설정을 클릭하여 IBM SPSS Modeler에서 사용하는 기본 디렉토리를 지정하십시오.

일부 또는 모든 스트림에 적용되는 옵션도 설정할 수 있습니다. 자세한 정보는 [34 페이지의 『스트림을 위한 옵션 설정』](#)의 내용을 참조하십시오.

시스템 옵션

도구 > 옵션 메뉴에서 시스템 옵션을 클릭하여 IBM SPSS Modeler의 선호하는 언어 또는 로케일을 지정할 수 있습니다. 여기서 SPSS Modeler의 최대 메모리 사용량을 설정하고 자동으로 스트림을 저장하는 빈도를 지정할 수도 있습니다. 이 대화 상자에서 작성된 변경사항은 SPSS Modeler를 다시 시작하기 전까지는 적용되지 않음을 유의하십시오.

최대 메모리. IBM SPSS Modeler의 메모리 사용량의 한계를 부과하려면 선택하십시오(MB 단위). 일부 플랫폼에서는 SPSS Modeler는 해당 프로세스 크기를 자원이 제한되어 있거나 로드가 많은 컴퓨터에서 통행량을 줄이기 위해 제한합니다. 대량의 데이터를 처리하고 있으면 이로 인해 "메모리 부족" 오류가 발생할 수도 있습니다. 새 임계값을 지정하여 메모리 로드를 완화할 수 있습니다.

예를 들어, 매우 큰 의사결정 트리를 표시하려고 시도하면 메모리 오류가 발생할 수 있습니다. 이 경우에는 메모리를 최대 4096Mb로 늘리는 것이 권장됩니다. 대량의 데이터를 처리하게 되는 경우에는 메모리 허용량을 늘린 후에 SPSS Modeler를 종료한 후 명령행에서 시작하여 데이터 처리 시 최대 메모리 양을 사용하는지 확인하십시오.

명령행에서 시작하려면(SPSS Modeler가 기본 위치에 설치된 것으로 가정), 명령 프롬프트 창에서 다음을 입력하십시오.

```
C:\Program Files\IBM\SPSS\Modeler\18.3.0\bin\modeleclient.exe" -J-Xss4096M
```

시스템 로케일 사용. 이 옵션은 기본적으로 선택되어 있고 영어(미국)로 설정됩니다. 사용 가능한 언어 및 로케일의 목록에서 또 다른 언어를 지정하려면 선택 취소하십시오.

스트림 자동 저장 간격(분). SPSS Modeler에서 자동으로 스트림을 저장할 빈도를 지정하십시오. 최대값은 60분이고 최소값은 1분이며 기본값은 5분입니다.

메모리 관리

시스템 옵션 대화 상자에 지정된 **최대 메모리** 설정에 추가로 메모리 사용을 최적화할 수 있는 몇 가지 방법이 있습니다.

- 스트림 특성 대화 상자에서 **명목 필드의 최대 멤버** 옵션을 조정하십시오. 이 옵션은 필드의 측정 수준이 유형 없음이 된 후에 명목 필드의 최대 멤버 수를 지정합니다. 자세한 정보는 35 페이지의 [『스트림의 일반 옵션 설정』](#)의 내용을 참조하십시오.
- IBM SPSS Modeler가 사용 중인 메모리와 할당된 양이 표시되는 창(xxMB / xxMB) 오른쪽 하단 구석을 클릭하여 IBM SPSS Modeler가 메모리를 강제로 해제하도록 하십시오. 이 영역을 클릭하면 메모리 할당 수치가 떨어진 후에 어두운 그림자로 변합니다. 영역이 일반 색상으로 돌아가면 IBM SPSS Modeler가 가능한 모든 메모리를 해제한 것입니다.

기본 디렉토리 설정

파일 메뉴에서 **디렉토리 설정** 또는 **서버 디렉토리 설정**을 선택하여 파일 브라우저 및 출력에 사용하는 기본 디렉토리를 지정할 수 있습니다.

- 디렉토리 설정.** 이 옵션을 사용하여 작업 디렉토리를 설정할 수 있습니다. 기본 작업 디렉토리는 IBM SPSS Modeler 버전 설치 경로 또는 명령행에서 IBM SPSS Modeler를 실행하는 데 사용하는 경로를 기반으로 합니다. 로컬 모드에서 작업 디렉토리는 모든 클라이언트 측 작업 및 출력 파일(상대 경로로 참조된 경우)에 사용하는 경로입니다.
- 서버 디렉토리 설정.** 원격 서버 연결이 있을 때마다 파일 메뉴에서 서버 디렉토리 설정 옵션을 사용할 수 있습니다. 이 옵션을 사용하여 입력 또는 출력에 대해 지정된 모든 서버 파일 및 데이터 파일의 기본 디렉토리를 지정하십시오. 기본 서버 디렉토리는 \$CLEO/data입니다(여기서 \$CLEO는 서버 버전의 IBM SPSS Modeler가 설치된 디렉토리임). 명령행을 사용하는 경우 -server_directory 플래그를 modelerclient 명령행 인수와 함께 사용하여 이 기본값을 대체할 수도 있습니다.

사용자 옵션 설정

도구 > 옵션 메뉴에서 사용자 옵션을 선택하여 IBM SPSS Modeler의 일반 옵션을 설정할 수 있습니다. 이러한 옵션은 IBM SPSS Modeler에서 사용된 모든 스트림에 적용됩니다.

다음 옵션 유형은 해당하는 탭을 클릭하여 설정할 수 있습니다.

- 모델 덮어쓰기 및 오류 메시지 등과 같은 알림 옵션.
- 그래프 및 배경 색상 등과 같은 표시 옵션.
- 구문 색상 표시 옵션.
- 모델을 예측 모델 마크업 언어(PMML)로 내보낼 때 사용하는 PMML 내보내기 옵션.
- 사용자의 이름, 이니셜 및 이메일 주소 등과 같은 사용자 또는 작성자 정보. 이 정보는 노드 및 사용자가 작성하는 기타 오브젝트의 주석(Annotation) 탭에 표시될 수 있습니다.
- 기존 모드와 분석 서버 모드 간에 전환

소수점 구분자, 시간 및 데이터 형식, 최적화, 스트림 레이아웃 및 스트림 스크립트 등과 같은 스트림 특정 옵션을 설정하려면 파일 및 도구 메뉴에서 사용 가능한 스트림 특성 대화 상자를 사용하십시오.

알림 옵션 설정

사용자 옵션 대화 상자의 알림 탭을 사용하면 IBM SPSS Modeler에서 발생과 경고 유형 및 확인 창과 관련하여 다양한 옵션을 설정할 수 있습니다. 새 출력과 모델이 생성될 때 관리자 분할창에서 출력과 모델 탭의 작동을 지정할 수도 있습니다.

스트림 실행 피드백 대화 상자 표시 스트림이 3초 동안 실행된 후 진행 표시기를 포함하는 대화 상자를 표시하려면 선택하십시오. 대화 상자에는 또한 스트림이 작성한 출력 오브젝트의 세부사항도 포함됩니다.

- 완료 시 대화 상자 닫기** 기본적으로 스트림이 실행을 완료하면 대화 상자가 닫힙니다. 스트림이 마칠 때 대화 상자가 표시 상태로 남아 있게 하려면 이 상자를 지우십시오.

노드가 파일을 덮어쓸 때 경고 노드 작업이 기존 파일을 덮어쓸 때 오류 메시지와 함께 경고하려면 선택하십시오.

노드가 데이터베이스 테이블을 덮어쓸 때 경고 노드 작업이 기존 데이터베이스 테이블을 덮어쓸 때 오류 메시지와 함께 경고하려면 선택하십시오.

사운드 알림

이벤트나 오류가 발생할 때 사운드로 알릴지 여부를 지정하려면 목록을 사용하십시오. 여러 사운드를 사용할 수 있습니다. 선택한 사운드를 재생하려면 재생(확성기) 단추를 사용하십시오. 사운드를 찾아서 선택하려면 생략 기호(...)를 사용하십시오.

참고: IBM SPSS Modeler에서 사운드를 작성하는 데 사용하는 .wav 파일은 설치의 /media/sounds 디렉토리에 저장됩니다.

- 모든 사운드 음소거 모든 이벤트의 사운드 알림을 끄려면 선택하십시오.

시각적 알림

이 그룹의 옵션은 새 항목이 생성될 때 표시의 오른쪽 상단에 있는 관리자 분할창의 출력 및 모델 탭의 작동을 지정하는 데 사용합니다. 해당 탭의 작동을 지정하려면 목록에서 새 모델 또는 새 출력을 선택하십시오.

다음 옵션을 새 모델에 사용 가능합니다.

이전 모델 바꾸기 선택된 경우(기본값) 모델 탭의 이 스트림과 스트림 캔버스에서 기존 모델을 덮어씁니다. 이 상자가 선택 해제되면 모델은 탭과 캔버스에서 기존 모델에 추가됩니다. 이 설정은 모델 링크의 모델 대체 설정으로 대체됨을 유의하십시오.

다음 옵션을 새 출력에 사용 가능합니다.

출력이 [n] 초과 시 경고 출력 탭의 항목 수가 사전에 지정된 양을 초과할 때 경고를 표시할지 여부를 선택하십시오. 기본 양은 20입니다. 그러나 필요한 경우 이를 변경할 수 있습니다.

다음 옵션은 모든 경우에 사용 가능합니다.

선택 탭 스트림이 실행되는 동안 해당하는 오브젝트가 생성될 때 출력 또는 모델 탭으로 전환할지 여부를 선택합니다.

- 관리자 분할창에서 해당하는 탭으로 전환하려면 **항상**을 선택하십시오.
- 캔버스에서 현재 표시되는 스트림이 생성한 오브젝트에 대해서만 해당 탭으로 전환하려면 **현재 스트림에 의해 생성된 경우**를 선택하십시오.
- 소프트웨어가 생성된 출력이나 모델에 대해 사용자에게 알리기 위해 해당 탭으로 전환되는 것을 막으려면 **사용 안 함**을 선택하십시오.

플래시 탭 새 출력이나 모델이 생성될 때 관리자 분할창에서 출력이나 모델 탭을 플래시할지 여부를 선택하십시오.

- 관리자 분할창에서 새 오브젝트가 생성될 때마다 해당 탭을 플래시하려면(아직 선택되지 않은 경우) **선택되지 않은 경우**를 선택하십시오.
- 소프트웨어가 생성된 오브젝트에 대해 사용자에게 알리기 위해 해당 탭을 플래시하는 것을 막으려면 **사용 안 함**을 선택하십시오.

팔레트를 스크롤하여 표시(새 모델에만 해당). 관리자 분할창에서 가장 최신 모델을 표시하기 위해 모델 탭을 자동으로 스크롤하는지 여부를 선택하십시오.

- 스크롤링을 사용 설정하려면 **항상**을 선택하십시오.
- 캔버스에서 현재 표시되는 스트림이 생성한 오브젝트에 대해서만 스크롤하려면 **현재 스트림이 생성한 경우**를 선택하십시오.
- 소프트웨어가 모델 탭을 자동으로 스크롤링하는 것을 막으려면 **사용 안 함**을 선택하십시오.

창 열기(새 출력에만 해당). 생성 시 출력 창을 자동으로 열지 여부를 선택하십시오.

- 새 출력 창을 항상 열려면 **항상**을 선택하십시오.
- 캔버스에서 현재 표시되는 스트림이 생성한 출력에 대해 새 창을 열려면 **현재 스트림이 생성한 경우**를 선택하십시오.
- 소프트웨어가 생성된 출력에 대해 자동으로 새 창을 여는 것을 막으려면 **사용 안 함**을 선택하십시오.

이 탭의 시스템 기본 설정으로 돌아가려면 **기본값**을 클릭하십시오.

표시 옵션 설정

사용자 옵션 대화 상자의 표시 탭을 사용하면 IBM SPSS Modeler에서 글꼴과 색상의 표시에 옵션을 설정할 수 있습니다.

시작할 때 시작 대화 상자 표시: 시작 시 시작 대화 상자를 표시하려면 선택하십시오. 시작 대화 상자에는 애플리케이션 예제 자습서를 실행하고, 시연 스트림이나 기존 스트림 또는 프로젝트를 열거나 새 스트림을 작성하기 위한 옵션이 있습니다.

스트림 및 수퍼노드 마크업 보이기: 선택할 경우 스트림 및 수퍼노드에 대한 마크업(있는 경우)이 기본적으로 표시됩니다. 마크업에는 스트림 주석, 모델 링크 및 스코어링 분기의 강조 표시가 포함됩니다.

표준 글꼴 및 색상(재시작 시 적용): 이 제어 상자의 옵션은 IBM SPSS Modeler 화면 디자인, 색 구성표 및 표시되는 글꼴의 크기를 지정하는 데 사용됩니다. 여기에서 선택한 옵션은 IBM SPSS Modeler를 닫고 다시 시작할 때 까지 적용되지 않습니다.

- **모양과 느낌:** 표준 색 구성표와 화면 디자인을 선택하십시오. 다음에서 선택할 수 있습니다.
 - **SPSS Standard** - 기본 디자인입니다.
 - **SPSS Classic** - SPSS Modeler의 이전 버전 사용자에게 친숙한 디자인.
 - **Windows** - 스트림 캔버스와 팔레트에서 대비 증가에 사용할 수 있는 Windows 디자인.
 - **Analytics Carbon** - 매끈한 아이콘과 색상을 사용한 모던 디자인입니다.
- **노드의 기본 글꼴 크기:** 노드 팔레트 및 스트림 캔버스에 표시되는 노드에 사용할 글꼴 크기를 지정하십시오.
- **고정 너비 글꼴 지정:** 고정 너비 글꼴 및 스크립팅과 CLEM 표현식 제어에 사용할 연관된 글꼴 크기를 선택하려면 이 확인 상자를 선택하십시오. 기본 글꼴은 Monospace plain입니다. 선택할 수 있는 다른 글꼴 목록을 표시하려면 **변경...**을 클릭하십시오.

참고: 스트림 특성 대화 상자의 옵션 탭의 레이아웃 분할창에서 스트림의 노드 아이콘 크기를 설정할 수 있습니다. 기본 메뉴에서 **도구 > 스트림 특성 > 옵션 > 레이아웃**을 선택하십시오.

사용자 정의 색상: 이 테이블에는 다양한 표시 항목에 사용되는 현재 선택된 색상이 나열됩니다. 테이블에 나열된 각 항목마다 색상 열에서 해당하는 행을 두 번 클릭하고 목록에서 색상을 선택하여 현재 색상을 변경할 수 있습니다. 사용자 정의 색상을 지정하려면 목록에서 맨 아래로 스크롤하고 색상... 항목을 클릭하십시오.

차트 범주 색상 순서: 이 테이블에는 새로 작성된 그래프에 표시하는 데 사용되는 현재 선택된 색상이 나열됩니다. 색상의 순서는 차트에 사용하는 순서를 반영합니다. 예를 들어, 색상 오버레이로 사용하는 명목 필드에 4개의 고유한 값이 포함되는 경우에는 여기에 나열되는 처음 4개의 색상만을 사용합니다. 테이블에 나열된 각 항목마다 색상 열에서 해당하는 행을 두 번 클릭하고 목록에서 색상을 선택하여 현재 색상을 변경할 수 있습니다. 사용자 정의 색상을 지정하려면 목록에서 맨 아래로 스크롤하고 색상... 항목을 클릭하십시오. 여기에서 작성한 변경 사항은 이전에 작성한 그래프에는 영향을 미치지 않습니다.

이 탭의 시스템 기본 설정으로 돌아가려면 **기본값**을 클릭하십시오.

구문 표시 옵션 설정

사용자 옵션 대화 상자의 구문 탭을 사용하면 IBM SPSS Modeler에서 작성하는 스크립트에서 글꼴 속성과 표시 색상의 옵션을 설정할 수 있습니다.

구문 강조 표시. 이 테이블은 글꼴과 표시 창 둘 모두를 포함하여 다양한 구문 항목에 사용하는 현재 선택된 색상을 나열합니다. 테이블에 나열되는 각 항목에 대해 행에서 해당하는 드롭 다운 목록을 클릭하고 목록에서 색상을 선택하여 색상을 변경할 수 있습니다. 또한 글꼴 항목의 경우 굵은체와 기울임꼴 강조를 추가할 수 있습니다.

미리보기. 이 테이블은 구문 강조 표시 테이블에서 선택한 색상 및 글꼴 속성을 사용하는 예제 구문 표시를 보여 줍니다. 이 미리보기는 선택을 변경하는 즉시 업데이트합니다.

이 탭의 시스템 기본 설정으로 돌아가려면 **기본값**을 클릭하십시오.

PMML 내보내기 옵션 설정

PMML 탭에서 IBM SPSS Modeler가 모델을 예측 모델 마크업 언어(PMML)로 내보내는 방법을 제어할 수 있습니다. 자세한 정보는 193 페이지의 『모델을 PMML로 가져오기 및 내보내기』의 내용을 참조하십시오.

PMML 내보내기. 여기에서 목표 애플리케이션과 가장 잘 작동하는 PMML의 변형을 구성할 수 있습니다.

- 표준 PMML에 상당하는 표준이 없는 특수한 경우 PMML 확장을 허용하려면 확장 사용을 선택하십시오. 대부분의 경우 이는 표준 PMML과 동일한 결과를 생성함을 유의하십시오.
- PMML 표준을 가능한 가장 밀접하게 고수하는 PMML을 내보내려면 표준 PMML로...를 선택하십시오.

표준 PMML 옵션. 표준 PMML로... 옵션이 선택된 경우에는 선형 및 로지스틱 회기 모델을 내보내기 위한 두 개의 유효한 방법 중 하나를 선택할 수 있습니다.

- PMML <GeneralRegression> 모델로
- PMML <Regression> 모델로

PMML에 대한 자세한 정보는 데이터 마이닝 그룹 웹 사이트(<http://www.dmg.org>)를 참조하십시오.

사용자 정보 설정

사용자/작성자 정보. 여기에 입력하는 정보는 사용자가 작성하는 노드와 기타 오브젝트의 주석(Annotation) 탭에 표시할 수 있습니다.

모드 설정

Modeler 모드 설정. 모드 탭의 다음 모드 중에서 선택할 수 있습니다.

- 기존 **SPSS Modeler** 모드에서는 사용 가능한 모든 노드와 표현식을 사용자 인터페이스에 표시합니다.
- Analytic Server** 모드에서는 Analytic Server에서 지원하는 노드와 표현식만 표시합니다. Analytic Server에서 완전히 지원되지 않아도 여전히 일부 노드와 일부 CLEM이 표시됩니다. 다음 표에서는 Analytic Server에서 지원하는 노드, 부분적으로 지원하는 노드, 지원하지 않는 노드에 대한 일반 정보를 제공합니다.

Analytic Server에 대한 자세한 정보는 [Analytic Server](#) 문서를 참조하십시오.

SPSS Modeler가 데이터베이스 모델링 팔레트에 데이터베이스 노드를 표시하도록 구성된 경우, 모드 전환 시 영향을 받지 않습니다. 데이터베이스 노드는 항상 표시됩니다. z/OS용 IBM Db2 또는 IBM Netezza 통합을 사용할 때 Analytic Server 모드로 전환된 후에 해당 노드가 데이터베이스 모델링 팔레트에서 사라지는 경우가 있을 수 있습니다. 이런 경우, 도구 > 옵션 > 헬퍼 애플리케이션으로 이동하여 확인 상자를 재설정하십시오.

표 38. 노드 지원

노드 유형(팔레트 이름)	Analytic Server에서 지원	Analytic Server에서 부분적으로 지원	Analytic Server에서 지원하지 않음
소스	<ul style="list-style-type: none"> Analytic Server 소스 노드 		<ul style="list-style-type: none"> 데이터베이스 가변 파일 고정 파일 통계 파일 데이터 수집 IBM Cognos TWC 가져오기 TM1 가져오기 SAS 파일 Excel XML 사용자 입력 시뮬레이션 생성 데이터 보기 지리 공간적 오브젝트 스토리지 확장 가져오기 R 가져오기 SNA(분산 분석 및 그룹 분석)
레코드 작업	<ul style="list-style-type: none"> 선택 정렬 균형 구분 RFM 통합 붙여쓰기 TS 스트리밍 확장 변환 TCM 스트리밍 	<ul style="list-style-type: none"> 샘플(단순 방법에만 Random%를 지원하고, 복합은 지원되지 않습니다.) 병합(키와 조건별 결합 만 지원) 통합(1분위수, 3분위수 및 중앙값은 Analytic Server에서 지원하지 않음) 	<ul style="list-style-type: none"> Space-Time-Boxes CPLEX 최적화

표 38. 노드 지원 (계속)

노드 유형(팔레트 이름)	Analytic Server에서 지원	Analytic Server에서 부분적으로 지원	Analytic Server에서 지원하지 않음
필드 작업	<ul style="list-style-type: none"> • 유형 • 필터 • 파생 • 채움 • 재분류 • 양상블 • 이분변환 • 구조변환 • 필드 다시 정렬 • 재투영 • 시간 간격 	<ul style="list-style-type: none"> • 자동 데이터 준비(변환만 지원) • 구간화(경계값 설정 최신으로 유지를 선택하면 구간화 방법 equalFreq가 지원되지 않음) • RFM 분석(경계값 설정 최신으로 유지를 선택하면 구간화 방법 Tiles가 지원되지 않음) • 파티션(파티션에 반복적으로 행을 지정하기 위해 고유 필드를 사용하지 않으면 Analytic Server에서 지원하지 않음) 	<ul style="list-style-type: none"> • 값 익명화 • 히스토리 • 전치
그래프	<ul style="list-style-type: none"> • 도표 • 다중 도표 • 시간 도표 • 분포 • 히스토그램 • 요약도표 • 웹 • 평가 • 맵 시각화 • E-Plot(베타) • t-SNE 	<ul style="list-style-type: none"> • 그래프 보드(측정 수준이 이산, 명목, 순서 또는 플래그인 필드의 통합 모드 기능만 지원) 	

표 38. 노드 지원 (계속)

노드 유형(팔레트 이름)	Analytic Server에서 지원	Analytic Server에서 부분적으로 지원	Analytic Server에서 지원하지 않음
모델링	<ul style="list-style-type: none"> 시계열 TCM 등위-AS 임의 트리 Tree-AS Linear-AS GLE LSVM STP TwoStep-AS 연관 규칙 XGBoost-AS K-평균-AS 	<ul style="list-style-type: none"> 자동 분류자 자동 숫자(두 노드에서 분할만 지원하며, 자동 분류자 옵션 Analytic Server에서 실행(분할 사용)을 사용할 경우 분할 역할이 있는 필드만 제공해야 함) 확장(R 구문 작성 모델은 Analytic Server에서 지원되지 않음) 다음 노드에서는 분할과 PSM만 지원합니다. <ul style="list-style-type: none"> - C&R 트리 - 선형 - 신경망 - CHAID - Quest 	<ul style="list-style-type: none"> 자동 군집 의사결정 목록 C5.0 회귀분석 PCA/요인 필드선택 차별 로지스틱 GenLin GLMM Bayes Net Apriori Carma 순차규칙 K-평균 코호넨 TwoStep Anomaly KNN R 임의 포리스트 다음 노드에는 asl이 있지만 읽기-쓰기 asl뿐입니다. <ul style="list-style-type: none"> - Cox - SVM - SLRM
출력	<ul style="list-style-type: none"> 테이블 교차표 분석 데이터 검토 변환 통계량 평균 보고서 글로벌 설정 	<ul style="list-style-type: none"> 확장 출력 R 구문 출력 	<ul style="list-style-type: none"> 시뮬레이션 평가 시뮬레이션 적합 R 출력

표 38. 노드 지원 (계속)

노드 유형(팔레트 이름)	Analytic Server에서 지원	Analytic Server에서 부분적으로 지원	Analytic Server에서 지원하지 않음
내보내기	<ul style="list-style-type: none"> Analytic Server 내보내기 노드 	<ul style="list-style-type: none"> 확장 내보내기 R 구문 내보내기 	<ul style="list-style-type: none"> 데이터베이스 플랫 파일 통계 내보내기 데이터 수집 Excel IBM Cognos 내보내기 TM1 내보내기 SAS XML 내보내기 오브젝트 스토리지 R 내보내기
IBM SPSS Statistics			<ul style="list-style-type: none"> 통계 파일 통계 변환 통계 모델 통계 출력 통계 내보내기
IBM SPSS Text Analytics	<ul style="list-style-type: none"> 텍스트 링크 분석 텍스트 마이닝 언어 노드 		<ul style="list-style-type: none"> 파일 목록 웹 피드 텍스트 링크 분석 변환 텍스트 마이닝 파일 뷰어
Python			<ul style="list-style-type: none"> SMOTE One-Class SVM XGBoost Tree XGBoost Linear t-SNE 임의 포리스트 HDBSCAN
Spark	모두 지원됨		

노드 팔레트 사용자 정의

스트림은 노드를 사용하여 작성됩니다. IBM SPSS Modeler 창 맨 아래의 노드 팔레트는 스트림 작성에서 사용 가능한 모든 노드를 포함합니다. 자세한 정보는 [12 페이지의 『노드 팔레트』](#)의 내용을 참조하십시오.

다음 두 가지 방법으로 노드 팔레트를 재구성할 수 있습니다.

- 팔레트 관리자 사용자 정의. 자세한 정보는 [214 페이지의 『팔레트 관리자 사용자 정의』](#)의 내용을 참조하십시오.

- 하위 팔레트가 포함된 팔레트 탭이 노드 팔레트에 표시되는 방법 변경. 자세한 정보는 [215 페이지의 『하위 팔레트 작성』](#)의 내용을 참조하십시오.

팔레트 관리자 사용자 정의

팔레트 관리자는 IBM SPSS Modeler의 사용법에 맞춰 사용자 정의할 수 있습니다. 예를 들어, 데이터베이스에서 시계열 데이터를 자주 분석하는 경우에는 데이터베이스 소스 노드, 시간 간격 노드, 시계열 노드 및 시간 도표 그래프 노드가 함께 고유한 팔레트 탭에서 사용 가능한지 확인하려고 할 수 있습니다. 팔레트 관리자를 사용하면 노드 팔레트에서 사용자 정의 팔레트 탭을 작성하여 이러한 조정을 쉽게 할 수 있습니다.

팔레트 관리자를 사용하면 다양한 작업을 수행할 수 있습니다.

- 어떤 팔레트 탭이 스트림 캔버스 아래 노드 팔레트에 표시되는지를 제어합니다.
- 노드 팔레트에 팔레트 탭이 표시되는 순서를 변경합니다.
- 사용자 고유의 팔레트 탭 및 연관된 하위 팔레트를 작성하고 편집합니다.
- 탭에서 기본 노드 선택을 편집합니다.

팔레트 관리자에 액세스하려면 도구 메뉴에서 **팔레트 관리**를 클릭하십시오.

팔레트 이름. 노드 팔레트에 표시되는지 여부와 관계없이 사용 가능한 각 팔레트 탭이 나열됩니다. 여기에는 사용자가 작성한 모든 팔레트 탭이 포함됩니다. 자세한 정보는 [214 페이지의 『팔레트 탭 작성』](#)의 내용을 참조하십시오.

노드 수. 각 팔레트 탭에 표시되는 노드 수입니다. 숫자가 높을 수록 탭에서 노드를 나누는 하위 팔레트를 작성하는 것이 보다 편리할 수 있습니다. 자세한 정보는 [215 페이지의 『하위 팔레트 작성』](#)의 내용을 참조하십시오.

표시 여부 노드 팔레트에서 팔레트 탭을 표시하려면 이 필드를 선택하십시오. 자세한 정보는 [214 페이지의 『노드 팔레트에 팔레트 탭 표시』](#)의 내용을 참조하십시오.

하위 팔레트. 팔레트 탭에 표시할 하위 팔레트를 선택하려면 필요한 **팔레트 이름**을 강조 표시하고 이 단추를 클릭하여 하위 팔레트 대화 상자를 표시하십시오. 자세한 정보는 [215 페이지의 『하위 팔레트 작성』](#)의 내용을 참조하십시오.

기본값 복원. 팔레트와 하위 팔레트에 작성한 모든 변경사항과 추가사항을 완전히 제거하고 기본 팔레트 설정으로 돌아가려면 이 단추를 클릭하십시오.

팔레트 탭 작성

사용자 정의 팔레트 탭 작성하기

- 도구 메뉴에서 팔레트 관리자를 여십시오.
 - 표시? 열의 오른쪽에서 팔레트 추가 단추를 클릭하십시오. 팔레트 작성/편집 대화 상자가 표시됩니다.
 - 고유한 팔레트 이름을 입력하십시오.
 - 사용 가능한 노드 영역에서 팔레트 탭에 추가할 노드를 선택하십시오.
 - 노드 추가 오른쪽 화살표 단추를 클릭하여 강조 표시된 노드를 선택 노드 영역으로 이동하십시오. 원하는 모든 노드를 추가할 때까지 반복하십시오.
- 필요한 모든 노드를 추가한 후에는 팔레트 탭에 노드가 표시되는 순서를 변경할 수 있습니다.
- 단순 화살표 단추를 사용하여 노드를 한 행 위나 아래로 이동하십시오.
 - 선 화살표 단추를 사용하여 노드를 목록의 맨 아래나 맨 위로 이동하십시오.
 - 팔레트에서 노드를 제거하려면 노드를 강조 표시하고 선택 노드 영역 오른쪽의 삭제 단추를 클릭하십시오.

노드 팔레트에 팔레트 탭 표시

IBM SPSS Modeler 내에 사용자가 결코 사용하지 않는 옵션이 사용 가능할 수 있습니다. 이 경우 팔레트 관리자를 사용하여 이러한 노드를 포함하는 탭을 숨기면 됩니다.

노드 팔레트에 어떤 탭이 표시되는지 선택

- 도구 메뉴에서 팔레트 관리자를 여십시오.

2. 표시? 열의 확인 상자를 사용하여 각 팔레트 탭을 포함하거나 숨길지 여부를 선택합니다.

노드 팔레트에서 팔레트 탭을 영구적으로 제거하려면 노드를 강조 표시하고 표시? 열 오른쪽의 삭제 단추를 클릭하십시오. 팔레트 탭은 한 번 삭제하면 복구할 수 없습니다.

참고: 즐겨찾기 탭을 제외하고는 IBM SPSS Modeler와 함께 제공된 기본 팔레트 탭은 삭제할 수 없습니다.

노드 팔레트에서 표시 순서 변경

표시할 팔레트 탭을 선택한 후에는 노드 팔레트에 표시되는 순서를 변경할 수 있습니다.

1. 단순 화살표 단추를 사용하여 팔레트 탭을 한 행씩 위나 아래로 이동하십시오. 위로 이동하면 노드 팔레트의 왼쪽으로 이동되고, 역 또한 같습니다.
2. 선 화살표 단추를 사용하여 팔레트 탭을 목록의 맨 아래나 맨 위로 이동하십시오. 목록의 맨 위에 있는 팔레트는 노드 팔레트의 왼쪽에 표시됩니다.

팔레트 탭에 하위 팔레트 표시

같은 방법으로 노드 팔레트에 어떤 팔레트 탭이 표시되는지를 제어하고, 해당 상위 팔레트 탭에서 어떤 하위 팔레트를 사용 가능한지를 제어할 수 있습니다.

팔레트 탭에 표시할 하위 팔레트 선택

1. 도구 메뉴에서 팔레트 관리자를 여십시오.
2. 필요한 팔레트를 선택하십시오.
3. 하위 팔레트 단추를 클릭하면 하위 팔레트 대화 상자가 표시됩니다.
4. 표시? 열의 확인 상자를 사용하여 각 하위 팔레트를 팔레트 탭에 포함할지 여부를 선택합니다. **모두** 하위 팔레트는 항상 표시되고 삭제할 수 없습니다.
5. 노드 팔레트에서 하위 팔레트를 영구적으로 제거하려면 하위 팔레트를 강조 표시하고 표시? 열 오른쪽의 삭제 단추를 클릭하십시오.

참고: 모델링 팔레트 탭과 함께 제공된 기본 하위 팔레트는 삭제할 수 없습니다.

팔레트 탭에서 표시 순서 변경

표시할 하위 팔레트를 선택한 후에는 상위 팔레트 탭에 표시되는 순서를 변경할 수 있습니다.

1. 단순 화살표 단추를 사용하여 하위 팔레트를 한 행 위나 아래로 이동하십시오.
2. 선 화살표 단추를 사용하여 하위 팔레트를 목록의 맨 아래나 맨 위로 이동하십시오.

사용자가 작성하는 하위 팔레트는 상위 팔레트 탭을 선택할 때 노드 팔레트에 표시됩니다. 자세한 정보는 [216 페이지의 『팔레트 탭 보기 변경』](#)의 내용을 참조하십시오.

하위 팔레트 작성

기존 노드를 사용자가 작성하는 사용자 정의 팔레트 탭에 추가할 수 있으므로 스크롤링하지 않고도 화면에 쉽게 표시할 수 있는 노드보다 많은 노드를 선택하는 것이 가능합니다. 스크롤링 할 필요를 막기 위해서는 팔레트 탭에서 선택한 노드를 배치할 하위 팔레트를 작성하면 됩니다. 예를 들어, 스트림을 작성하기 위해 자주 사용하는 노드를 포함하는 팔레트 탭을 작성한 경우에는 선택사항을 소스 노드, 필드 작업, 모델링 및 출력으로 구분하는 4개의 하위 팔레트를 작성할 수 있습니다.

참고: 상위 팔레트 탭에 추가한 하위 팔레트 노드에서만 선택할 수 있습니다.

하위 팔레트 작성

1. 도구 메뉴에서 팔레트 관리자를 여십시오.
2. 하위 팔레트를 추가하려는 팔레트를 선택하십시오.
3. 하위 팔레트 단추를 클릭하면 하위 팔레트 대화 상자가 표시됩니다.
4. 표시? 열의 오른쪽에서 하위 팔레트 추가 단추를 클릭하십시오. 하위 팔레트 작성/편집 대화 상자가 표시됩니다.
5. 고유한 하위 팔레트 이름을 입력하십시오.
6. 사용 가능한 노드 영역에서 하위 팔레트 탭에 추가할 노드를 선택하십시오.

7. 노드 추가 오른쪽 화살표 단추를 클릭하여 선택한 노드를 선택 노드 영역으로 이동하십시오.
 8. 필수 노드를 추가한 후 확인을 클릭하여 하위 팔레트 대화 상자로 돌아가십시오.
- 사용자가 작성하는 하위 팔레트는 상위 팔레트 탭을 선택할 때 노드 팔레트에 표시됩니다. 자세한 정보는 [216 페이지의 『팔레트 탭 보기 변경』](#)의 내용을 참조하십시오.

팔레트 탭 보기 변경

IBM SPSS Modeler에서 사용 가능한 노드가 아주 많기 때문에, 노드 팔레트의 왼쪽 또는 오른쪽으로 스크롤하지 않으면 작은 화면에 노드가 모두 표시되지는 않을 수 있습니다. 모델링 팔레트 탭에서는 특히 그렇습니다. 스크롤 필요성을 줄이기 위해 하위 팔레트(사용 가능한 경우)에 포함된 노드만 표시하도록 선택할 수 있습니다. 자세한 정보는 [215 페이지의 『하위 팔레트 작성』](#)의 내용을 참조하십시오.

팔레트 탭에 표시된 노드를 변경하려면 팔레트 탭을 선택한 후 왼쪽 메뉴에서 모든 노드 또는 특정 하위 팔레트의 해당 노드만 표시하도록 선택하십시오.

제 16 장 스트림과 노드의 성능 고려사항

노드를 가장 효율적인 구성으로 배열하고 해당하는 경우 노드 캐시를 사용 가능으로 설정하고, 이 섹션에 설명된 다른 고려사항에 유의하여 성능을 최대화하도록 스트림을 계획할 수 있습니다.

여기에서 설명된 고려사항은 제외하고, 추가적 및 보다 실질적인 성능 개선사항은 일반적으로 특히 SQL 최적화를 통해 데이터베이스를 효율적으로 사용함으로써 얻을 수 있습니다.

노드의 순서

SQL 최적화를 사용 중이 아니더라도 스트림에서 노드의 순서는 성능에 영향을 미칩니다. 일반 목표는 다운스트림 처리를 최소화하는 것입니다. 그러므로 데이터 양을 줄이는 노드가 있으면 이를 스트림의 시작 부분 근처에 두십시오. IBM SPSS Modeler Server는 컴파일링 중에 몇몇 순서 조정 규칙을 자동으로 적용하여 특정 노드를 앞으로 가져와도 안전한 것으로 증명할 수 있을 때는 그렇게 합니다. (이 기능은 기본적으로 사용 안함으로 설정됩니다. 시스템 관리자에게 문의하여 설치에서 사용 가능한지 확인하십시오.)

SQL 최적화를 사용할 때 가용성과 효율을 극대화하려고 합니다. 최적화는 스트림에 데이터베이스에서 수행할 수 없는 작업이 포함된 경우 정지되므로 SQL 최적화된 작업은 스트림의 시작 부분에 함께 그룹화하는 것이 좋습니다. 이 전략은 데이터베이스에서 더 많은 처리를 보존하므로 IBM SPSS Modeler에 전달되는 데이터 양은 줄어듭니다.

다음 작업은 대부분의 데이터베이스에서 수행할 수 있습니다. 스트림의 시작 부분에 이들을 그룹화하십시오.

- 키를 기준으로 병합(결합)
- 선택
- 통합
- 정렬
- 샘플
- 붙여쓰기
- 모든 필드가 선택된 포함 모드에서 고유 작업
- 채움 작업
- (데이터베이스가 어떤 작업을 지원하는지에 따라) 표준 산술 또는 문자열 조작을 사용하는 기본 파생 작업
- 설정 플래그

다음 작업은 대부분의 데이터베이스에서 수행할 수 없습니다. 이들은 앞의 목록에서 작업 후 스트림에 배치해야 합니다.

- 플랫 파일 등과 같은 비데이터베이스 데이터에 대한 작업
- 순서별 병합
- 균형
- 삭제 모드 또는 필드의 서브세트만이 고유로 선택된 곳에서 고유 작업
- 처리 중인 레코드가 아닌 레코드에서 데이터에 액세스해야 하는 모든 작업
- 상태 및 개수 필드 파생
- 히스토리 노드 작업
- "@"(시계열) 함수와 관련된 작업
- 유형 확인 모드 경고 및 중단
- 모델 구성, 애플리케이션 및 분석

참고: 의사결정 트리, 규칙세트, 선형 회귀 및 요소 생성 모델은 SQL을 생성할 수 있으므로 데이터베이스로 다시 넣을 수 있습니다.

- 데이터를 처리 중인 동일 데이터베이스가 아닌 다른 데이터베이스로 데이터 출력

노드 캐시

실행 중인 스트림을 최적화하기 위해 비터미널 노드에서 캐시를 설정할 수 있습니다. 노드에서 캐시를 설정하면 다음 번에 데이터 스트림을 실행할 때 노드를 통과하는 데이터로 캐시가 채워집니다. 그때부터는 데이터 소스가 아닌 캐시(디스크의 임시 디렉토리에 저장됨)에서 데이터를 읽습니다.

캐싱은 정렬, 병합 또는 통합과 같이 시간이 걸리는 작업 다음으로 가장 유용합니다. 예를 들어, 데이터베이스에서 판매 데이터를 읽는 소스 노드 세트와 지역별로 판매액을 요약하는 통합 노드가 있다고 가정하십시오. 전체 데이터 세트가 아니라 통합된 데이터를 저장할 캐시를 원하기 때문에 소스 노드가 아닌 통합 노드에서 캐시를 설정할 수 있습니다.

참고: 소스 노드에서 캐싱하면(IBM SPSS Modeler로 데이터를 읽을 때 원 데이터 세트의 사본만 저장함) 대부분의 상황에서 성능이 향상되지는 않습니다.

캐싱을 사용하는 노드에는 오른쪽 상단 모서리에 작은 문서 아이콘이 표시됩니다. 노드에서 데이터가 캐싱될 때 문서 아이콘은 녹색입니다.

캐시 사용 방법

1. 스트림 캔버스에서 노드를 마우스 오른쪽 단추로 클릭하고 메뉴에서 캐시를 클릭하십시오.
2. 캐싱 하위 메뉴에서 사용을 클릭하십시오.
3. 노드를 마우스 오른쪽 단추로 클릭하고 캐싱 하위 메뉴에서 사용 안함을 클릭하여 캐시를 설정 해제할 수 있습니다.

데이터베이스의 노드 캐싱

데이터베이스에서 실행되는 스트림의 경우, 스트림 중간에 파일 시스템이 아닌 데이터베이스의 임시 테이블에 데이터를 캐싱할 수 있습니다. SQL 최적화와 결합되면 이로 인해 성능이 상당히 향상됩니다. 예를 들어, 다중 테이블을 병합하여 데이터 마이닝 보기を作성하는 스트림의 출력을 캐싱하고 필요에 따라 재사용할 수 있습니다. 모든 다운스트림 노드에 대해 SQL을 자동으로 생성하면 성능이 추가로 향상될 수 있습니다.

데이터베이스 캐싱을 이용하려면 SQL 최적화와 데이터베이스 캐싱 둘 다를 사용으로 설정해야 합니다. 서버 최적화 설정은 클라이언트에서 해당 설정을 대체합니다. 자세한 정보는 [37 페이지의 『스트림의 최적화 옵션 설정』](#)의 내용을 참조하십시오.

데이터베이스 캐싱을 사용으로 설정한 상태에서는 비터미널 노드를 마우스 오른쪽 단추로 클릭하여 해당 지점에서 데이터를 캐싱하면 되며, 다음 번에 스트림이 실행될 때 데이터베이스에 자동으로 캐시가 직접 작성됩니다. 데이터베이스 캐싱 또는 SQL 최적화를 사용할 수 없으면 캐시는 파일 시스템에 대신 기록됩니다.

참고: Db2, Oracle, SQL Server 및 Teradata에서는 캐싱 목적으로 임시 테이블을 지원합니다. Netezza와 같은 기타 데이터베이스는 데이터베이스 캐싱에 정규 테이블을 사용합니다. 특정 데이터베이스에 대해 SQL 코드를 사용자 정의할 수 있습니다. 지원을 받으려면 서비스팀에 문의하십시오.

성능: 프로세스 노드

정렬. 정렬 노드는 정렬하려면 먼저 전체 입력 데이터 세트를 읽어야 합니다. 데이터는 일부 한계까지 메모리에 저장되고 초과분은 디스크로 유출됩니다. 정렬 알고리즘은 조합 알고리즘입니다. 데이터는 한계까지 메모리로 읽혀지고 빠른 하이브리드 빠른 정렬 알고리즘을 사용하여 정렬됩니다. 모든 데이터가 메모리에 맞는 경우에는 정렬이 완료됩니다. 그렇지 않은 경우에는 병합-정렬 알고리즘이 적용됩니다. 정렬된 데이터가 파일에 작성되고 데이터의 다음 청크가 메모리에 읽혀지고, 정렬된 후 디스크에 작성됩니다. 이는 모든 데이터가 읽혀질 때까지 반복된 다음 정렬된 청크가 병합됩니다. 병합하려면 디스크에 저장된 데이터 위에서 반복적으로 전달해야 합니다. 최고 사용량 시에, 정렬 노드에는 디스크에 두 개의 전체 데이터 사본이 생깁니다(정렬 및 정렬되지 않음).

알고리즘의 전체 실행 시간은 대략 $N \times \log(N)$ 입니다. 여기서 N 은 레코드 수입니다. 메모리에서 정렬하는 것은 디스크에서 병합하는 것보다 빠르므로 실제 실행 시간은 정렬에 더 많은 메모리를 할당하여 줄일 수 있습니다. 알고리즘은 IBM SPSS Modeler Server 구성 옵션 메모리 사용 승수에 의해 제어되는 물리적 RAM의 일부를 자신에게 할당합니다. 정렬에 사용된 메모리를 늘리기 위해서는 더 많은 물리적 RAM을 제공하거나 이 값을 늘리십시오. 사용한 메모리의 비율이 프로세스의 작업 세트를 초과하면 메모리의 일부가 디스크에 페이지징될 수 있도록 성능이 저하됨을 유의하십시오. 인메모리 정렬 알고리즘의 메모리 액세스 패턴은 임의적이고 과도한 페이지징을 야기할 수 있기 때문입니다. 정렬 알고리즘은 정렬 노드가 아닌 몇몇 노드에서 사용하지만 동일한 성능 규칙이 적용됩니다.

구간화. 구간화 노드는 구간에 레코드를 할당하기 전에 구간 경계를 계산하기 위해 전체 입력 데이터 세트를 읽습니다. 경계가 계산되는 동안 데이터 세트가 캐싱됩니다. 그런 다음 할당을 위해 다시 스캔됩니다. 구간화 방법이 고정 너비이거나 평균+표준 편차이면 데이터 세트는 디스크에 직접 캐싱됩니다. 이러한 방법에는 선형 실행 시간이 있고 전체 데이터 세트를 저장하기 위한 충분한 디스크 공간이 필요합니다. 구간화 방법이 순위이거나 분위수이면 데이터 세트는 앞서 설명된 정렬 알고리즘을 사용하여 정렬되고 정렬된 데이터 세트는 캐시로 사용합니다. 정렬은 이러한 방법에 $M \times N \times \log(N)$ 실행 시간을 제공합니다. 여기서 M 은 구간화 필드의 수이고 N 은 레코드 수입니다. 데이터 세트 크기의 두 배에 해당하는 디스크 공간이 필요합니다.

생성된 구간을 기반으로 파생 노드를 생성하면 후속 전달 시에 성능이 향상됩니다. 파생 작업은 구간화보다 훨씬 빠릅니다.

키별로 병합(결합). 병합 방법이 키이면(데이터베이스 결합에 해당) 병합 노드는 해당 입력 데이터 세트를 키 필드별로 정렬합니다. 프로시저의 이 부분은 실행 시간이 $M \times N \times \log(N)$ 입니다. 여기서 M 은 입력 수이고 N 은 최신 입력에 있는 레코드 수입니다. 해당 입력 데이터 세트와 가장 큰 데이터 세트의 두 번째 사본을 모두 저장하려면 충분한 디스크 공간이 필요합니다. 병합 자체의 실행 시간은 출력 데이터 세트의 크기에 비례하며 이는 매치하는 키의 빈도에 따라 다릅니다. 출력이 입력의 데카르트 곱인 최악의 경우 실행 시간은 NM 에 접근할 수 있습니다. 이는 드문 경우이며 대부분의 결합은 매치하는 키 수가 훨씬 적습니다. 하나의 데이터 세트가 다른 데이터 세트보다 상대적으로 크거나 수신 데이터가 이미 키 필드별로 정렬되어 있으면 최적화 탭을 사용하여 이 노드의 성능을 개선할 수 있습니다.

통합. 키가 인접합 옵션이 설정되지 않은 경우 이 노드는 통합 출력을 생성하기 전에 전체 입력 데이터 세트를 읽지만 저장하지는 않습니다. 통합 데이터의 크기가 한계(IBM SPSS Modeler Server 구성 옵션 메모리 사용 승수에 의해 결정됨)에 도달하는 보다 극단적인 경우 나머지 데이터 세트는 키가 인접합 옵션이 설정된 것처럼 정렬되고 처리됩니다. 이 옵션이 설정되면 통합 출력 레코드는 입력 데이터를 읽을 때 생성되므로 데이터가 저장되지 않습니다.

고유. 고유 노드는 모든 고유한 키 필드를 입력 데이터 세트에 저장합니다. 모든 필드가 키 필드이고 모든 레코드가 고유한 경우 이는 전체 데이터 세트를 저장합니다. 기본적으로 고유 노드는 키 필드에서 데이터를 정렬한 다음 각 그룹에서 맨 처음 고유 레코드를 선택 또는 삭제합니다. 고유 키 개수가 적은 더 작은 데이터 세트 또는 미리 정렬된 데이터 세트의 경우 처리의 속도와 효율성을 높이는 옵션을 선택할 수 있습니다.

유형. 몇몇 경우에는 유형 노드는 값을 읽을 때 입력 데이터를 캐싱합니다. 캐시는 다운스트림 처리에 사용합니다. 캐시에는 전체 데이터 세트를 저장하려면 충분한 디스크 공간이 필요하지만 처리 속도를 높입니다.

평가. 평가 노드는 분위수를 계산하기 위해 입력 데이터를 정렬해야 합니다. 스코어와 결과적 레코드 순서가 서로 다르기 때문에 평가된 각 모델마다 정렬이 반복됩니다. 실행 시간은 $M \times N \times \log(N)$ 입니다. 여기서 M 은 모델 수이고 N 은 레코드 수입니다.

성능: 모델링 노드

신경망 및 코호넨. 코호넨 알고리즘을 포함하여 신경망 학습 알고리즘은 학습 데이터 위에서 전달하는 것이 많습니다. 데이터는 일부 한계까지 메모리에 저장되고 초과분은 디스크로 유출됩니다. 디스크에서 학습 데이터에 액세스하는 방법은 액세스 방법이 임의적이어서 과도한 디스크 활동으로 이어질 수 있으므로 비용이 많이 듭니다. 이러한 알고리즘의 경우 노드의 대화 상자의 모델 탭에서 속도 최적화 옵션을 선택하여 디스크 저장 공간 사용을 사용 안함으로 설정하여 모든 데이터가 메모리에 저장되도록 강제할 수 있습니다. 데이터를 저장하는 데 필요한 메모리 양이 서버가 처리하는 작업 세트보다 큰 경우 데이터의 일부는 디스크로 페이징되고 성능도 이에 따라 저하됩니다.

메모리 최적화가 사용 가능하면 물리적 RAM의 퍼센트는 IBM SPSS Modeler Server 구성 옵션 모델링 메모리 한계 퍼센트의 값에 따라 알고리즘에 할당됩니다. 신경망 훈련을 위해 더 많은 메모리를 사용하려면 더 많은 RAM을 제공하거나 이 옵션의 값을 늘리십시오. 그러나 값을 너무 높게 설정하면 페이징이 발생함을 유의하십시오.

신경망 알고리즘의 실행 시간은 필요한 정확도 수준에 따라 다릅니다. 노드의 대화 상자에서 중지 조건을 설정하여 실행 시간을 제어할 수 있습니다.

K-평균. K-평균 군집 알고리즘에는 메모리 사용을 제어하기 위한 신경망 알고리즘과 동일한 옵션이 있습니다. 그러나 데이터에 대한 액세스가 순차적이므로 디스크에 저장된 데이터에 대한 성능이 더 낮습니다.

성능: CLEM 표현식

데이터 스트림을 되돌아보는 CLEM 순차규칙 함수("@ 함수")는 가장 긴 되돌아보기를 충족하기 위해서는 충분한 데이터를 저장해야 합니다. 되돌아보기의 정도가 제한이 없는 작업의 경우에는 필드의 모든 값을 저장해야 합니다. 제한이 없는 작업은 오프셋 값이 리터럴 정수가 아닌 작업입니다(예: @OFFSET(Sales, Month)). 오프셋 값은 값이 실행될 때까지 알려지지 않은 필드 이름 월입니다. 서버는 정확한 결과를 위해서는 판매 필드의 모든 값을 저장해야 합니다. 상한선이 알려진 곳에서는 값을 추가 인수로서 제공해야 합니다(예: @OFFSET(Sales, Month, 12)). 이 작업은 서버에게 판매의 최근 12개 값을 초과하여 저장하지 않도록 지시합니다. 순차규칙 함수는 제한이 있거나, 그렇지 않은 경우에는 거의 항상 SQL 생성을 금지합니다.

제 17 장 IBM SPSS Modeler의 내게 필요한 옵션

IBM SPSS Modeler의 내게 필요한 옵션 개요

IBM SPSS Modeler는 시각적 및 기타 기능적 장애가 있는 사용자에 대한 특정 지원은 물론 모든 사용자에게 내게 필요한 옵션 지원을 제공합니다. 이 절에서는 스크린 리더, 키보드 단축키와 같은 내게 필요한 옵션 향상을 사용하여 작업하는 방법 및 기능에 대해 설명합니다.

내게 필요한 옵션 지원 유형

시각 장애가 있거나 키보드에 조작을 의존하는 경우 이 데이터 마이닝 툴킷을 사용하기 위한 매우 다양한 대체 방법이 있습니다. 예를 들어, 스트림 빌드, 옵션 지정, 출력 읽기를 마우스를 사용하지 않고 모두 수행할 수 있습니다. 사용 가능한 키보드 단축기는 뒤따르는 주제에 나열되어 있습니다. 또한 IBM SPSS Modeler는 Windows용 JAWS와 같은 스크린 리더에 대한 광범위한 지원을 제공합니다. 또한 색 구성표를 최적화하여 추가 대비를 제공할 수 있습니다. 이러한 유형의 지원은 다음 주제에서 설명합니다.

시각 장애인을 위한 내게 필요한 옵션

소프트웨어 사용 능력을 향상시킬 다수의 특성을 IBM SPSS Modeler에서 지정할 수 있습니다.

표시 옵션

그래프 표시 색상을 선택할 수 있습니다. 또한 소프트웨어 자체에 특정 Windows 설정을 사용하도록 선택할 수 있습니다. 이렇게 하면 시각적 대비를 증대하는 데 도움이 됩니다.

1. 표시 옵션을 설정하려면 도구 메뉴에서 **사용자 옵션**을 클릭하십시오.
2. **디스플레이** 탭을 클릭하십시오. 이 탭의 옵션으로는 소프트웨어 색 구성표, 차트 색상, 노드 글꼴 크기가 있습니다.

참고: 스크린 리더에서 그래프를 읽을 수 없으므로 시각 장애 사용자가 액세스할 수 없습니다.

알림을 위한 소리 사용

소리를 켜거나 꺼서 소프트웨어의 특정 조작에 대한 경고를 받는 방법을 제어할 수 있습니다. 예를 들어, 노드 작성 및 삭제 또는 새 출력이나 모델 생성과 같은 이벤트에 대한 소리를 활성화할 수 있습니다.

1. 알림 옵션을 설정하려면 도구 메뉴에서 **사용자 옵션**을 클릭하십시오.
2. 알림 탭을 클릭하십시오.

새 창 자동 시작 제어

사용자 옵션 대화 상자의 알림 탭은 테이블 및 차트와 같이 새로 생성된 출력이 별도의 창에서 시작되는지 여부를 제어하는 데도 사용합니다. 이 옵션을 사용 안함으로 설정하고 필요할 때만 출력 창을 여는 것이 용이합니다.

1. 이러한 옵션을 설정하려면 도구 메뉴에서 **사용자 옵션**을 클릭하십시오.
2. 알림 탭을 클릭하십시오.
3. 대화 상자의 **시각적 알림** 그룹에 있는 목록에서 **새 출력**을 선택하십시오.
4. 창 열기 아래에서 사용 안 함을 선택하십시오.

노드 크기

표준 또는 작은 크기를 사용하여 노드를 표시할 수 있습니다. 필요에 맞게 이러한 크기를 조정할 수 있습니다.

1. 노드 크기 옵션을 설정하려면 파일 메뉴에서 **스트림 특성**을 클릭하십시오.
2. **레이아웃** 탭을 클릭하십시오.

3. 아이콘 크기 목록에서 표준을 선택하십시오.

맹인 사용자를 위한 내게 필요한 옵션

맹인 사용자를 위한 지원은 대부분 Windows용 JAWS와 같은 스크린 리더 사용에 의존합니다. IBM SPSS Modeler에서 스크린 리더 사용을 최적화하기 위해 다수의 설정을 지정할 수 있습니다.

표시 옵션

스크린 리더는 화면에서 시각적 대비가 클 때 최상의 성능을 보이는 경향이 있습니다. 이미 고대비 Windows 설정이 있는 경우, 소프트웨어 자체에 이러한 Windows 설정을 사용하도록 선택할 수 있습니다.

1. 표시 옵션을 설정하려면 도구 메뉴에서 **사용자 옵션**을 클릭하십시오.
2. **디스플레이** 탭을 클릭하십시오.

참고: 스크린 리더에서 그래프를 읽을 수 없으므로 맹인 사용자가 액세스할 수 없습니다.

알림을 위한 소리 사용

소리를 켜거나 꺼서 소프트웨어의 특정 조작에 대한 경고를 받는 방법을 제어할 수 있습니다. 예를 들어, 노드 작성 및 삭제 또는 새 출력이나 모델 생성과 같은 이벤트에 대한 소리를 활성화할 수 있습니다.

1. 알림 옵션을 설정하려면 도구 메뉴에서 **사용자 옵션**을 클릭하십시오.
2. **알림** 탭을 클릭하십시오.

새 창 자동 시작 제어

사용자 옵션 대화 상자의 알림 탭은 새로 생성된 출력이 별도의 창에서 시작되는지 여부를 제어하는 데도 사용합니다. 이 옵션을 사용 안함으로 설정하고 필요에 따라 출력 창을 여는 것이 용이합니다.

1. 이러한 옵션을 설정하려면 도구 메뉴에서 **사용자 옵션**을 클릭하십시오.
2. **알림** 탭을 클릭하십시오.
3. 대화 상자의 시각적 알림 그룹에 있는 목록에서 새 출력을 선택하십시오.
4. 창 열기 아래에서 사용 안 함을 선택하십시오.

내게 필요한 옵션의 키보드 기능

키보드에서 제품 기능에 액세스할 수 있습니다. 가장 기본적인 레벨에서 Alt + 해당 키를 눌러 창 메뉴를 활성화 (예: Alt+F를 눌러 파일 메뉴에 액세스)하거나 Tab 키를 눌러 대화 상자 제어를 스크롤할 수 있습니다. 그러나 제품의 각 기본 창과 관련된 특정 문제와 대화 상자 탐색에 유용한 힌트가 있습니다.

이 절에서는 스트림 열기에서 노드 대화 상자를 사용하여 출력에 대해 작업하는 방법에 이르기까지 내게 필요한 옵션의 키보드 기능의 강조 표시에 대해 설명합니다. 또한 보다 효율적인 탐색을 위해 키보드 단축키 목록도 제공됩니다.

기본 창 탐색을 위한 단축키

IBM SPSS Modeler의 기본 창에서 데이터 마이닝 작업을 대부분 수행합니다. 기본 영역을 스트림 캔버스라고 하며 데이터 스트림을 빌드하고 실행하는 데 사용합니다. 창의 하단부는 사용 가능한 노드를 포함하는 노드 팔레트를 포함합니다. 팔레트는 노드의 각 그룹에 대한 데이터 마이닝 작업 유형에 해당하는 탭에 구성됩니다. 예를 들어, 데이터를 IBM SPSS Modeler로 가져오는 데 사용하는 노드는 소스 탭에 그룹화되고 필드 파생, 필터링 또는 입력에 사용하는 노드는 필드 작업 탭에 그룹화됩니다.

창의 오른쪽에는 스트림, 출력, 프로젝트 관리를 위한 몇 개의 도구가 있습니다. 오른쪽의 상단 절반은 관리자를 포함하며 스트림, 출력, 생성된 모델을 관리하는 데 사용하는 세 개의 탭이 있습니다. 탭과 오브젝트를 목록에서 선택하여 이러한 오브젝트에 액세스할 수 있습니다. 오른쪽의 하단 절반은 프로젝트 분할창을 포함하는데, 이를 사용하여 프로젝트에 대한 작업을 구성할 수 있습니다. 이 영역에는 프로젝트에 대한 서로 다른 두 개의 보기로 반영하는 두 개의 탭이 있습니다. **클래스 보기**는 프로젝트 오브젝트를 유형별로 정렬하는 반면, **CRISP-DM 보기**

는 데이터 준비 또는 모델링과 같은 관련 데이터 마이닝 단계별로 오브젝트를 정렬합니다. IBM SPSS Modeler 창의 이러한 다양한 측면은 도움말 시스템과 사용자 안내서 전반에 설명되어 있습니다.

다음은 기본 IBM SPSS Modeler 창 내에서의 이동 및 스트림 빌드에 사용하는 단축키 표입니다. 대화 상자 및 출력 단축키는 뒤따르는 주제에 나열되어 있습니다. 이러한 단축키는 기본 창에서만 사용할 수 있습니다.

표 39. 기본 창 단축키	
단축키	함수
Ctrl+F5	노드 팔레트로 초점을 이동합니다.
Ctrl+F6	스트림 캔버스로 초점을 이동합니다.
Ctrl+F7	관리자 분할창으로 초점을 이동합니다.
Ctrl+F8	프로젝트 분할창으로 초점을 이동합니다.

표 40. 노드 및 스트림 단축키	
단축키	함수
Ctrl+N	빈 스트림 캔버스를 새로 작성합니다.
Ctrl+O	열기 대화 상자를 표시합니다(여기서 기존 스트림을 선택하고 열 수 있음).
Ctrl+숫자 키	창 또는 분할창의 해당 탭으로 초점을 이동합니다. 예를 들어, 탭 지정된 분할창 또는 창 내에서 Ctrl+1은 왼쪽에서 시작하여 첫 번째 탭으로 이동하고 Ctrl+2는 두 번째 탭으로 이동하는 식입니다.
Ctrl+아래로 화살표	팔레트 탭에서 해당 탭 아래의 첫 번째 노드로 초점을 이동하기 위해 노드 팔레트에서 사용합니다.
Ctrl+위쪽 화살표	노드에서 팔레트 탭으로 초점을 이동하기 위해 노드 팔레트에서 사용합니다.
Enter	노드 팔레트에서 노드가 선택된 경우(생성된 모델 팔레트의 세분화된 모델 포함) 이 키 입력은 스트림 캔버스에 노드를 추가합니다. 캔버스에서 노드가 이미 선택된 경우 Enter 를 누르면 해당 노드에 대한 대화 상자가 열립니다.
Ctrl+Enter	팔레트에서 노드가 선택된 경우 노드를 선택하지 않고 스트림 캔버스에 해당 노드를 추가하고 팔레트의 첫 번째 노드로 초점을 이동합니다.
Alt+Enter	팔레트에서 노드가 선택된 경우 스트림 캔버스에 해당 노드를 추가하고 팔레트의 첫 번째 노드로 초점을 이동하는 동안 노드를 선택합니다.
Shift+스페이스바	노드나 주석이 팔레트에서 초점을 가진 경우 해당 노드나 주석 선택과 선택 취소 간에 토글합니다. 다른 노드나 주석도 선택된 경우 선택 취소됩니다.
Ctrl+Shift+스페이스바	노드나 주석이 스트림에서 초점을 가진 경우 또는 노드나 주석이 팔레트에서 초점을 가진 경우 노드나 주석 선택과 선택 취소 간에 토글합니다. 선택된 다른 노드나 주석에는 영향을 주지 않습니다.
왼쪽/오른쪽 화살표	스트림 캔버스가 초점을 가진 경우 전체 스트림을 화면에서 수평으로 이동합니다. 팔레트 탭이 초점을 가진 경우 탭 간에 순환합니다. 팔레트 노드가 초점을 가진 경우 팔레트의 노드 간에 이동합니다.

표 40. 노드 및 스트림 단축키 (계속)

단축키	함수
위로/아래로 화살표	스트림 캔버스가 초점을 가진 경우 전체 스트림을 화면에서 수직으로 이동합니다. 팔레트 노드가 초점을 가진 경우 팔레트의 노드 간에 이동합니다. 하위 팔레트가 초점을 가진 경우 이 팔레트 탭의 다른 하위 팔레트 간에 이동합니다.
Alt+왼쪽/오른쪽 화살표	스트림 캔버스에서 선택된 노드와 주석을 화살표 키 방향으로 수평으로 이동합니다.
Alt+위로/아래로 화살표	스트림 캔버스에서 선택된 노드와 주석을 화살표 키 방향으로 수직으로 이동합니다.
Ctrl+A	스트림의 모든 노드를 선택합니다.
Ctrl+Q	노드가 초점을 가진 경우 해당 노드와 모든 노드 다운스트림을 선택하고 모든 노드 업스트림을 선택 취소합니다.
Ctrl+W	선택된 노드가 초점을 가진 경우 해당 노드와 선택된 모든 노드 다운스트림을 선택 취소합니다.
Ctrl+Alt+D	선택된 노드를 복제합니다.
Ctrl+Alt+L	스트림에서 모델 너깃이 선택된 경우 삽입 대화 상자를 열어 저장된 모델을 .nod 파일에서 스트림으로 로드할 수 있게 합니다.
Ctrl+Alt+R	선택된 노드에 대한 주석(Annotation) 탭을 표시하여 노드 이름을 변경할 수 있게 합니다.
Ctrl+Alt+U	사용자 입력 소스 노드를 작성합니다.
Ctrl+Alt+C	노드 캐시 설정 또는 해제를 토글합니다.
Ctrl+Alt+F	노드 캐시를 비웁니다.
Tab	스트림 캔버스에서 현재 스트림의 모든 소스 노드와 주석을 순환합니다. 노드 팔레트에서 팔레트의 노드 간에 이동합니다. 선택된 하위 팔레트에서 하위 팔레트의 첫 번째 노드로 이동합니다.
Shift+Tab	Tab과 동일한 작업을 역순으로 수행합니다.
Ctrl+Tab	관리자 분할창 또는 프로젝트 분할창에 초점을 맞춘 상태에서 스트림 캔버스로 초점을 이동합니다. 노드 팔레트에 초점을 맞춘 상태에서 노드와 해당 팔레트 탭 간에 초점을 이동합니다.
임의의 알파벳 키	현재 스트림의 노드에 초점을 맞춘 상태에서 이름이 눌려진 키로 시작하는 다음 노드에 초점을 맞추고 이동합니다.
F1	초점 관련 주제에서 도움말 시스템을 엽니다.
F2	캔버스에서 선택된 노드에 대한 연결 프로세스를 시작합니다. Tab 키를 사용하여 캔버스의 필수 노드로 이동하고 Shift+스페이스바를 눌러 연결을 마치십시오.
F3	캔버스에서 선택된 노드에 대한 연결을 모두 삭제합니다.
F6	관리자 분할창, 프로젝트 분할창, 노드 팔레트 간에 초점을 이동합니다.
F10	파일 메뉴를 엽니다.
Shift+F10	노드 또는 스트림에 대한 팝업 메뉴를 엽니다.
Delete	캔버스에서 선택된 노드를 삭제합니다.
Esc	팝업 메뉴 또는 대화 상자를 닫습니다.

표 40. 노드 및 스트림 단축키 (계속)

단축키	함수
Ctrl+Alt+X	수퍼노드를 펼칩니다.
Ctrl+Alt+Z	수퍼노드를 확대합니다.
Ctrl+Alt+Shift+Z	수퍼노드를 축소합니다.
Ctrl+E	스트림 캔버스에 초점을 맞춘 상태에서 현재 스트림을 실행합니다.

IBM SPSS Modeler에서는 다수의 표준 단축키도 사용합니다(예: 복사의 경우 Ctrl+C). 자세한 정보는 [18 페이지의 『단축키 사용』](#)의 내용을 참조하십시오.

대화 상자 및 테이블 단축키

대화 상자, 테이블, 대화 상자의 테이블에 대한 작업 시 몇 개의 단축키와 스크린 리더 키가 유용합니다. 특수 키보드 및 스크린 리더 단축키 전체 목록은 다음과 같습니다.

표 41. 대화 상자 및 표현식 작성기 단축키

단축키	함수
Alt+4	열린 대화 상자 또는 출력 창을 모두 종료하는 데 사용합니다. 관리자 분할창의 출력 탭에서 출력을 검색할 수 있습니다.
Ctrl+End	표현식 작성기의 임의 제어에 초점을 맞춘 상태에서 표현식의 끝으로 삽입점을 이동합니다.
Ctrl+1	표현식 작성기에서 표현식 편집 제어로 초점을 이동합니다.
Ctrl+2	표현식 작성기에서 함수 목록으로 초점을 이동합니다.
Ctrl+3	표현식 작성기에서 필드 목록으로 초점을 이동합니다.

테이블 단축키

테이블 단축기는 유형, 필터, 병합과 같은 노드에 대해 대화 상자의 테이블 제어에는 물론 출력 테이블에도 사용합니다. 일반적으로 테이블 셀 간의 이동에는 Tab 키를 사용하고 테이블 제어를 종료하는 데는 Ctrl+Tab을 사용합니다. 참고: 때때로 스크린 리더가 셀 내용 읽기를 즉시 시작하지 않을 수도 있습니다. 화살표 키를 한 번 또는 두 번 누르면 소프트웨어가 재설정되고 음성을 시작합니다.

표 42. 테이블 단축키

단축키	함수
Ctrl+W	테이블에 대해 선택된 행에 대한 간단한 설명을 읽습니다. 예를 들어, "선택된 행 2 값은 성, 플래그, 남/여 등입니다."
Ctrl+Alt+W	테이블에 대해 선택된 행에 대한 자세한 설명을 읽습니다. 예를 들어, "선택된 행 2 값은 필드 = 성, 유형 = 플래그, 성 = 남/여 등입니다."
Ctrl+D	테이블에 대해 선택된 영역에 대한 간단한 설명을 읽습니다(예: "선택은 6열당 1행임").
Ctrl+Alt+D	테이블에 대해 선택된 영역에 대한 자세한 설명을 제공합니다(예: "선택은 6열당 1행임. 선택된 열은 필드, 유형, 결측값임. 선택 행은 1임").
Ctrl+T	테이블에 대해 선택된 열에 대한 간단한 설명을 제공합니다(예: "필드, 유형, 결측값").

표 42. 테이블 단축키 (계속)	
단축키	함수
Ctrl+Alt+T	테이블에 대해 선택된 열에 대한 자세한 설명을 제공합니다(예: "선택된 열은 필드, 유형, 결측값임").
Ctrl+R	테이블에 대해 테이블의 레코드 수를 제공합니다.
Ctrl+Alt+R	테이블에 대해 테이블의 레코드 수는 물론 열 이름도 제공합니다.
Ctrl+I	테이블에 대해 초점이 있는 셀의 셀 정보 또는 내용을 읽습니다.
Ctrl+Alt+I	테이블에 대해 초점이 있는 셀의 셀 정보(열 이름 및 셀 내용)에 대한 자세한 설명을 읽습니다.
Ctrl+G	테이블에 대해 간단한 일반 선택 정보를 제공합니다.
Ctrl+Alt+G	테이블에 대해 자세한 일반 선택 정보를 제공합니다.
Ctrl+Q	테이블에 대해 테이블 셀 빠른 토클을 제공합니다. Ctrl+Q는 화살표 키를 사용하여 테이블을 이동할 때 "성=여성"과 같은 자세한 설명을 읽습니다. Ctrl+Q를 다시 선택하면 간단한 설명(셀 내용)으로 토클 됩니다.
F8	테이블에 대해 초점이 테이블이면 열 헤더로 초점을 설정합니다.
스페이스바	테이블에 대해 초점이 열 헤더이면 열 정렬을 사용 가능하게 합니다.

주석 단축키

온스크린 주석에 대한 작업 시 다음 단축키를 사용할 수 있습니다.

표 43. 주석 단축키	
단축키	함수
Alt+C	주석 표시/숨기기 기능을 토글합니다.
Alt+M	주석이 현재 표시된 경우 새 주석을 삽입합니다. 주석이 현재 숨겨진 경우 주석을 표시합니다.
Tab	스트림 캔버스에서 현재 스트림의 모든 소스 노드와 주석을 순환합니다.
Enter	주석에 초점이 있으면 편집 시작을 표시합니다.
Alt+Enter 또는 Ctrl+Tab	편집을 종료하고 편집 변경사항을 저장합니다.
Esc	편집을 취소합니다. 편집 중에 작성된 변경사항은 유실됩니다.
Alt+Shift+위쪽 화살표	눈금에 정렬이 설정(또는 설정 해제)되어 있으면 1눈금 셀(또는 1픽셀)씩 텍스트 영역의 높이를 줄입니다.
Alt+Shift+아래로 화살표	눈금에 정렬이 설정(또는 설정 해제)되어 있으면 1눈금 셀(또는 1픽셀)씩 텍스트 영역의 높이를 늘립니다.
Alt+Shift+왼쪽 화살표	눈금에 정렬이 설정(또는 설정 해제)되어 있으면 1눈금 셀(또는 1픽셀)씩 텍스트 영역의 너비를 줄입니다.
Alt+Shift+오른쪽 화살표	눈금에 정렬이 설정(또는 설정 해제)되어 있으면 1눈금 셀(또는 1픽셀)씩 텍스트 영역의 너비를 늘립니다.

군집 뷰어 및 모델 뷰어 단축키

군집 뷰어 및 모델 뷰어 창 주위를 이동하는 데 단축키를 사용할 수 있습니다.

표 44. 일반 단축키 - 군집 뷰어 및 모델 뷰어	
단축키	함수
Tab	다음 화면 제어로 초점을 이동합니다.
Shift+Tab	이전 화면 제어로 초점을 이동합니다.
아래로 화살표	드롭 다운 목록에 초점이 있으면 목록을 열거나 목록의 다음 항목으로 이동합니다. 메뉴에 초점이 있으면 메뉴의 다음 항목으로 이동합니다. 썸네일 그래프에 초점이 있으면 세트의 다음 항목(또는 마지막 썸네일에 초점이 있으면 첫 번째 항목)으로 이동합니다.
위쪽 화살표	드롭 다운 목록이 열려 있으면 목록의 이전 항목으로 이동합니다. 메뉴에 초점이 있으면 메뉴의 이전 항목으로 이동합니다. 썸네일 그래프에 초점이 있으면 세트의 이전 항목(또는 첫 번째 썸네일에 초점이 있으면 마지막 항목)으로 이동합니다.
Enter	열려 있는 드롭 다운 목록을 닫거나 열린 메뉴를 선택합니다.
F6	창의 왼쪽 및 오른쪽 분할창 사이에서 초점을 토글합니다.
왼쪽 및 오른쪽 화살표	탭에 초점이 있으면 이전 또는 다음 탭으로 이동합니다. 메뉴에 초점이 있으면 이전 또는 다음 메뉴로 이동합니다.
Alt+문자	해당 이름의 이 문자에 밑줄이 쳐져 있는 단추 또는 메뉴를 선택합니다.
Esc	열린 메뉴 또는 드롭 다운 목록을 닫습니다.

군집 뷰어만

군집 뷰어에는 기능별 군집 눈금이 포함된 군집 뷰어가 있습니다.

모델 요약 보기 대신 군집 보기 선택하려면 다음을 수행하십시오.

1. 보기 단추가 선택될 때까지 반복해서 탭을 누르십시오.
 2. 아래로 화살표를 두 번 눌러 군집을 선택하십시오.
- 여기서부터 눈금 내 개별 셀을 선택할 수 있습니다.
3. 시각화 도구 모음의 마지막 아이콘에 도착할 때까지 반복해서 탭을 누르십시오.



그림 19. 시각화 트리 표시 아이콘

4. Tab을 한 번 더 누른 후 스페이스바를 누르고 화살표 키를 누르십시오.

이제 다음 키보드 단축키를 사용할 수 있습니다.

표 45. 군집 뷰어 단축키	
단축키	함수
화살표 키	눈금의 개별 셀 간에 초점을 이동합니다. 초점이 이동함에 따라 오른쪽 분할창의 셀 분포 표시가 변경됩니다.
Ctrl+,(쉼표)	셀에 초점이 있는 눈금의 전체 열을 선택하거나 선택 취소합니다. 선택에 열을 추가하려면 화살표 키를 사용하여 해당 열의 셀로 이동한 후 Ctrl+,를 다시 누르십시오.
Tab	초점이 눈금에서 벗어나 다음 화면 제어로 이동합니다.
Shift+Tab	초점이 눈금에서 벗어나 이전 화면 제어로 다시 이동합니다.
F2	편집 모드에 진입합니다(레이블 및 설명 셀만).
Enter	편집 변경사항을 저장하고 편집 모드를 종료합니다(레이블 및 설명 셀만).
Esc	편집 변경사항을 저장하지 않고 편집 모드를 종료합니다(레이블 및 설명 셀만).

단축키 예제: 스트림 빌드

키보드나 스크린 리더에 의존하는 사용자를 위해 스트림 빌드 프로세스를 보다 명확하게 하기 위해 다음은 마우스를 사용하지 않고 스트림을 빌드하는 예제입니다. 이 예제에서는 다음 단계를 사용하여 가변파일 노드, 파생 노드, 히스토그램 노드가 포함된 스트림을 빌드합니다.

1. **IBM SPSS Modeler 시작.** IBM SPSS Modeler가 처음 시작될 때 초점은 노드 팔레트의 즐겨찾기 탭에 맞춰져 있습니다.
 2. **Ctrl+아래로 화살표.** 탭 자체에서 탭 본문으로 초점을 이동합니다.
 3. **오른쪽 화살표.** 가변파일 노드로 초점을 이동합니다.
 4. **스페이스바.** 가변파일 노드를 선택합니다.
 5. **Ctrl+Enter.** 스트림 캔버스에 가변파일 노드를 추가합니다. 이 키 조합은 가변파일 노드에 대한 선택도 유지하므로 추가된 다음 노드는 가변파일 노드에 연결됩니다.
 6. **Tab.** 노드 팔레트로 초점을 다시 이동합니다.
 7. **오른쪽 화살표 네 번.** 파생 노드로 이동합니다.
 8. **스페이스바.** 파생 노드를 선택합니다.
 9. **Alt+Enter.** 캔버스에 파생 노드를 추가하고 파생 노드로 선택을 이동합니다. 이제 이 노드는 추가된 다음 노드에 연결할 준비가 되었습니다.
 10. **Tab.** 노드 팔레트로 초점을 다시 이동합니다.
 11. **오른쪽 화살표 다섯 번.** 팔레트의 히스토그램 노드로 초점을 이동합니다.
 12. **스페이스바.** 히스토그램 노드를 선택합니다.
 13. **Enter.** 스트림에 노드를 추가하고 스트림 캔버스로 초점을 이동합니다.
- 다음 예제를 계속하거나 나중에 다음 예제를 시도하려면 스트림을 저장하십시오.

단축키 예제: 노드 편집

이 예제에서는 이전 예제에 빌드된 스트림을 사용합니다. 스트림은 가변파일 노드, 파생 노드, 히스토그램 노드로 구성되어 있습니다. 지시사항은 스트림의 세 번째 노드인 히스토그램 노드에 초점을 맞추고 시작합니다.

1. **Ctrl+왼쪽 화살표 두 번.** 초점을 가변파일 노드로 다시 이동합니다.

2. **Enter.** 가변파일 대화 상자를 엽니다. Tab 키를 사용하여 파일 필드로 이동하고 텍스트 파일 경로와 이름을 입력하여 해당 파일을 선택하십시오. Ctrl+Tab을 눌러 대화 상자의 하단 부분으로 이동하고 확인 단추까지 Tab 키로 이동한 후 Enter를 눌러 대화 상자를 닫으십시오.
3. **Ctrl+오른쪽 화살표.** 두 번째 노드인 파생 노드에 초점을 맞춥니다.
4. **Enter.** 파생 노드 대화 상자를 엽니다. Tab 키로 이동하여 필드를 선택하고 파생 조건을 지정하십시오. Ctrl +Tab을 눌러 확인 단추까지 이동하고 Enter를 눌러 대화 상자를 닫으십시오.
5. **Ctrl+오른쪽 화살표.** 세 번째 노드인 히스토그램 노드에 초점을 맞춥니다.
6. **Enter.** 히스토그램 노드 대화 상자를 엽니다. Tab 키로 이동하여 필드를 선택하고 그래프 옵션을 지정하십시오. 드롭 다운 목록의 경우 아래로 화살표를 눌러 목록을 열고 목록 항목을 강조 표시한 후 Enter를 눌러 목록 항목을 선택하십시오. 확인 단추까지 Tab 키로 이동한 후 Enter를 눌러 대화 상자를 닫으십시오.

이 시점에서 추가 노드를 추가하거나 현재 스트림을 실행할 수 있습니다. 스트림을 빌드하고 있을 때 다음 팁을 명심하십시오.

- 수동으로 노드를 연결하는 경우 F2를 사용하여 연결 시작점을 작성하고 Tab 키를 사용하여 끝점으로 이동한 후 Shift+스페이스바를 사용하여 연결을 마무리하십시오.
- 캔버스의 선택된 노드에 대한 연결을 모두 영구 삭제하려면 F3을 사용하십시오.
- 스트림을 작성했으면 Ctrl+E를 사용하여 현재 스트림을 실행하십시오.

단축키 전체 목록을 사용할 수 있습니다. 자세한 정보는 222 페이지의 [『기본 창 탐색을 위한 단축키』](#)의 내용을 참조하십시오.

스크린 리더 사용

시장에서 다수의 스크린 리더를 사용할 수 있습니다. IBM SPSS Modeler는 IBM SPSS Modeler와 함께 설치되는 Java Access Bridge를 사용하여 Windows용 JAWS를 지원하도록 구성되어 있습니다. JAWS가 설치된 경우에는 IBM SPSS Modeler를 시작하기 전에 JAWS를 시작하여 이 제품을 사용하면 됩니다.

참고: SPSS Modeler에서 JAWS를 실행하려면 최소 6GB의 공간을 권장합니다.

IBM SPSS Modeler의 데이터 마이닝 프로세스에 대한 고유한 그래픽 표시 특성으로 인해 차트와 그래프는 시각적으로 최적으로 사용합니다. 그러나 스크린 리더를 사용하여 텍스트로 보여지는 출력과 모델을 기반으로 이해하고 결정할 수 있습니다.

참고: 64비트 클라이언트 시스템을 사용할 경우 일부 지원 기술 기능이 작동하지 않습니다. 그 이유는 Java Access Bridge가 64비트 작업에 맞게 설계되지 않았기 때문입니다.

IBM SPSS Modeler 사전 파일 사용

IBM SPSS Modeler 사전 파일(Awt.JDF)을 JAWS에 사용할 수 있습니다. 이 파일을 사용하려면 다음을 수행하십시오.

1. IBM SPSS Modeler 설치의 /accessibility 서브디렉토리로 이동하여 사전 파일(Awt.JDF)을 복사하십시오.
2. JAWS 스크립트가 있는 디렉토리로 복사하십시오.

다른 JAVA 애플리케이션이 실행 중인 경우에는 시스템에 Awt.JDF 파일이 이미 있을 수 있습니다. 이 경우, 사전 파일을 수동으로 편집해야 이 사전 파일을 사용할 수 있습니다.

HTML 출력에 스크린 리더 사용

스크린 리더를 사용하여 IBM SPSS Modeler에 HTML로 표시된 출력을 볼 때 일부 난관에 봉착할 수 있습니다. 다음을 포함하여 다수의 출력이 영향을 받습니다.

- 회귀분석, 로지스틱 회귀분석, 요인/PCA 노드의 고급 탭에서 보여지는 출력
- 보고서 노드 출력

이러한 창이나 대화 상자 각각에는 기본 브라우저에서 출력을 실행하는 데 사용할 수 있는 도구가 도구 모음에 있습니다. 그러면 스크린 리더를 사용하여 출력 정보를 전달할 수 있습니다.

대화식 트리 창에서 내게 필요한 옵션

대화식 트리 창에서 의사결정 트리 모형의 표준 표시는 스크린 리더에 문제를 일으킬 수 있습니다. 액세스 가능한 버전에 액세스하려면 대화식 트리 메뉴에서 다음을 클릭하십시오.

보기 > 액세스 가능 창

이는 표준 트리 맵과 유사한 보기를 표시하지만 JAWS가 올바르게 읽을 수 있는 보기입니다. 표준 화살표 창을 사용하여 위, 아래, 오른쪽 또는 왼쪽으로 이동할 수 있습니다. 액세스 가능 창을 탐색할 때 대화식 트리 창의 초점도 이에 따라서 이동합니다. 스페이스바를 사용하여 선택을 변경하거나 Ctrl+스페이스바를 사용하여 현재 선택을 확장하십시오.

사용 팁

IBM SPSS Modeler 환경에 보다 쉽게 액세스할 수 있는 몇 가지 팁이 있습니다. IBM SPSS Modeler에서 작업 시 일반적인 힌트는 다음과 같습니다.

- **확장 텍스트 상자 종료.** Ctrl+Tab을 사용하여 확장 텍스트 상자를 종료하십시오. 참고: Ctrl+Tab은 테이블 제어를 종료하는 데도 사용합니다.
- **화살표 키보다 Tab 키 사용.** 대화 상자 옵션 선택 시 Tab 키를 사용하여 옵션 단추 간에 이동하십시오. 화살표 키는 이 컨텍스트에서 작동하지 않습니다.
- **드롭 다운 목록.** 대화 상자에 대한 드롭 다운 목록에서 Escape 키 또는 스페이스바를 사용하여 항목을 선택한 후 목록을 닫을 수 있습니다. 또한 Escape 키를 사용하여 다른 제어에 템을 지정한 경우 닫지 않는 드롭 다운 목록을 닫을 수 있습니다.
- **실행 상태.** 대형 데이터베이스에서 스트림을 실행 중인 경우 스트림 상태를 읽을 때 JAWS가 뒤쳐질 수 있습니다. 주기적으로 Ctrl 키를 눌러 상태 보고를 업데이트하십시오.
- **노드 팔레트 사용.** 노드 팔레트의 템에 처음 진입할 때 JAWS가 노드 이름 대신 "그룹 상자"를 읽는 경우가 있습니다. 이 경우, Ctrl+오른쪽 화살표와 Ctrl+왼쪽 화살표를 차례로 사용하여 스크린 리더를 재설정하고 노드 이름을 들을 수 있습니다.
- **메뉴 읽기.** 때때로 메뉴를 처음 열 때 JAWS가 첫 번째 메뉴 항목을 읽지 않을 수 있습니다. 이런 상황이 발생했다고 의심되면 아래로 화살표와 위쪽 화살표를 차례로 사용하여 메뉴의 첫 번째 항목을 들으십시오.
- **계단식 메뉴.** JAWS는 계단식 메뉴의 첫 번째 레벨을 읽지 않습니다. 메뉴를 이동하는 동안 소리가 중단되면 오른쪽 화살표 키를 눌러 하위 메뉴 항목을 들으십시오.

또한 IBM SPSS Modeler Text Analytics가 설치된 경우 다음 팁을 사용하면 대화식 워크벤치 인터페이스에 보다 쉽게 액세스할 수 있습니다.

- **대화 상자 진입.** 대화 상자에 진입할 때 첫 번째 제어에 초점을 맞추려면 Tab 키를 눌러야 합니다.
- **확장 텍스트 상자 종료.** Ctrl+Tab을 사용하여 확장 텍스트 상자를 종료하고 다음 제어로 이동하십시오. 참고: Ctrl+Tab은 테이블 제어를 종료하는 데도 사용합니다.
- **트리 목록의 요소를 찾기 위해 첫 번째 문자 입력.** 범주 분할창, 추출된 결과 분할창 또는 라이브러리 트리에서 요소를 찾을 때 분할창이 초점을 가진 경우 요소의 첫 번째 문자를 입력할 수 있습니다. 이렇게 하면 입력한 문자로 시작하는 요소의 다음 발생이 선택됩니다.
- **드롭 다운 목록.** 대화 상자에 대한 드롭 다운 목록에서 스페이스바를 사용하여 항목을 선택한 후 목록을 닫을 수 있습니다.

다른 소프트웨어와의 충돌

IBM SPSS Modeler에서 JAWS와 같은 스크린 리더 사용을 검정할 때 개발 팀은 조직 내에서 SMS(Systems Management Server)를 사용하면 IBM SPSS Modeler와 같은 Java 기반 애플리케이션을 읽을 수 있는 JAWS의 기능을 방해할 수 있음을 발견했습니다. SMS를 사용하지 않으면 이런 상황이 발생하지 않습니다. SMS에 대한 자세한 정보를 알려면 Microsoft 웹 사이트를 방문하십시오.

JAWS 및 Java

여러 버전의 JAWS는 Java 기반 소프트웨어 애플리케이션에 대한 다양한 지원 레벨을 제공합니다. IBM SPSS Modeler에서 최신 버전의 모든 JAWS가 작동하긴 하지만 일부 버전에서는 Java 기반 시스템에서 사용할 경우

사소한 문제점이 발생할 수 있습니다. Windows용 JAWS 웹 사이트(<http://www.FreedomScientific.com>)를 방문하십시오.

IBM SPSS Modeler에서 그래프 사용

히스토그램, 평가 차트, 다중 도표, 산점도와 같은 정보에 대한 시각적 표시는 스크린 리더로 해석하기 어렵습니다. 그러나 출력 창에서 사용 가능한 텍스트 요약을 사용하여 해당 웹 그래프와 분포를 볼 수는 있습니다.

제 18 장 유니코드 지원

IBM SPSS Modeler에서 유니코드 지원

IBM SPSS Modeler는 IBM SPSS Modeler 및 IBM SPSS Modeler Server 둘 모두에 대해 완전히 유니코드 사용이 가능합니다. 이로 인해 다중 언어 데이터베이스를 포함하여 유니코드를 지원하는 다른 애플리케이션과 데이터를 로케일 특정 인코딩 스키마로/에서의 변환으로 초래될 수 있는 데이터의 유실 없이 교환하는 것이 가능합니다.

- IBM SPSS Modeler는 유니코드 데이터를 내부적으로 저장하고 유니코드로 저장된 다중 언어 데이터를 유실 없이 데이터베이스에 씁니다.
- IBM SPSS Modeler는 UTF-8 인코딩 텍스트 파일을 읽고 쓸 수 있습니다. 텍스트 파일 가져오기 및 내보내기는 기본적으로 로케일 인코딩으로 설정되지만 대안으로 UTF-8을 지원합니다. 이 설정은 파일 가져오기와 내보내기 노드에 지정되거나 기본 인코딩을 스트림 특성 대화 상자에서 변경할 수 있습니다. 자세한 정보는 [35 페이지의 『스트림의 일반 옵션 설정』](#)의 내용을 참조하십시오.
- 로케일 인코딩에 저장된 Statistics, SAS 및 텍스트 데이터 파일은 가져올 때 UTF-8으로 변환되고 내보낼 때 다시 돌아갑니다. 파일에 쓸 때, 로케일 문자 세트에 없는 유니코드 문자가 있는 경우에는 이들은 대체되고 경고가 표시됩니다. 이는 데이터를 유니코드를 지원하고(데이터베이스 또는 UTF-8 텍스트 파일) 다른 로케일이나 다중 로케일 또는 문자 세트의 문자를 포함하는 데이터 소스로부터 가져오는 경우에만 발생합니다.
- IBM SPSS Modeler Solution Publisher 이미지는 UTF-8 인코딩이 되어 있고 플랫폼과 로케일 사이에 휴대가 가능합니다.

유니코드 정보

유니코드 표준의 목표는 경계선, 로케일 및 애플리케이션 간에 쉽게 공유할 수 있도록 다국어 텍스트를 인코딩하는 일관된 방법을 제공합니다. 현재 버전 4.0.1인 유니코드 표준은 오늘날 세계에서 공통으로 사용하고 있는 모든 문자 세트의 수퍼세트인 문자 세트를 정의하고 각 문자에 고유한 이름과 코드 포인트를 지정합니다. 문자 및 해당 코드 포인트는 ISO-10646에 의해 정의된 범용 문자 세트(UCS)의 그것과 동일합니다. 자세한 정보는 [유니코드 홈 페이지](#)를 참조하십시오.

주의사항

이 정보는 미국에서 제공되는 제품 및 서비스용으로 작성된 것입니다. 이 자료는 IBM에서 다른 언어로 사용 가능합니다. 그러나 자료에 접근하기 위해서는 해당 언어로 된 제품 또는 제품 버전의 사본이 필요할 수 있습니다.

IBM은 다른 국가에서 이 책에 기술된 제품, 서비스 또는 기능을 제공하지 않을 수도 있습니다. 현재 사용할 수 있는 제품 및 서비스에 대한 정보는 한국 IBM 담당자에게 문의하십시오. 이 책에서 IBM 제품, 프로그램 또는 서비스를 언급했다고 해서 해당 IBM 제품, 프로그램 또는 서비스만을 사용할 수 있다는 것을 의미하지는 않습니다. IBM의 지적 재산권을 침해하지 않는 한, 기능상으로 동등한 제품, 프로그램 또는 서비스를 대신 사용할 수도 있습니다. 그러나 비IBM 제품, 프로그램 또는 서비스의 운영에 대한 평가 및 검증은 사용자의 책임입니다.

IBM은 이 책에서 다루고 있는 특정 내용에 대해 특허를 보유하고 있거나 현재 특허 출원 중일 수 있습니다. 이 책을 제공한다고 해서 특허에 대한 라이센스까지 부여하는 것은 아닙니다. 라이센스에 대한 의문사항은 다음으로 문의하십시오.

07326

서울특별시 영등포구
국제금융로 10, 3IFC
한국 아이.비.엠 주식회사
대표전화서비스: 02-3781-7114

2바이트(DBCS) 정보에 관한 라이센스 문의는 한국 IBM에 문의하거나 다음 주소로 서면 문의하시기 바랍니다.

*Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan*

IBM은 타인의 권리 비침해, 상품성 및 특정 목적에의 적합성에 대한 묵시적 보증을 포함하여(단, 이에 한하지 않음) 묵시적이든 명시적이든 어떠한 종류의 보증 없이 이 책을 "현상태대로" 제공합니다. 일부 국가에서는 특정 거래에서 명시적 또는 묵시적 보증의 면책사항을 허용하지 않으므로, 이 사항이 적용되지 않을 수도 있습니다.

이 정보에는 기술적으로 부정확한 내용이나 인쇄상의 오류가 있을 수 있습니다. 이 정보는 주기적으로 변경되며, 변경된 사항은 최신판에 통합됩니다. IBM은 이 책에서 설명한 제품 및/또는 프로그램을 사전 통지 없이 언제든지 개선 및/또는 변경할 수 있습니다.

이 정보에서 언급되는 비IBM 웹 사이트는 단지 편의상 제공된 것으로, 어떤 방식으로든 이들 웹 사이트를 옹호하고자 하는 것은 아닙니다. 해당 웹 사이트의 자료는 본 IBM 제품 자료의 일부가 아니므로 해당 웹 사이트 사용으로 인한 위험은 사용자 본인이 감수해야 합니다.

IBM은 귀하의 권리를 침해하지 않는 범위 내에서 적절하다고 생각하는 방식으로 귀하가 제공한 정보를 사용하거나 배포할 수 있습니다.

(i) 독립적으로 작성된 프로그램과 기타 프로그램(본 프로그램 포함) 간의 정보 교환 및 (ii) 교환된 정보의 상호 이용을 목적으로 본 프로그램에 관한 정보를 얻고자 하는 라이센스 사용자는 다음 주소로 문의하십시오.

07326

서울특별시 영등포구
국제금융로 10, 3IFC
한국 아이.비.엠 주식회사
대표전화서비스: 02-3781-7114

이러한 정보는 해당 조건(예를 들면, 사용료 지불 등)하에서 사용될 수 있습니다.

이 정보에 기술된 라이센스가 부여된 프로그램 및 프로그램에 대해 사용 가능한 모든 라이센스가 부여된 자료는 IBM이 IBM 기본 계약, IBM 프로그램 라이센스 계약(IPLA) 또는 이와 동등한 계약에 따라 제공한 것입니다.

인용된 성능 데이터와 고객 예제는 예시 용도로만 제공됩니다. 실제 성능 결과는 특정 구성과 운영 조건에 따라 다를 수 있습니다.

비IBM 제품에 관한 정보는 해당 제품의 공급업체, 공개 자료 또는 기타 범용 소스로부터 얻은 것입니다. IBM에서는 이러한 제품들을 테스트하지 않았으므로, 비IBM 제품과 관련된 성능의 정확성, 호환성 또는 기타 청구에 대해서는 확신할 수 없습니다. 비IBM 제품의 성능에 대한 의문사항은 해당 제품의 공급업체에 문의하십시오.

IBM이 제시하는 방향 또는 의도에 관한 모든 언급은 특별한 통지 없이 변경될 수 있습니다.

이 정보에는 일상의 비즈니스 운영에서 사용되는 자료 및 보고서에 대한 예제가 들어 있습니다. 이들 예제에는 개념을 가능한 완벽하게 설명하기 위하여 개인, 회사, 상표 및 제품의 이름이 사용될 수 있습니다. 이들 이름은 모두 가공의 것이며 실제 인물 또는 기업의 이름과 유사하더라도 이는 전적으로 우연입니다.

상표

IBM, IBM 로고 및 ibm.com은 전세계 여러 국가에 등록된 International Business Machines Corp.의 상표 또는 등록상표입니다. 기타 제품 및 서비스 이름은 IBM 또는 타사의 상표입니다. 현재 IBM 상표 목록은 웹 "저작권 및 상표 정보"(www.ibm.com/legal/copytrade.shtml)에 있습니다.

Adobe, Adobe 로고, PostScript 및 PostScript 로고는 미국 및/또는 기타 국가에서 사용되는 Adobe Systems Incorporated의 등록상표 또는 상표입니다.

Intel, Intel 로고, Intel Inside, Intel Inside 로고, Intel Centrino, Intel Centrino 로고, Celeron, Intel Xeon, Intel SpeedStep, Itanium 및 Pentium은 미국 또는 기타 국가에서 사용되는 Intel Corporation 또는 그 계열사의 상표 또는 등록상표입니다.

Linux는 미국 또는 기타 국가에서 사용되는 Linus Torvalds의 등록상표입니다.

Microsoft, Windows, Windows NT 및 Windows 로고는 미국 또는 기타 국가에서 사용되는 Microsoft Corporation의 상표입니다.

UNIX는 미국 및 기타 국가에서 사용되는 The Open Group의 등록상표입니다.

Java 및 모든 Java 기반 상표와 로고는 Oracle 및/또는 그 계열사의 상표 또는 등록상표입니다.

제품 문서의 이용 약관

다음 이용 약관에 따라 이 책을 사용할 수 있습니다.

적용성

본 이용 약관은 IBM 웹 사이트의 모든 이용 약관에 추가됩니다.

개인적 사용

모든 소유권 사항을 표시하는 경우에 한하여 귀하는 이 책을 개인적, 비상업적 용도로 복제할 수 있습니다. 귀하는 IBM의 명시적 동의 없이 본 발행물 또는 그 일부를 배포 또는 전시하거나 2차적 저작물을 만들 수 없습니다.

상업적 사용

모든 소유권 사항을 표시하는 경우에 한하여 귀하는 이 책을 귀하 기업집단 내에서만 복제, 배포 및 전시할 수 있습니다. 귀하는 귀하의 기업집단 외에서는 IBM의 명시적 동의 없이 이 책의 2차적 저작물을 만들거나 이 책 또는 그 일부를 복제, 배포 또는 전시할 수 없습니다.

권한

본 허가에서 명시적으로 부여된 경우를 제외하고, 이 책이나 이 책에 포함된 정보, 데이터, 소프트웨어 또는 기타 지적 재산권에 대한 어떠한 허가나 라이센스 또는 권한도 명시적 또는 묵시적으로 부여되지 않습니다.

IBM은 이 책의 사용이 IBM의 이익을 해친다고 판단되거나 위에서 언급된 지시사항이 준수되지 않는다고 판단하는 경우 언제든지 부여한 허가를 철회할 수 있습니다.

귀하는 미국 수출법 및 관련 규정을 포함하여 모든 적용 가능한 법률 및 규정을 철저히 준수하는 경우에만 본 정보를 다운로드, 송신 또는 재송신할 수 있습니다.

IBM은 이 책의 내용과 관련하여 아무런 보장을 하지 않습니다. 타인의 권리 침해, 상품성 및 특정 목적에의 적합성에 대한 묵시적 보증을 포함하여(단, 이에 한하지 않음) 묵시적이든 명시적이든 어떠한 종류의 보증 없이 현 상태대로 제공합니다.

색인

특수 문자

'계속' 표시
피벗 테이블 114
@BLANK 함수 125, 150, 173
@DIFF 함수 167
@FIELD 함수 125, 173
@FIELDS_BETWEEN 함수 125, 132, 173
@FIELDS_MATCHING 함수 125, 132, 173
@INDEX 함수 167
@LAST_NON_BLANK 함수 167, 173
@MAX 함수 167
@MEAN 함수 167
@MIN 함수 167
@MULTI_RESPONSE_SET 함수 133, 173
@NULL 함수 125, 150, 173
@OFFSET 함수
 성능 고려사항 220
@PARTITION_FIELD 함수 173
@PREDICTED 함수 173
@SDEV 함수 167
@SINCE 함수 167
@SUM 함수 167
@TARGET 함수 173
@TESTING_PARTITION 함수 173
@THIS 함수 167
@TODAY 함수 163
@TRAINING_PARTITION 함수 173
@VALIDATION_PARTITION 함수 173

숫자

3차원 막대형 차트 60
3차원 산점도 83
3차원 차트 59
508 준수 221

A

abs 함수 153
allbutfirst 함수 158
allbutlast 함수 158
alphabefore 함수 158
and 연산자 153
arccos 함수 155
arccosh 함수 155
arcsin 함수 155
arcsinh 함수 155
arctan 함수 155
arctan2 함수 155
arctanh 함수 155
area
 공간 함수 156
area 함수 156

bitwise 함수 156
BMP 파일
 차트 내보내기 98, 103

cdf_chisq 함수 155
cdf_f 함수 155
cdf_normal 함수 155
cdf_t 함수 155
Champion Challenger 분석 175, 186
CLEM
 데이터 유형 143, 144
 소개 19, 127
 언어 143
 예 127
 표현식 129, 143
 표현식 검사 138
 표현식 작성 134
 함수 134, 135
CLEM 표현식
 성능 220
CLEM 표현식 검사 138
CLEM 표현식에서 백슬래시 문자 144
CLEM 함수
 결측값 125
 공간 156
 공백 및 널 173
 글로벌 172
 날짜/시간 163
 논리 153
 랜덤 157
 문자열 158
 변환 150
 비교 151
 사용 가능한 목록 148
 삼각 155
 숫자 153
 시퀀스 167
 정보 150
 특수 함수 173
 획률 155
 bitwise 156
close_to
 공간 함수 156
close_to 함수 156
Cognos 활성 보고서 100
COP 9
COP에서 연결 검색 9
cos 함수 155
cosh 함수 155
count_equal 함수 132, 151
count_greater_than 함수 132, 151
count_less_than 함수 132, 151
count_non_nulls 함수 151

B

bitwise 함수 156
BMP 파일
 차트 내보내기 98, 103

cdf_chisq 함수 155
cdf_f 함수 155
cdf_normal 함수 155
cdf_t 함수 155
Champion Challenger 분석 175, 186
CLEM

 데이터 유형 143, 144
 소개 19, 127
 언어 143
 예 127
 표현식 129, 143
 표현식 검사 138
 표현식 작성 134
 함수 134, 135

CLEM 표현식
 성능 220
CLEM 표현식 검사 138
CLEM 표현식에서 백슬래시 문자 144
CLEM 함수

 결측값 125
 공간 156
 공백 및 널 173
 글로벌 172
 날짜/시간 163
 논리 153
 랜덤 157
 문자열 158
 변환 150
 비교 151
 사용 가능한 목록 148
 삼각 155
 숫자 153
 시퀀스 167
 정보 150
 특수 함수 173
 획률 155
 bitwise 156

close_to
 공간 함수 156

close_to 함수 156

Cognos 활성 보고서 100

COP 9

COP에서 연결 검색 9

cos 함수 155

cosh 함수 155

count_equal 함수 132, 151

count_greater_than 함수 132, 151

count_less_than 함수 132, 151

count_non_nulls 함수 151

count_not_equal 함수 132, 151
count_nulls 함수 125, 132, 151
count_substring 함수 158
CRISP-DM
 프로젝트 보기 197
CRISP-DM 프로세스 모델 23
crosses
 공간 함수 156
crosses 함수 156

D

date_before 함수 151
datetime_date 함수 150
DIFF 함수 167
distance
 공간 함수 156
distance 함수 156
div 함수 153
DTD 193

E

endstring 함수 158
EPS 파일
 차트 내보내기 98, 104
equals 연산자 151
Excel 형식
 출력 내보내기 98, 100

F

f 분포
 확률 함수 155
first_index 함수 133, 151
first_non_null 함수 133, 151
first_non_null_index 함수 133, 151
fracof 함수 153

G

greater than 연산자 151

H

hasendstring 함수 158
hasmidstring 함수 158
hasstartswith 함수 158
hassubstring 함수 158
HTML
 출력 내보내기 98, 99
HTML 출력
 스크린 리더 229

I

IBM SPSS Analytic Server
 다중 연결 9
 연결 9
IBM SPSS Collaboration and Deployment Services 175
IBM SPSS Collaboration and Deployment Services Repository

IBM SPSS Collaboration and Deployment Services Repository (계속)
 검색 위치 181
 신임 정보 176
 싱글 사인온 176
 연결 176
 오브젝트 검색 180
 오브젝트 및 버전 삭제 183
 오브젝트 잠금 및 잠금 해제 183
 오브젝트 저장 177
 오브젝트 특성 184
 찾아보기 176
 폴더 183, 184
 프로젝트 전송 199

IBM SPSS Collaboration and Deployment Services Repository 오브젝트 잠금 183
IBM SPSS Collaboration and Deployment Services Repository 오브젝트 잠금 해제 183
IBM SPSS Collaboration and Deployment Services Repository에 오브젝트 저장 177
IBM SPSS Collaboration and Deployment Services Repository에서 오브젝트 검색 180, 181

IBM SPSS Modeler
 개요 7, 205
 내게 필요한 옵션 기능 221
 명령행에서 실행 7
 문서 3
 시작하기 7
 옵션 205
 팁 및 단축키 54

IBM SPSS Modeler Advantage 175

IBM SPSS Modeler Server
 도메인 이름(Windows) 8
 비밀번호 8
 사용자 ID 8
 포트 번호 8
 호스트 이름 8

IBM SPSS Modeler Server 연결 추가 8, 9

IBM SPSS Modeler Server에 로그인 8

if, then, else 함수 153

INDEX 함수 167
integer_bitcount 함수 156
integer_leastbit 함수 156
integer_length 함수 156
intof 함수 153
is_date 함수 150
is_datetime 함수 150
is_integer 함수 150
is_number 함수 150
is_real 함수 150
is_string 함수 150
is_time 함수 150
is_timestamp 함수 150
isalphacode 함수 158
isendstring 함수 158
islowercode 함수 158
ismidstring 함수 158
isnumbercode 함수 158
isstartswith 함수 158
issubstring 함수 158
issubstring_count 함수 158
issubstring_lim 함수 158
isuppercode 함수 158

J

Java 230
JAWS 221, 229, 230
JPEG 파일
 차트 내보내기 98, 103

K

K-평균 노드
 성능 219
 큰 변수군 35

L

last_index 함수 133, 151
LAST_NON_BLANK 함수 167
last_non_null 함수 133, 151
last_non_null_index 함수 133, 151
length 함수 158
less than 연산자 151
locchar 함수 158
locchar_back 함수 158
log 함수 153
log10 함수 153
lowerToUpper 함수 158

M

matches 함수 158
max 함수 151
MAX 함수 167
max_index 함수 133, 151
max_n 함수 132, 151
MEAN 함수 167
mean_n 함수 132, 153
min 함수 151
MIN 함수 167
min_index 함수 133, 151
min_n 함수 132, 151
mod 함수 153

N

negate 함수 153
not 연산자 153
not equal 연산자 151
num_points
 공간 함수 156
num_points 함수 156

O

OFFSET 함수 167
oneof 함수 157
or 연산자 153
overlap
 공간 함수 156
overlap 함수 156

P

PDF
 출력 내보내기 98, 101
pi 함수 155
PMML
 내보내기 옵션 208
 모델 가져오기 193, 194
 모델 내보내기 193, 194
PMML 모델
 로지스틱 회귀분석 208
 선형 회귀 208
PNG 파일
 차트 내보내기 98, 104
PowerPoint
 PowerPoint로 출력 내보내기 101
PowerPoint 형식
 출력 내보내기 98

Q

Q-Q 도표 80

R

random0 함수 157
rem 함수 153
replace 함수 158
replicate 함수 158
round 함수 153

S

SAS 파일
 인코딩 233
SDEV 함수 167
sdev_n 함수 132, 153
sign 함수 153
sin 함수 155
SINCE 함수 167
sinh 함수 155
skipchar 함수 158
skipchar_back 함수 158
soundex 함수 163
soundex_difference 함수 163
SPLOMs 83
SQL 생성
 로깅 38
 미리보기 38
sqrt 함수 153
startstring 함수 158
stripchar 함수 158
strmember 함수 158
subscr 함수 158
substring 함수 158
substring_between 함수 158
SUM 함수 167
sum_n 함수 132, 153

T

t 분포
 확률 함수 155

t-SNE 차트 90
tan 함수 155
tanh 함수 155
testbit 함수 156
THIS 함수 167
TIFF 파일
 차트 내보내기 98, 104
time_before 함수 151
to_date 함수 150, 163
to_dateline 함수 163
to_datetime 함수 150
to_integer 함수 150
to_number 함수 150
to_real 함수 150
to_string 함수 150
to_time 함수 150, 163
to_timestamp 함수 150, 163
trim 함수 158
trim_start 함수 158
trimend 함수 158

U

undef 함수 173
unicode_char 함수 158
unicode_value 함수 158
uppertolower 함수 158
URL
 IBM SPSS Analytic Server 9
UTF-8 인코딩 35, 233

V

value_at 함수 133, 151

W

within
 공간 함수 156
within 함수 156
Word 형식
 넓은 테이블 98
 출력 내보내기 98, 100

가

가져오기
 PMML 193, 194
값
 데이터 검토로부터 보기 138
 CLEM 표현식에 추가 138
값 레이블
 개요 분할창 120
 피벗 테이블 120
개요
 뷰어 94
 수준 변경 95
 접기 95
 펼치기 95
거듭제곱(지수) 함수 153
거품형 차트 63
검색
 스트림에서 노드 43

검색 및 바꾸기
 뷰어 문서 96
격자선
 피벗 테이블 116
결측값
 레코드에 122
 시스템 122
 CLEM 표현식 125
경계선
 숨겨진 경계선 표시 116
경고
 옵션 설정 206
고유 노드
 성능 218
공간 함수 156
공백
 문자열에서 제거 130, 158
공백 처리
 CLEM 함수 173
관계 차트 82
관리자 13
구간화 노드
 성능 218
규칙 149
규칙 세트
 평가 35
규칙 작성 노드
 로드 52
그래프
 출력 저장 52
 프로젝트에 추가 198
그룹 기호
 숫자 표시 형식 35
그룹 레이블 108
그룹 레이블 삽입 108
그룹 레이블 제거 108
그룹 레이블 추가 108
그룹 선형 차트 71
그룹화 산점도 83
글꼴
 개요 영역 95
글로벌 값
 CLEM 표현식에서 138
글로벌 함수 172
기본 창 11
기본값
 프로젝트 단계 197
꼬리말
 번호 재지정 115
 표식 113

나

날짜
 변환 167
 조작 167
날짜 함수
 @TODAY 함수 163
 date_before 151, 163
 date_days_difference 163
 date_in_days 163
 date_in_months 163
 date_in_weeks 163
 date_in_years 163

날짜 함수 (계속)
date_months_difference 163
date_weeks_difference 163
date_years_difference 163
날짜 형식 36, 144, 145
날짜/시간 값 131
날짜/시간 함수
datetime_date 163
datetime_day 163
datetime_day_name 163
datetime_day_short_name 163
datetime_hour 163
datetime_in_seconds 163
datetime_minute 163
datetime_month 163
datetime_month_name 163
datetime_month_short_name 163
datetime_now datetime_second 163
datetime_time 163
datetime_timestamp 163
datetime_weekday 163
datetime_year 163
내게 필요한 옵션
예제 228
IBM SPSS Modeler에서의 팁 230
IBM SPSS Modeler의 기능 221
내보내기
스트림 설명 45
PMML 193, 194
너깃
정의된 13
널 131
넓은 테이블
Microsoft Word에 붙여넣기 97
노드
검색 43
데이터 미리보기 33
로드 52
사용 31
사용 안함 31, 32
사용자 정의 팔레트 작성 214
사용자 정의 하위 팔레트 작성 215
삭제 30
성능 218, 219
소개 29
순서 217
스트림에서 사용 안함 31
스트림에서 연결 30
스트림에서 우회 31
실행 시간 41
연결 삭제 31
옵션 설정 32
잡금 34
저장 51
주석 46, 50
주석 추가 46
중복 32
추가 30, 31
팔레트 탭 사용자 정의 216
팔레트에 표시 214
팔레트에서 제거 214
편집 32
프로젝트에 추가 198

노드 (계속)
IBM SPSS Collaboration and Deployment Services Repository에 저장 179
노드 사용 31
노드 사용 안함 31, 32
노드 이름 50
노드 잡금 34
노드 팔레트 선택 214
노드와 스트림 이름 지정 50
논리 함수 153
누적 막대형 차트 60
누적 오버플로우 오류 205
다
다른 애플리케이션으로 출력 복사 및 붙여넣기 97
다른 애플리케이션으로 출력 붙여넣기 97
다중 계열 차트 77
다중 데이터 파일 열기 118
다중 범주형 세트
CLEM 표현식에서 133
다중 선형 차트 71
다중 이분형 세트
CLEM 표현식에서 133
다중 차트 76
다중 축 차트 65
다중반응 세트
CLEM 표현식에서 133, 138
단어 클라우드 차트 90
단축키
일반 사용법 54
키보드 18, 222, 225-227
대시보드
차트 91
대화 상자
변수 레이블 표시 118
변수 이름 표시 118
변수 표시 순서 118
대화식 트리 창
내게 필요한 옵션 230
대화형 출력 98
데이터
미리보기 33
데이터 검토 노드
데이터 마이닝에 사용 22
탐색에서 사용 21
데이터 마이닝
애플리케이션 예제 28
전략. 23
데이터 마이닝의 애플리케이션 22
데이터 맵핑 54
데이터 맵핑 도구 53
데이터 스트림
작성 29
데이터 유형
모수 42
데이터 파일
다중 데이터 파일 열기 118
데이터 편집기
기술통계 옵션 120
다중 데이터 파일 열기 118
데이터베이스
함수 134, 135
데이터베이스 함수

데이터베이스 함수 (계속)
 사용자 정의 함수(UDF) [135](#)
 CLEM 표현식에서 [135](#)
도구 모음 [15](#)
도구 팁
 노드에 주석달기 [50](#)
도메인 이름(Windows)
 IBM SPSS Modeler Server [8](#)
디렉토리
 기본 [206](#)

라

라디안
 측정 단위 [37](#)
랜덤 함수 [157](#)
레거시 테이블 [118](#)
레이블
 값 [193](#)
 그룹 레이블 삽입 [108](#)
 변수 [193](#)
 삭제 [108](#)
 표시 [35](#)
레이블 회전 [108](#)
레이블, IBM SPSS Collaboration and Deployment Services Repository 오브젝트 [185](#)
레이어
 인쇄 [104, 112, 114](#)
 작성 [110](#)
 표시 [110](#)
 피벗 테이블 [110](#)
레코드
 결측값 [122](#)
 시스템 결측값 [122](#)
로그 파일
 생성된 SQL 표시 [38](#)
로드
 노드 [52](#)
 상태 [52](#)
로지스틱 회귀분석
 PMML로서 내보내기 [208](#)
로케일
 옵션 [205](#)
롤오버 일 수 [36](#)

마

마우스
 IBM SPSS Modeler에서 사용 [18, 30](#)
마우스 가운데 단추
 시뮬레이션 [18, 30](#)
마침표 [35](#)
막대형 차트 [60](#)
맞추기
 출력 [94, 119](#)
맞춤
 출력 [94, 119](#)
맵 서비스 [74](#)
맵 차트 [72](#)
맵핑 필드 [53](#)
머리말 [105](#)
메모리
 관리 [205](#)

메모리 (계속)
 누적 오버플로우 오류 [205](#)
메모리 부족 [205](#)
메시지
 생성된 SQL 표시 [38](#)
메타파일
 차트 내보내기 [98](#)
멤버 함수 [151](#)
명령행
 IBM SPSS Modeler 시작 [7](#)
모델
 내보내기 [208](#)
 바꾸기 [206](#)
 새로 고침 [188](#)
 프로젝트에 추가 [198](#)
 IBM SPSS Collaboration and Deployment Services Repository에 저장 [180](#)
모델 너깃 [46](#)
모델 바꾸기 [206](#)
모델 새로 고침 [186, 188](#)
모델 팔레트 [180](#)
모델링
 분기 [46](#)
모델링 노드
 모델링 팔레트 탭 사용자 정의 [216](#)
 성능 [219](#)
모수
 런타임 프롬프트 [42](#)
 모델 작성 [187](#)
 세션 [42](#)
 스코어링 [187](#)
 스트림 [42](#)
 유형 [42](#)
 CLEM 표현식에서 [138](#)
목록 [143, 144](#)
문서 [3](#)
문자 [143, 144](#)
문자열
 매치 [130](#)
 바꾸기 [130](#)
 CLEM 표현식에서 조작 [130](#)
문자열 연결 [150](#)
문자열 함수 [158](#)
미리보기
 노드 데이터 [33](#)

바

방사형 차트 [81](#)
배경 색상 [114](#)
배율
 피벗 테이블 [112, 114](#)
배포 [175](#)
배포 옵션 [186](#)
배포 유형 [186](#)
백업 스트림 파일
 복원 [51](#)
버전 레이블, IBM SPSS Collaboration and Deployment Services Repository 오브젝트 [185](#)
변수
 대화 상자에 표시 순서 [118](#)
변수 레이블
 개요 분할창 [120](#)
 대화 상자 [118](#)

변수 레이블 (계속)
 피벗 테이블 [120](#)
 변수 이름
 대화 상자 [118](#)
 변수군 [35](#)
 변환 함수 [150](#)
 보고서
 생성 [201](#)
 출력 저장 [52](#)
 특성 설정 [201](#)
 프로젝트에 추가 [198](#)
 보기 렌더링 오류
 메모리 부족 [205](#)
 보기와 맞춰 스트림 스케일링 [17](#)
 복사 [15](#)
 본질적 필드 [53, 54](#)
 분기, 모델링 및 스코어링 [46, 187–189](#)
 분포 함수 [155](#)
 붙여넣기 [15](#)
 뷰어
 값 레이블 표시 [120](#)
 개요 [94](#)
 개요 글꼴 변경 [95](#)
 개요 분할창 [93](#)
 개요 수준 변경 [95](#)
 개요 접기 [95](#)
 개요 크기 변경 [95](#)
 개요 펼치기 [95](#)
 결과 분할창 [93](#)
 결과 숨김 [93](#)
 데이터 값 표시 [120](#)
 문서 저장 [106](#)
 변수 레이블 표시 [120](#)
 변수 이름 표시 [120](#)
 정보 검색 및 바꾸기 [96](#)
 정보 찾기 및 바꾸기 [96](#)
 출력 삭제 [94](#)
 출력 항목 간의 공백 [105](#)
 표시 옵션 [119](#)
 비교 함수 [151](#)
 비밀번호
 IBM SPSS Analytic Server [9](#)
 IBM SPSS Modeler Server [8](#)
 비주얼 프로그래밍 [11](#)
 빈도 다각형 [70](#)
 빠른 피벗 테이블 [119](#)

사

사용자 옵션 [206](#)
 사용자 정의 차트 [65](#)
 사용자 정의 팔레트 작성
 하위 팔레트 작성 [215](#)
 사용자 정의 함수(UDF) [134, 135](#)
 사용자 ID
 IBM SPSS Modeler Server [8](#)
 사전 파일 [229](#)
 산점도
 1차원 [83](#)
 3차원 [83](#)
 그룹 [83](#)
 단순 [83](#)
 오버레이 [83](#)
 점도표 [83](#)

산점도 (계속)
 행렬 [83, 84](#)
 삼각 함수 [155](#)
 상자도표 [62](#)
 상태
 로드 [52](#)
 저장 [51](#)
 새로 고침
 소스 노드 [35](#)
 새로운 기능 [5](#)
 색상
 설정 [208](#)
 생산 기능
 저널 파일의 명령 구문 사용 [118](#)
 생성된 모델 팔레트 [13](#)
 서버
 기본 디렉토리 [206](#)
 로그인 [8](#)
 연결 추가 [8](#)
 COP에서 서버 검색 [9](#)
 선택 방법
 피벗 테이블에서 행과 열 선택 [116](#)
 선택하여 복사 [97](#)
 선형 차트
 시간 도표 [86](#)
 하락-선 [83, 84](#)
 선형 회귀
 PMML로서 내보내기 [208](#)
 성능
 모델링 노드 [219](#)
 프로세스 노드 [218](#)
 CLEM 표현식 [220](#)
 세션 모수 [42](#)
 셀 특성 [114](#)
 소개
 IBM SPSS Modeler [7, 205](#)
 소수점 기호
 숫자 표시 형식 [35](#)
 소수점이하자리수
 표시 형식 [37](#)
 소스 노드
 데이터 맵핑 [53](#)
 새로 고침 [35](#)
 속성 [21](#)
 수직 레이블 텍스트 [108](#)
 수평배열 막대형 차트 [60](#)
 수학 곡선 차트 [75](#)
 숨김
 꼬리말 [115](#)
 제목 [111](#)
 차원 레이블 [111](#)
 캡션 [115](#)
 프로시저 결과 [93](#)
 행 및 열 [110](#)
 숫자 [131, 143](#)
 숫자 표시 형식 [37](#)
 숫자 함수 [153](#)
 쉼표 [35](#)
 스코어링
 분기 [46, 187–189](#)
 스크롤링
 옵션 설정 [39](#)
 스크린 리더
 예제 [228](#)

스크립트 색상
설정 208

스크립팅 19, 127

스트림
노드 사용 안함 31
노드 연결 30
노드 우회 31
노드 추가 30, 31
로드 52
배포 옵션 186
백업 파일 51
보기 스케일링 17
실행 45
실행 시간 보기 41
옵션 34-39
이름 변경 44, 50
작성 29
저장 51
주석 46, 50
주석 추가 46
지리 공간적 좌표계 40
프로젝트에 추가 198
IBM Cloud Pak for Data에 저장 191
IBM SPSS Collaboration and Deployment Services Repository에 저장 179

스트림 기본 인코딩 35

스트림 모수 42

스트림 설명 44, 45

스트림 실행 45

스트림 이름 50

스트림 캔버스
설정 39

스트림 특성
Analytic Server 39

스트림에 대한 주석 모두 나열 49

시간 도표 86

시간 및 날짜 함수 144, 145

시간 필드
변환 167

시간 함수
time_before 151, 163
time_hours_difference 163
time_in_hours 163
time_in_mins 163
time_in_secs 163
time_mins_difference 163
time_secs_difference 163

시간 형식 36, 144, 145

시스템
옵션 205

시스템 결측값
레코드에 122

시스템 학습 21

시작 대화 상자 208

시퀀스 함수 167

신경망 노드
성능 219
큰 변수군 35

신임 정보
IBM SPSS Collaboration and Deployment Services Repository 176

실수 143

실행 시간, 보기 41

실행 중단 15

실행 취소 15

싱글 사인온 8

싱글 사인온, IBM SPSS Collaboration and Deployment Services Repository 175, 176

아

아이콘
옵션 설정 17, 39

알림
옵션 설정 206

애플리케이션 21

애플리케이션 예제 3

언어
옵션 205
출력 언어 변경 109

여러 IBM SPSS Modeler 세션 10

연결
서버 군집 9
IBM SPSS Analytic Server 9
IBM SPSS Collaboration and Deployment Services Repository에 176
IBM SPSS Modeler Server로 8, 9

연산자
문자열 결합 150
CLEM 표현식에서 134

연산자 우선순위 146

열
파벗 테이블에서 너비 변경 116
파벗 테이블에서 선택 116

열 너비
기본 너비 제어 119
줄 바꾼 텍스트의 너비 제어 112
최대 너비 제어 112
파벗 테이블 116

열 차트 60

열기
노드 52
모델 52
상태 52
스트림 52
출력 52
프로젝트 198

예제
개요 3
애플리케이션 안내서 3

오류 메시지 40

오버레이 산점도 83

오브젝트
특성 200

오브젝트 특성, IBM SPSS Collaboration and Deployment Services Repository 184

오차 막대 차트 60, 66

옵션
구문 208
데이터 편집기의 기술통계량 120
뷰어 119
사용자 206
스트림 특성 34-41
일반 118
출력 레이블 120
표시 208
파벗 테이블 모양 119
IBM SPSS Modeler용 205

옵션 (계속)
 PMML 208
요약 포인트 도표 83
요인 229
우선순위 146
원 채우기 차트 64
원통형 차트 63
원형 차트 78
웹 보고서
 출력 내보내기 100
유니코드 지원 233
유형 노드
 결측값 124
 성능 218
의사결정 트리
 내게 필요한 옵션 230
이름 변경
 노드 50
 스트림 44
이중 Y축 차트 65
이중축 차트 65
인구 피라미드 70
인구 피라미드 차트 79
인쇄
 레이어 104, 112, 114
 머리말과 꼬리말 105
 스트림 17, 32
 인쇄 미리보기 104
 차트 104
 차트 크기 105
 출력 항목 간의 공백 105
 데이터 구분 제어 117
 데이터 배율 112, 114
 텍스트 출력 104
 페이지 번호 105
 피벗 테이블 104
인코딩 35, 233
일반 애플리케이션 21
임시 디렉토리 10

자

자동화 127
잘라내기 15
잡음 데이터 22
저장
 노드 51
 다중 오브젝트 51
 상태 51
 스트림 51
 출력 오브젝트 52
정규 분포
 확률 함수 155
정렬
 피벗 테이블 행 108
정렬 노드
 성능 218
정보 함수 150
정수 143
제목
 뷰어에 추가 95
조건 129
주석
 노드 46, 50

주석 (계속)
 노드와 스트림 46
 스트림 46, 50
 스트림에 모두 나열 49
 키보드 단축키 226, 227
주석(Annotation)
 주석으로 변환 50
 풀더 200
 프로젝트 200
줄바꿈
 줄 바꾼 텍스트의 열 너비 제어 112
지리 공간적 좌표
 시스템 선택 40
 표시 형식 39
지리 공간적 좌표 형식 39
지리 공간적 좌표계 선택 40
지수 표기법
 결과에 출력않음 118
 표시 형식 37
지수 함수 153
지식 검색 21

차

차수
 측정 단위 37
차트
 3차원 59
 거품형 63
 관계 82
 내보내기 98
 다중 계열 77
 다중 차트 76
 다중 축 65
 단어 클라우드 90
 대시보드 91
 막대 60
 맵 72, 74
 방사형 81
 빈도 다각형 70
 사용자 정의 65
 산점도 83
 상자도표 62
 선 71
 수학 곡선 75
 숨김 93
 시간 도표 86
 열 60
 오차 막대 60, 66
 요약 포인트 도표 83
 원 78
 원 채우기 64
 원통형 63
 이중 Y축 65
 인구 피라미드 70, 79
 점도표 83
 출력 저장 52
 테마 리버 87
 템플리트 91
 트리 88
 트리 맵 89
 평가 67
 평형 78
 피벗 테이블에서 작성 118

차트 (계속)
하락-선 83
햇살 85
히스토그램 70
히트 맵 69
Q-Q 도표 80
t-SNE 90
차트 내보내기 98, 103, 104
차트 작성기
 갤러리 58
 레이아웃 57
 용어 57
차트 저장
 메타파일 98
 포스트스크립트 파일 104
 BMP 파일 98, 103
 EMF 파일 98
 EPS 파일 98, 104
 JPEG 파일 98, 103
 PICT 파일 98
 PNG 파일 104
 TIFF 파일 104
찾기 및 바꾸기
 뷰어 문서 96
채움 노드
 결측값 124
최소화 16
추가
 프로젝트에 198
출력
 가운데 맞춤 94, 119
 내보내기 98
 다른 애플리케이션으로 붙여넣기 97
 대화식 98
 맞추기 94, 119
 복사 94
 뷰어 93
 삭제 94
 숨김 93
 암호화 106
 이동 94
 저장 106
 출력 언어 변경 109
 표시 93
출력 내보내기
 웹 보고서 100
 Excel 형식 98, 100
 HTML 99
 HTML 형식 98
 PDF 형식 98, 101
 PowerPoint 형식 98
 Word 형식 98, 100
출력 노드 29
출력 삭제 94
출력 오브젝트
 IBM SPSS Collaboration and Deployment Services Repository에 저장 180
출력 저장
 웹 보고서 100
 텍스트 형식 98, 102
 Excel 형식 98, 100
 HTML 98, 99
 HTML 형식 98
 PDF 형식 98, 101

출력 저장 (계속)
 PowerPoint 형식 98, 101
 Word 형식 98, 100
출력 탭
 출력 이동 94
출력 파일
 저장 52
 출력을 가운데 맞춤 94, 119
 측정 시스템 118

카

카이제곱 분포
 확률 함수 155
캐시
 비우기 32, 35
 사용 205
 저장 32
 캐시 설정 32
캐시 파일 노드
 로드 52
캔버스 11
캡션 114, 115
캡슐화된 포스트스크립트 파일
 차트 내보내기 98, 104
케이스 21
코호넨 노드
 성능 219
 큰 변수군 35
크기
 개요 95
 크기 조정 16
클라이언트
 기본 디렉토리 206
클래스 14, 197, 198
키보드 단축키 222, 225–227
키워드
 노드에 주석달기 50

타

탐색
 키보드 단축키 222
터미널 노드 29
테넌트
 IBM SPSS Analytic Server 9
테마 리버 차트 87
테이블
 글꼴 114
 맞추기 114
 배경 색상 114
 셀 특성 114
 여백 114
 출력 저장 52
 테이블 구분 제어 117
 프로젝트에 추가 198
테이블 구분 117
테이블 분할
 테이블 구분 제어 117
테이블 차트 118
테이블 모양
 작성 111
 적용 111

텍스트
 뷰어에 추가 95
 뷰어에 텍스트 파일 추가 96
 텍스트로 출력 내보내기 98, 102
텍스트 데이터 파일
 인코딩 233
텍스트 인코딩 35
템플리트
 차트 91
 템플리트 펠드 54
통계량 파일
 인코딩 233
통합 노드
 성능 218
통화 표시 형식 37
트리 기반 분석
 일반 애플리케이션 21
트리 맵 차트 89
트리 차트 88
특성
 데이터 스트림 34
 보고서 단계 201
 테이블 112
 프로젝트 폴더 200
 피벗 테이블 112
특수 문자
 문자열에서 제거 130
특수 함수 173
팁
 내가 필요한 옵션에 대한 230
 일반 사용법 54

파

파워포인트 파일 198
파일
 뷰어에 텍스트 파일 추가 96
팔레트
 사용자 정의 214
팔레트 템 사용자 정의 216
페이지 번호 매김 105
페이지 설정
 머리말과 꼬리말 105
 차트 크기 105
평가 노드
 성능 218
평가 차트 67
평형 차트 78
포트 번호
 IBM SPSS Modeler Server 8
폴더, IBM SPSS Collaboration and Deployment Services Repository 183, 184
표시
 결과 93
 꼬리말 115
 제목 111
 차원 레이블 111
 캡션 115
 행 또는 열 111
 표시 순서 107, 108
 표시 형식
 그룹 기호 37
 소수점이하자리수 37
 숫자 37

표시 형식 (계속)
 지리 공간적 좌표 39, 40
 지수표기 37
 통화 37
표시할 행 수 제어 112
표현식 143
표현식 작성기
 개요 134
 사용 134
 액세스 134
품질 노드
 결측값 122
프로세스 노드
 성능 218
프로세스 조정자 9
프로젝트
 기본 폴더 설정 197
 닫기 201
 보고서 생성 201
 새로 작성 198
 오브젝트 추가 198
 오브젝트 특성 200
 작성 198
 주석 200
 클래스 보기 198
 특성 설정 199
 폴더 특성 200
 CRISP-DM 보기 197
 IBM SPSS Collaboration and Deployment Services Repository에 저장 179
 IBM SPSS Collaboration and Deployment Services Repository에서 199
프롬프트, 런타임 42
피벗 테이블
 '계속' 표시 114
 값 레이블 109
 격자선 116
 경계선 113
 그룹 레이블 삭제 108
 그룹 레이블 삽입 108
 글꼴 114
 기본값 열 너비 조정 119
 꼬리말 114-116
 꼬리말 특성 113
 다른 애플리케이션으로 붙여넣기 97
 레거시 테이블 118
 레이블 회전 108
 레이어 110
 레이어 인쇄 104
 맞추기 114
 모양 변경 111
 배경 색상 114
 변경 실행 취소 110
 변수 레이블 109
 빠른 피벗 테이블 119
 새 테이블의 기본값 모양 119
 셀 너비 116
 셀 특성 114
 셀 표시 및 숨김 110
 셀 형식 113
 숨겨진 경계선 표시 116
 숨김 93
 아이콘 사용 107
 언어 109

피벗 테이블 (계속)
 여백 [114](#)
 일반 특성 [112](#)
 조작 [107](#)
 캡션 [114, 115](#)
 큰 테이블 인쇄 [117](#)
 테이블 구분 제어 [117](#)
 테이블로 붙여넣기 [97](#)
 테이블에서 차트 작성 [118](#)
 테이블을 더 빠르게 렌더링 [119](#)
 특성 [112](#)
 페이지에 맞춤 [112, 114](#)
 편집 [107](#)
 표시 순서 변경 [107, 108](#)
 표시할 행 수 제어 [112](#)
 피벗 [107](#)
 행 또는 열 그룹화 [108](#)
 행 및 열 삽입 [109](#)
 행 색상 변경 [113](#)
 행 정렬 [108](#)
 행과 열 선택 [116](#)
 행과 열 이동 [108](#)
 행과 열 전치 [108](#)
 행이나 열 그룹 해제 [108](#)
 HTML로 내보내기 [98](#)

피벗 테이블의 색상
 경계선 [113](#)

피벗 테이블의 셀
 너비 [116](#)
 선택 [116](#)
 숨김 [110](#)
 표시 [110](#)
 형식 [113](#)

필드
 값 보기 [138](#)
 CLEM 표현식에서 [138](#)

필드 맵핑 해제 [53](#)

필드선택 노드
 결측값 [122](#)

필수 필드 [54](#)

하

하락-선 차트 [83, 84](#)

하위 팔레트
 작성 [215](#)
 팔레트 탭에 표시 [215](#)
 팔레트 탭에서 제거 [215](#)

함수
 @BLANK [124](#)
 @FIELD [134, 173](#)
 @GLOBAL_MAX [172](#)
 @GLOBAL_MEAN [172](#)
 @GLOBAL_MIN [172](#)
 @GLOBAL_SDEV [172](#)
 @GLOBAL_SUM [172](#)
 @PARTITION [173](#)
 @PREDICTED [134, 173](#)
 @TARGET [134, 173](#)
 결측값 처리 [125](#)
 데이터베이스 [134, 135](#)
 사용자 정의 함수(UDF) [134](#)
 예 [127](#)
 CLEM 표현식에서 [134](#)

합치기 노드
 성능 [218](#)

햇살 차트 [85](#)

행
 피벗 테이블에서 선택 [116](#)
 행 또는 열 그룹화 [108](#)
 행 색상 변경
 피벗 테이블 [113](#)
 행과 열 순서 변경 [108](#)
 행과 열 이동 [108](#)
 행과 열 전치 [108](#)
 행렬 산점도 [83, 84](#)
 호스트 이름
 IBM SPSS Modeler Server [8](#)

확대/축소 [15](#)

확률 함수 [155](#)

회귀분석 [229](#)

히스토그램 [70](#)

히트 맵 차트 [69](#)

힌트
 일반 사용법 [54](#)

IBM.[®]