

**IBM SPSS Modeler 18.1 ユー
ザーズ・ガイド**

IBM

注記

本書および本書で紹介する製品をご使用になる前に、229 ページの『特記事項』に記載されている情報をお読みください。

本書は、IBM SPSS Modeler バージョン 18 リリース 0 モディフィケーション 0 および新しい版で明記されない限り、以降のすべてのリリースおよびモディフィケーションに適用されます。

お客様の環境によっては、資料中の円記号がバックスラッシュと表示されたり、バックスラッシュが円記号と表示されたりする場合があります。

原典： IBM SPSS Modeler 18.1 User's Guide

発行： 日本アイ・ビー・エム株式会社

担当： トランスレーション・サービス・センター

目次

第 1 章 IBM SPSS Modeler について . . . 1

IBM SPSS Modeler 製品	1
IBM SPSS Modeler	1
IBM SPSS Modeler Server	2
IBM SPSS Modeler Administration Console	2
IBM SPSS Modeler Batch	2
IBM SPSS Modeler Solution Publisher	2
IBM SPSS Modeler Server Adapters for IBM SPSS Collaboration and Deployment Services	2
IBM SPSS Modeler のエディション	3
IBM SPSS Modeler ドキュメント	3
SPSS Modeler Professional ドキュメント	3
SPSS Modeler Premium ドキュメント	4
アプリケーションの例	4
Demos フォルダ	5
ライセンスの追跡	5

第 2 章 IBM SPSS Modeler 18.1 の新機能 7

第 3 章 IBM SPSS Modeler の概要 9

はじめに	9
IBM SPSS Modeler の開始	9
コマンド・ラインからの起動	10
IBM SPSS Modeler Server に接続中	10
Analytic Server に接続中	12
一時ディレクトリの変更	13
複数の IBM SPSS Modeler セッションの起動	14
IBM SPSS Modeler インターフェースについて	14
IBM SPSS Modeler ストリーム領域	15
ノード・パレット	16
IBM SPSS Modeler マネージャ	17
IBM SPSS Modeler プロジェクト	18
IBM SPSS Modeler ツールバー	19
ツールバーのカスタマイズ	20
IBM SPSS Modeler ウィンドウのカスタマイズ	21
ストリームのアイコン・サイズの変更	22
IBM SPSS Modeler でのマウスの使用方法	22
ショートカット・キーの使用	22
印刷中	24
IBM SPSS Modeler の自動化	24

第 4 章 データ・マイニングについて . . . 25

データ・マイニングの概要	25
データへの評価	26
データ・マイニングの戦略的方法	27
CRISP-DM プロセス・モデル	28
モデルの種類	29
データ・マイニングの例	34

第 5 章 ストリームの構築 35

ストリームの構築の概要	35
データ・ストリームの作成	35
ノードの作業	36
ストリームの作業	41
ストリームの説明	55
ストリームの実行	57
モデルの処理	58
ノードおよびストリームへのコメントおよび注釈 の追加	58
データ・ストリームの保存	64
ファイルのロード	66
データ・ストリームのマッピング	66
ヒントとショートカット	69

第 6 章 出力の処理 71

ビューア	71
結果の表示と非表示	71
出力の移動、削除、およびコピー	72
初期位置合わせの変更	72
出力項目の位置合わせの変更	72
ビューアのアウトライン	73
ビューアへの項目の追加	74
ビューアの情報の検索および置換	74
他のアプリケーションへの出力のコピー	76
インタラクティブ出力	77
出力のエクスポート	78
HTML のオプション	79
Web レポートのオプション	80
Word/RTF のオプション	80
Excel のオプション	81
PowerPoint のオプション	82
PDF のオプション	83
テキストのオプション	83
グラフィックスのみのオプション	84
グラフィックス形式オプション	84
ビューアの印刷	85
出力とグラフを印刷するには	85
印刷プレビュー	86
ページ属性: ヘッダーとフッター	86
ページ属性: オプション	87
出力の保存	87
ビューア文書を保存するには	87
ピボット・テーブル	89
ピボット・テーブル	89
ピボット・テーブルの操作	89
層の処理	92
項目の表示と非表示	93
テーブル・ルック	94
テーブル・プロパティ	94
セル・プロパティ	97

脚注とキャプション	98
データ・セルの幅	100
列の幅の変更	100
ピボット・テーブルの隠れた罫線を表示する	100
ピボット・テーブルの行、列、およびセルの選択	100
ピボット・テーブルの印刷	100
ピボット・テーブルからのグラフの作成	102
レガシー・テーブル	102
オプション	102
オプション	102
全般オプション	103
ビューアーのオプション	103
ピボット・テーブル・オプション	103
出力オプション	104

第 7 章 欠損値の処理 105

欠損値の概要	105
欠損値の処理	106
欠損値を含むレコードの処理	106
欠損値を含むフィールドの処理	106
システム欠損値を含むレコードの処理	107
欠損値の代入または置換	109
欠損値用 CLEM 関数	110

第 8 章 CLEM 式の作成 111

CLEM について	111
CLEM の例	111
値とデータ型	113
式と条件	114
ストリーム、セッション、およびスーパーノード・パラメーター	114
文字列の処理	115
空白および欠損値の処理	116
数値の処理	116
時間と日付の処理	116
複数フィールドの要約	117
複数回答データの処理	118
Clem 式ビルダー	119
Clem 式ビルダーへのアクセス	119
CLEM 式の作成	119
関数の選択	119
フィールド、パラメーター、およびグローバル変数の選択	122
値の表示または選択	123
CLEM 式の検査	123
検索と置換	124

第 9 章 CLEM 言語リファレンス 127

CLEM リファレンス概要	127
CLEM データ型	127
整数値	127
実数	128
文字	128
文字列	128
リスト	128
フィールド	129

日付	129
時間	130
CLEM 演算子	130
関数のリファレンス	132
関数の表記方法について	133
情報関数	134
変換関数	135
比較関数	136
論理関数	138
数値関数	138
三角関数	139
確率関数	140
空間処理関数	140
ビット単位の整数演算	141
乱数関数	142
文字列関数	143
SoundEx 関数	148
日付および時刻の関数	148
シーケンス関数	152
グローバル関数	158
空白値とヌル値処理関数	159
特殊フィールド	160

第 10 章 リポジトリでの IBM SPSS Modeler の使用 163

IBM SPSS Collaboration and Deployment Services Repository について	163
リポジトリ・オブジェクトの保存と展開	164
リポジトリへの接続	164
リポジトリの資格情報の入力	165
リポジトリ資格情報の参照	165
リポジトリ・コンテンツを参照	165
リポジトリのオブジェクトの保存	165
オブジェクトの特性の設定	166
ストリームの保存	168
プロジェクトの格納	168
ノードの格納	169
出力オブジェクトの格納	169
モデルおよびモデル・パレットの格納	169
リポジトリからのオブジェクトの取得	170
取得するオブジェクトの選択	171
オブジェクトのバージョンの選択	171
リポジトリ内のオブジェクトの検索	172
リポジトリのオブジェクトを変更	173
フォルダーの編集、名前の変更、削除	173
リポジトリ・オブジェクトのロックおよびロック解除	173
リポジトリのオブジェクトを削除	174
リポジトリ・オブジェクトのプロパティの管理	174
フォルダーのプロパティの表示	174
オブジェクトのプロパティの表示および編集	175
オブジェクトのバージョン・ラベルの管理	176
ストリームの展開	177
ストリーム展開のオプション	177
スコアリングブランチ	179

第 11 章 外部アプリケーションへのエクスポート	183
外部アプリケーションへのエクスポートについて	183
IBM SPSS Modeler Advantage でストリームを開く	183
PMML としてのモデルのインポートおよびエクスポート	184
PMML をサポートするモデルの種類	185
第 12 章 プロジェクトとレポート	187
プロジェクトの概要	187
CRISP-DM ビュー	187
クラス・ビュー	188
プロジェクトの作成	188
新規プロジェクトの作成	188
プロジェクトへの追加	189
プロジェクト IBM SPSS Collaboration and Deployment Services Repository への転送	190
プロジェクトのプロパティの設定	190
プロジェクトの注釈	191
オブジェクトのプロパティ	191
プロジェクトを閉じる	192
レポートの生成	192
生成されたレポートの保存とエクスポート	194
第 13 章 IBM SPSS Modeler のカスタマイズ	195
IBM SPSS Modeler オプションのカスタマイズ	195
IBM SPSS Modeler オプションの設定	195
システム・オプション	195
デフォルト・ディレクトリの設定	196
ユーザー・オプションの設定	196
ノード・パレットのカスタマイズ	204

パレット・マネージャのカスタマイズ	205
パレット・タブの表示の変更	207

第 14 章 ストリームおよびノードのパフォーマンスの考察	209
ノードの順序	209
ノードのキャッシュ	210
パフォーマンス : プロセス・ノード	211
パフォーマンス : モデル作成ノード	212
パフォーマンス : CLEM 式	213

第 15 章 IBM SPSS Modeler のアクセシビリティ	215
IBM SPSS Modeler のアクセシビリティの概要	215
アクセシビリティサポートの種類	215
視力の弱いユーザーのためのアクセシビリティ	215
視覚障害ユーザーのためのアクセシビリティ	216
キーボード・アクセシビリティ	217
画面読み上げソフトウェアの使用	224
使用上のヒント	225
他のソフトウェアとの相性	226
JAWS および Java	226
IBM SPSS Modeler でのグラフの使用	226

第 16 章 Unicode のサポート	227
IBM SPSS Modeler 内での Unicode のサポート	227

特記事項	229
商標	230
製品資料に関するご使用条件	230

索引	233
---------------------	------------

第 1 章 IBM SPSS Modeler について

IBM® SPSS® Modeler は、ビジネスの専門知識を活用して予測モデルを迅速に作成したり、また作成したモデルをビジネス・オペレーションに展開して意志決定を改善できるようにする、一連のデータ・マイニング・ツールです。IBM SPSS Modeler は業界標準の CRISP-DM モデルをベースに設計されたものであり、データ・マイニング・プロセス全体をサポートして、データに基づいてより良いビジネスの成果を達成できるようにします。

IBM SPSS Modeler には、マシン学習、人工知能、および統計に基づいたさまざまなモデル作成方法が用意されています。「モデル作成」パレットを利用して、データから新しい情報を引き出したり、予測モデルを作成することができます。各手法によって、利点や適した問題の種類が異なります。

SPSS Modeler は、スタンドアロン製品として購入または SPSS Modeler Server と組み合わせてクライアントとして使用することができます。後のセクションで説明されているとおり、多くの追加オプションも使用することができます。詳しくは、「<https://www.ibm.com/analytics/us/en/technology/spss/>」を参照してください。

IBM SPSS Modeler 製品

製品と関連するソフトウェアの IBM SPSS Modeler ファミリーの構成は次のとおりです。

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console (IBM SPSS Deployment Manager に付属)
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- IBM SPSS Modeler Server Adapters for IBM SPSS Collaboration and Deployment Services

IBM SPSS Modeler

SPSS Modeler はこの製品のすべての機能を搭載したバージョンであり、ユーザーのパーソナル・コンピューターにインストールし、そのコンピューターで実行します。スタンドアロン製品としてローカル・モードで SPSS Modeler を実行するか、大規模なデータ・セットを使用する場合にパフォーマンスを向上させるために IBM SPSS Modeler Server と組み合わせて実行することができます。

SPSS Modeler を使用して、プログラミングの必要なく、正確な予測モデルを迅速かつ直感的に構築することができます。独自のビジュアル・インターフェースを使用すると、データ・マイニング・プロセスを簡単に視覚化することができます。製品に組み込まれている高度な分析の支援を受けて、データ内に隠れたパターンやトレンドを発見することができます。結果をモデル化し、ビジネスチャンスを活用してリスクを軽減できるようになり、それらに影響を与える要因を理解することができます。

SPSS Modeler は SPSS Modeler Professional および SPSS Modeler Premium の 2 つのエディションで使用できます。詳しくは、トピック 3 ページの『IBM SPSS Modeler のエディション』を参照してください。

IBM SPSS Modeler Server

SPSS Modeler は、クライアント/サーバー・アーキテクチャーを使用して、リソース集中型の操作が必要な要求を、強力なサーバー・ソフトウェアへ分散します。

SPSS Modeler Server は、1 つまたは複数の IBM SPSS Modeler のインストールと組み合わせてサーバー・ホストで分散分析モードで継続的に実行する、別途ライセンスが必要な製品です。このように、SPSS Modeler Server では、メモリー集中型の操作を、クライアント コンピューターにデータをダウンロードせずにサーバー上で実行できるため、大きなデータ・セットで優れたパフォーマンスを発揮します。IBM SPSS Modeler Server は、パフォーマンスと自動化のさらなる利点を提供し、SQLの最適化とデータベース内のモデリング機能をサポートしています。

IBM SPSS Modeler Administration Console

Modeler Administration Console は、SPSS Modeler Server 構成オプションの多くを管理するグラフィカル・ユーザー・インターフェースです。それらの構成オプションは、オプション・ファイルで設定することも可能です。コンソールは、IBM SPSS Deployment Manager に含まれています。コンソールを使用すると、SPSS Modeler Server インストール済み環境をモニターしたり、構成したりできます。SPSS Modeler Server の現在の顧客は、コンソールを無料で利用できます。アプリケーションは Windows コンピューターにのみインストールできますが、サポートされる任意のプラットフォームにインストールされたサーバーを管理できます。

IBM SPSS Modeler Batch

データマイニングは、通常、対話型のプロセスですが、グラフィカル・ユーザー・インターフェースを必要とせずに、コマンドラインから SPSS Modeler を実行することも可能です。例えば、ユーザーの介入なしで実行する長期実行または反復的なタスクがあります。SPSS Modeler Batch は、通常のユーザー・インターフェースにアクセスせずに SPSS Modeler の完全な分析機能のサポートを提供する製品の特別バージョンです。SPSS Modeler Batch を使用するには、SPSS Modeler Server が必要です。

IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher は、外部ランタイムで実行することができ、外部アプリケーションに埋め込まれる SPSS Modeler ストリームのパッケージ版を作成することができるツールです。このように、SPSS Modeler がインストールされていない環境で使用するための完全な SPSS Modeler ストリームを公開して展開することができます。SPSS Modeler Solution Publisher は、個別のライセンスが必要とされている IBM SPSS Collaboration and Deployment Services - Scoring サービスの一部として配布されています。このライセンスを使用すると、SPSS Modeler Solution Publisher Runtime を受信し、公開されたストリームを実行することができます。

SPSS Modeler Solution Publisher について詳しくは、IBM SPSS Collaboration and Deployment Services の資料を参照してください。IBM SPSS Collaboration and Deployment Services Knowledge Center に『IBM SPSS Modeler Solution Publisher』と『IBM SPSS Analytics Toolkit』というセクションがあります。

IBM SPSS Modeler Server Adapters for IBM SPSS Collaboration and Deployment Services

さまざまな IBM SPSS Collaboration and Deployment Services 用のアダプターを使用すると、SPSS Modeler および SPSS Modeler Server を IBM SPSS Collaboration and Deployment Services リポジトリとインタラクティブに機能させることができます。このように、リポジトリに展開された SPSS

Modeler ストリームは、複数のユーザーで共有したり、シンククライアント アプリケーションである IBM SPSS Modeler Advantage からアクセスしたりできます。リポジトリをホストするシステムに、アダプターをインストールします。

IBM SPSS Modeler のエディション

SPSS Modeler は次のエディションで使用できます。

SPSS Modeler Professional

SPSS Modeler Professional は、CRM システムで追跡する行動や対話、人口統計データ、購入行動や販売データなど、多くの構造化データを処理するために必要なすべてのツールを提供しています。

SPSS Modeler Premium

SPSS Modeler Premium は、特化したデータ、または構造化されていないテキスト・データを処理するために SPSS Modeler Professional を拡張する、別途ライセンスが必要な製品です。SPSS Modeler Premium には、以下の IBM SPSS Modeler Text Analytics が含まれます。

IBM SPSS Modeler Text Analytics は、高度な言語技術と Natural Language Processing (NLP) を使用して、構造化されていない多様なテキスト・データをすばやく処理し、重要なコンセプトを抽出および組織化し、そしてそのコンセプトをカテゴリー別に分類します。抽出されたコンセプトとカテゴリーを、人口統計のような既存の構造化データと組み合わせ、IBM SPSS Modeler の豊富なデータ・マイニング・ツールを適用する方法で、焦点を絞ったより良い決定を下すことができます。

IBM SPSS Modeler ドキュメント

資料は、SPSS Modeler の「ヘルプ」メニューから参照できます。この「ヘルプ」メニューから SPSS Modeler Knowledge Center を開きます。Knowledge Center は、製品の外部で公に利用できます。

各製品の完全な資料 (インストール手順を含む) は、PDF 形式でも提供されており、製品ダウンロードの一部として、個別の圧縮フォルダーに格納されています。PDF 文書は、Web (<http://www.ibm.com/support/docview.wss?uid=swg27046871>) からダウンロードできます。

SPSS Modeler Professional ドキュメント

SPSS Modeler Professional のドキュメント スイート (インストール手順を除く) は次のとおりです。

- **IBM SPSS Modeler ユーザーズ・ガイド**: SPSS Modeler の使用への全体的な入門で、データ ストリームの作成方法、欠損値の処理方法、CLEM 式の作成方法、プロジェクトおよびレポートの処理方法と、IBM SPSS Collaboration and Deployment Services または IBM SPSS Modeler Advantage に展開するためのストリームのパッケージ方法が含まれています。
- 「**IBM SPSS Modeler 入力ノード、プロセス・ノード、出力ノード**」。各種形式のデータの読み取り、処理、および出力に使用するすべてのノードの説明です。これは、モデル作成ノード以外のすべてのノードについての説明です。
- 「**IBM SPSS Modeler モデル作成ノード**」。データ・マイニング・モデルの作成に使用するすべてのノードについての説明です。IBM SPSS Modeler には、マシン学習、人工知能、および統計に基づいたさまざまなモデル作成方法が用意されています。

- **IBM SPSS Modeler アプリケーション・ガイド**。このガイドの例では、特定のモデル作成手法および技法について、簡単に対象を絞って紹介します。本ガイドのオンラインバージョンは、「ヘルプ」メニューからも利用できます。詳しくは、トピック『アプリケーションの例』を参照してください。
- 「**IBM SPSS Modeler Python スクリプトとオートメーション**」。Python スクリプトによるシステムの自動化に関する情報です。ノードおよびストリームの操作に使用できるプロパティーを含めて説明します。
- **IBM SPSS Modeler 展開ガイド: IBM SPSS Deployment Manager** のもとで処理されるジョブ内のステップとして IBM SPSS Modeler ストリームを実行することに関する情報。
- **IBM SPSS Modeler CLEF 開発者ガイド: CLEF** では、IBM SPSS Modeler のノードとしてデータ処理ルーチンやモデル作成アルゴリズムなどのサード・パーティー製のプログラムを統合できます。
- 「**IBM SPSS Modeler データベース内 マイニング・ガイド**」。サード・パーティー製アルゴリズムを使用してご使用のデータベースの能力を利用してパフォーマンスを向上させ、分析機能の範囲を拡張する方法に関する情報を示します。
- **IBM SPSS Modeler Server 管理およびパフォーマンス・ガイド: IBM SPSS Modeler Server** の構成方法と管理方法に関する情報。
- 「**IBM SPSS Deployment Manager ユーザー・ガイド**」。IBM SPSS Modeler Server の監視や構成を行うための Deployment Manager アプリケーションに組み込まれている管理コンソール・ユーザー・インターフェースの使用法に関する情報。
- 「**IBM SPSS Modeler CRISP-DM ガイド**」。SPSS Modeler でのデータ・マイニングに対する CRISP-DM 方法の使用に関するステップバイステップのガイドです。
- 「**IBM SPSS Modeler Batch ユーザーズ・ガイド**」。IBM SPSS Modeler をバッチ・モードで使用するための完全ガイドで、バッチ・モードでの実行およびコマンド・ライン引数の詳細について説明します。このガイドは、PDF 形式のみです。

SPSS Modeler Premium ドキュメント

SPSS Modeler Premium のドキュメントスイート (インストール手順を除く) は次のとおりです。

- 「**SPSS Modeler Text Analytics ユーザーズ・ガイド**」。SPSS Modeler でテキスト分析を使用する場合の情報。テキスト・マイニング・ノード、インタラクティブ・ワークベンチ、テンプレートなどについて説明します。

アプリケーションの例

SPSS Modeler のデータ・マイニング・ツールは、多様なビジネスおよび組織の問題解決を支援しますが、アプリケーションの例では、特定のモデル作成手法および技術に関する簡単で、目的に沿った説明を行います。ここで使用されるデータセットは、データ・マイニング作業によって管理される巨大なデータ・ストアよりも非常に小さいですが、関係するコンセプトや方法は実際のアプリケーションの規模に応じて拡張できます。

例にアクセスするには、SPSS Modeler の「ヘルプ」メニューで「アプリケーションの例」をクリックします。

データ・ファイルとサンプル・ストリームは、製品のインストール・ディレクトリーの Demos フォルダにインストールされています。詳しくは、5 ページの『Demos フォルダ』を参照してください。

データベース・モデル作成の例：例は、『IBM SPSS Modeler データベース内マイニング・ガイド』を参照してください。

スクリプトの例：例は、『IBM SPSS Modeler スクリプトとオートメーション ガイド』を参照してください。

Demos フォルダ

アプリケーションの例で使用されるデータ・ファイルとサンプル・ストリームは、製品のインストール・ディレクトリーの Demos フォルダにインストールされています (例: C:\Program Files\IBM\SPSS\Modeler\<version>\Demos)。このフォルダには、Windowsの「スタート」メニューの IBM SPSS Modeler プログラム・グループから、または「ファイル」 > 「ストリームを開く」ダイアログ・ボックスの最近使ったディレクトリのリストで「Demos」をクリックしてアクセスすることもできます。

ライセンスの追跡

SPSS Modeler を使用すると、ライセンスの使用状況が一定の間隔で追跡され、ログに記録されます。ログに記録されるライセンスメトリックは `AUTHORIZED_USER` と `CONCURRENT_USER` であり、ログに記録されるメトリックのタイプは、SPSS Modeler に使用するライセンスのタイプによって決まります。

作成されたログファイルは IBM License Metric Tool によって処理可能であり、そのファイルからライセンス使用状況レポートを生成できます。

ライセンスログファイルは、SPSS Modeler クライアントログファイルが記録されるディレクトリと同じディレクトリに作成されます (デフォルトでは `%ALLUSERSPROFILE%/IBM/SPSS/Modeler/<version>/log`)。

第 2 章 IBM SPSS Modeler 18.1 の新機能

IBM SPSS Modeler は、本リリースで以下の機能を追加しています。

TWC ソース・ノード。新しい TWC 入力ソース・ノードは、IBM ビジネスの 1 つである The Weather Company から気象データをインポートします。これを使用して、ある場所の過去または予報の気象データを取得できます。これにより、使用可能な最も正確で高精度の気象データを利用して、意思決定を向上させるための気象主導のビジネス・ソリューションの開発に役立てることができます。

Python ノード。「ノード パレット」の新しい「Python」タブには、Python 固有のアルゴリズムを使用するためのノードがあります。SMOTE ノード、XGBoost Tree ノード、XGBoost Linear ノード、および One-Class SVM ノードが使用可能です。これらの新規ノードは、Windows 64、Linux64、および Mac でサポートされています。

Python ディストリビューション。 Python/Spark のスクリプトを SPSS Modeler Server (または SPSS Modeler Client (Windows 64、Linux64、または Mac が必要) に付属のローカル・サーバー) に対して実行する予定の場合、以前のリリースのように Python をインストールして options.cfg を編集する必要がなくなりました。SPSS Modeler に Python ディストリビューションが組み込まれるようになりました。ただし、デフォルトの SPSS Modeler Python ディストリビューションに含まれていない特定のモジュールが必要な場合は、<Modeler_installation_directory>/python にアクセスし、追加のパッケージをインストールできます。

言語ノード。 IBM SPSS Modeler Text Analytics の新しい言語ノードはプロセス ノードの 1 種であり、ソース テキストをスキャンして、書かれている自然言語を判別し、言語のマークを新しいフィールドに書き込みます。主に大量のデータとともに使用するよう設計されています。このノードは、データ ソースに複数の言語が存在するが、1 つの言語のみを処理したい場合に特に役立ちます。

ランダム ツリー ノードの新しい作成オプション。ランダム ツリー ノードでは、「報告する関心の高いルールの数 (Number of interesting rules to report)」という名前の高度な作成オプションが新たに用意されています。以前は、報告されるルールは最大 5 件でした。この新しいオプションでは、報告するルールの数を指定できます (最小値は 1、最大値は 1000、デフォルト値は 50 です)。

IBM Cognos TM1 Server の専用ビューのインポート。TM1 インポート・ノードでは、共有ビューのみではなく、専用ビューもインポートできるようになりました。

R を使用したデータのインポートとエクスポート。拡張用のカスタム・ダイアログ・ビルダーを使用して、カスタム ノードを作成し、R スクリプトを記述して、データ ソースがある任意の場所からデータを読み取り、R によってサポートされている任意のデータ形式にデータを書き出しができるようになりました。

Python for Spark を使用したデータのインポートとエクスポート。拡張用のカスタム・ダイアログ・ビルダーを使用して、カスタム ノードを作成し、Python for Spark スクリプトを記述して、データ ソースがある任意の場所からデータを読み取り、Apache Spark によってサポートされている任意のデータ形式にデータを書き出しができるようになりました。

Analytic Server モード。「ツール」 > 「オプション」 > 「ユーザー オプション」の下で「モード」タブを使用して、従来モードと Analytic Server モードの間を切り替えて、ユーザー・インターフェースに表示するノードおよび式をフィルタリングできます。

行列入替ノードの新しいモード。行列入替ノードの「設定」タブに新規ドロップダウンが追加され、行列入替方法を「フィールドとレコードの両方」、「レコードからフィールドへ」、「フィールドからレコードへ」の中から選択できます。

ストリーム自動保存。ストリームを自動保存する頻度を指定できるようになりました。「ツール」 > 「オプション」 > 「システム オプション」に移動します。

拡張ノード。IBM SPSS Modeler およびそのデータ・マイニング機能を補足するために、新しい拡張ノードでは、エキスパート・ユーザーが独自の **R** スクリプトまたは **Python for Spark** スクリプトを入力して、データ処理、モデリング作成およびモデル・スコアリングを実行できるようにしています。この拡張ノードは、以前有効だった R ノードの後継です。

パレット管理。パレット・マネージャー（「ツール」 > 「パレット管理」）を使用して、以前のリリースのように「お気に入り」タブだけでなく、任意のタブにおけるデフォルトのノード選択を編集できるようになりました。

Microsoft Office サポートの向上。グラフを Microsoft グラフィックス オブジェクトとしてコピーできるようになりました。これにより、グラフを Microsoft Office で操作できます。この形式をサポートするグラフは、Microsoft Office アプリケーションにコピーして、ネイティブの Microsoft Office グラフとしてこれらのアプリケーション内で編集できます。IBM SPSS Modeler のグラフと Microsoft Office のグラフの違いのため、IBM SPSS Modeler のグラフの機能によっては、コピーされたバージョンで維持されないものがあります。Microsoft Office グラフィックス オブジェクト形式のグラフを複数選択してコピーする操作は、サポートされていません。

CPLEX の最適化。新しい CPLEX の最適化ノードにより、OPL (Optimization Programming Language) モデル・ファイルを介した複雑な数学 (CPLEX) ベースの最適化の機能が提供されます。この機能は IBM Analytical Decision Management 製品でも使用可能ですが、IBM Analytical Decision Management の必要なしに、SPSS Modeler でも CPLEX ノードを使用できるようになりました。

オンライン・ヘルプ。以前のリリースでは、オンライン・ヘルプ・システムは IBM SPSS Modeler と一緒にインストールされていました。現在、オンライン・ヘルプ・システムは、IBM Knowledge Center に含まれます。IBM SPSS Modeler アプリケーション内部で F1 によってヘルプを呼び出すと、IBM Knowledge Center の最新バージョンの資料にリンクする別の Web ブラウザー・インスタンスで関連ヘルプ・トピックが開くようになりました。

RFE コミュニティー。IBM Knowledge Center が Request for Enhancement (RFE) コミュニティー 内に移動しました。ここでは、IBM 製品のユーザーが、IBM KC アプリケーションおよびそのコンテンツを改善するための提案を行えます。新規要求のオープンおよび追跡に加えて、ユーザーは既存の要求を検索したり、既存の要求にコメントしたりできます。RFE コミュニティーは、ユーザーが IBM KC および製品の開発チームとより直接的に対話するための機会を提供します。RFE コミュニティーの概要については、「ヘルプ」タブを参照してください。「ヘルプ」タブ以外の機能を使用するには、IBMid およびパスワードが必要です。

第 3 章 IBM SPSS Modeler の概要

はじめに

IBM SPSS Modeler はデータ・マイニング・アプリケーションで、大きなデータ・セット中の有益なリレーションシップを見つけだすための、戦略的なアプローチ手段を提供しています。従来の統計的な手法とは対照的に、作業開始時に何を見つけだそうとしているのかがわからなくてもかまいません。データの探索、さまざまなモデルの適合、およびさまざまなデータの関連の調査などの作業を行いながら、有益な情報を見つけだしていくことができます。

IBM SPSS Modeler の開始

アプリケーションを起動するには、以下のメニューをクリックします。

「スタート」 > 「(すべての) プログラム」 > 「IBM SPSS Modeler <バージョン>」 > 「IBM SPSS Modeler <バージョン>」

数秒後にメイン・ウィンドウが表示されます。

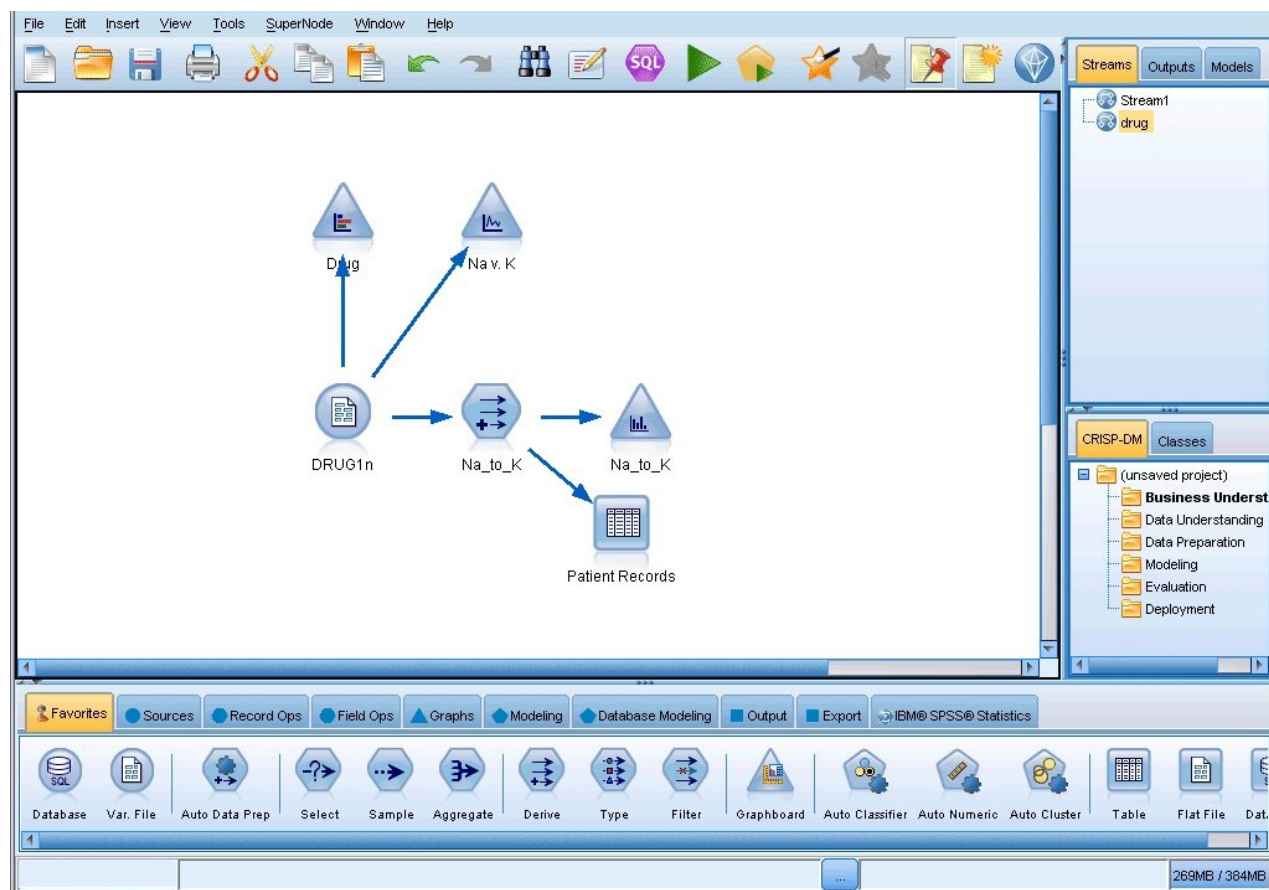


図 1. IBM SPSS Modeler のメイン・アプリケーション・ウィンドウ

コマンド・ラインからの起動

オペレーティング・システムのコマンド・ラインを使用し、次のようにして IBM SPSS Modeler を起動できます。

1. IBM SPSS Modeler がインストールされているコンピューターで、DOS つまりコマンド・プロンプト・ウィンドウを開きます。
2. IBM SPSS Modeler インターフェースをインタラクティブ・モードで起動するには、`modelerclient` コマンドを入力し、続いて例えば次のような適切な引数を入力します。

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

使用可能な引数 (フラグ) により、サーバーへの接続、ストリームのロード、スクリプトの実行、または必要に応じて他のパラメーターの指定を行うことができます。

IBM SPSS Modeler Server に接続中

IBM SPSS Modeler は、スタンドアロンのアプリケーションとして、または、IBM SPSS Modeler Server に直接または IBM SPSS Collaboration and Deployment Services から Coordinator of Processes プラグインを介して IBM SPSS Modeler Server またはサーバー・クラスターに接続されたクライアントとして、実行できます。現在の接続ステータスは、IBM SPSS Modeler ウィンドウの左下に表示されます。

サーバーに接続する場合は、接続するサーバー名を手動で入力するか、以前定義した名前を選択できます。ただし、IBM SPSS Collaboration and Deployment Services を使用する場合、「サーバーへのログイン」ダイアログ・ボックスからサーバーまたはサーバー・クラスターのリストを使用して検索することができます。ネットワーク上で実行する Statistics サービスを介して参照する機能は、Coordinator of Processes で使用できます。

サーバーに接続するには

1. 「ツール」メニューの「サーバーへのログイン」をクリックします。「サーバーへのログイン」ダイアログ・ボックスが開きます。または、IBM SPSS Modeler ウィンドウの接続ステータス領域をダブルクリックします。
2. ダイアログ・ボックスで、ローカル・サーバーのコンピューターに接続するオプションを指定するか、テーブルから接続を選択します。
 - 「追加」または「編集」をクリックして、接続を追加または編集します。詳しくは、トピック 11 ページの『IBM SPSS Modeler Server 接続の追加および編集』を参照してください。
 - 「検索」をクリックして、SPSS COP のサーバーまたはサーバー・クラスターにアクセスします。詳しくは、トピック 12 ページの『IBM SPSS Collaboration and Deployment Services のサーバーの検索』を参照してください。

「サーバー・テーブル」。このテーブルには、一連の定義済みのサーバー接続が含まれています。このテーブルには、デフォルト接続、サーバー名、説明、ポート番号が表示されています。既存の接続を選択または検索、あるいは新しい接続を手動で追加することができます。特定のサーバーをデフォルト接続として設定するには、接続のテーブルの「デフォルト」列のチェック・ボックスを選択します。

「デフォルト・データ・パス」。サーバー・コンピューター上のデータに使用されるパスを指定します。「...」ボタンをクリックして、目的の場所を指定することもできます。

「資格情報の設定」。このボックスのチェック・マークを外した状態にして、シングル・サインオン機能を有効にします。これにより、ローカル・コンピューターのユーザー名およびパスワードの詳細を

使用してサーバーへのログインを試みます。この でシングル・サインオンを使用できない場合、またはこのボックスをチェックしてシングル・サインオンを無効にした場合 (例えば、管理者アカウントにログインした場合)、資格情報を入力するための次のフィールドが表示されます。

「ユーザー ID」。サーバーにログインするユーザー名を入力します。

「パスワード」。指定したユーザー名に関連付けられたパスワードを入力します。

「ドメイン」。サーバーにログオンするために使用するドメインを指定します。サーバー・コンピューターが クライアント・コンピューターとは異なる Windows ドメインにある場合にのみ、ドメイン名が必要です。

3. 「OK」をクリックして、接続を完了します。

サーバーとの接続を切断するには

1. 「ツール」メニューの「サーバーへのログイン」 をクリックします。「サーバーへのログイン」ダイアログ・ボックスが開きます。または、IBM SPSS Modeler ウィンドウの接続ステータス領域をダブルクリックします。
2. ダイアログ・ボックスで、「ローカル・サーバー」を選択し、「OK」 をクリックします。

IBM SPSS Modeler Server 接続の追加および編集

「サーバーへのログイン」ダイアログ・ボックスでサーバー接続を手動で編集または追加することができます。「追加」をクリックすると、サーバー接続の詳細を入力できる空の「サーバーの追加/編集」ダイアログ・ボックスにアクセスすることができます。「サーバーへのログイン」ダイアログ・ボックスで既存の接続を選択して「編集」をクリックすると、「サーバーの追加/編集」ダイアログ・ボックスが開いて接続の詳細が表示され、その接続を変更することができます。

注: IBM SPSS Collaboration and Deployment Services から追加されたサーバー接続は、名前、ポート、およびそのほかの詳細が IBM SPSS Collaboration and Deployment Services で定義されているため、編集することができません。IBM SPSS Collaboration and Deployment Services と SPSS Modeler Client の両方との通信で同じポートを使用することが、ベスト・プラクティスとして規定されています。これらは options.cfg ファイル内で max_server_port および min_server_port として設定できます。

サーバー接続を追加するには

1. 「ツール」メニューの「サーバーへのログイン」 をクリックします。「サーバーへのログイン」ダイアログ・ボックスが開きます。
 2. ダイアログ・ボックスで、「追加」 をクリックします。「サーバーへのログイン: サーバーの追加/編集」ダイアログ・ボックスが表示されます。
 3. サーバー接続の詳細を入力して 「OK」 をクリックします。接続が保存され、「サーバーへのログイン」ダイアログ・ボックスに戻ります。
- サーバー。利用できるサーバーを指定するか、またはリストから選択します。サーバー・コンピューターは、英数字の名前 (例えば、*myserver*)、または、サーバー・コンピューターに割り当てられた IP アドレス (例えば、202.123.456.78) で識別できます。
 - ポート。サーバーが待機しているポート番号を入力します。デフォルトのポート番号がうまく動作しない場合は、システム管理者に問い合わせ正しいポート番号を取得してください。
 - 説明。サーバー接続の説明をオプションで入力します。
 - 「セキュア接続を確保 (SSL の使用)」。SSL (Secure Sockets Layer) 接続を使用するかどうかを指定します。SSL は、ネットワークを介してセキュアなデータ送信を行うために一般的に使用されているプ

ロトコルです。この機能を使用するには、IBM SPSS Modeler Server をホストするサーバー側で SSL を有効にする必要があります。必要な場合、詳細を各サイトの管理者にお問い合わせください。

サーバー接続を編集するには

1. 「ツール」メニューの「サーバーへのログイン」をクリックします。「サーバーへのログイン」ダイアログ・ボックスが開きます。
2. ダイアログ・ボックスで、編集する接続を選択し、「編集」をクリックします。「サーバーへのログイン: サーバーの追加/編集」ダイアログ・ボックスが表示されます。
3. サーバー接続の詳細を変更して「OK」をクリックします。変更が保存され、「サーバーへのログイン」ダイアログ・ボックスに戻ります。

IBM SPSS Collaboration and Deployment Services のサーバーの検索

サーバー接続を手動で入力する代わりに、IBM SPSS Collaboration and Deployment Services で使用できる Coordinator of Processes (COP) を介してネットワークで使用可能なサーバーまたはサーバー・クラスターを選択できます。サーバー・クラスターは、Coordinator of Processes が処理要求に応答するのに最適なサーバーを決定するサーバーのグループです。

「サーバーへのログイン」ダイアログ・ボックスで手動でサーバーを追加することもできますが、使用可能なサーバーの検索では、正しいサーバー名およびポート番号を把握していなくてもサーバーに接続できます。この情報は自動的に提供されます。この情報は、自動的に提供されます。ただし、ユーザー名、ドメインおよびパスワードなどの、正しいログオン情報が必要です。

注：Coordinator of Processes 機能へアクセスしていない場合、接続するサーバー名を手動で入力したり、以前定義した名前を選択することができます。詳しくは、トピック 11 ページの『IBM SPSS Modeler Server 接続の追加および編集』を参照してください。

サーバーおよびクラスターを検索するには

1. 「ツール」メニューの「サーバーへのログイン」をクリックします。「サーバーへのログイン」ダイアログ・ボックスが開きます。
2. ダイアログ・ボックスで「検索」をクリックすると、「サーバーの検索」ダイアログ・ボックスが表示されます。Coordinator of Processes をブラウズしようとする際に IBM SPSS Collaboration and Deployment Services にログオンしていない場合、ログオンを指示するメッセージが表示されます。
3. リストからサーバーまたはサーバー・クラスターを選択します。
4. 「OK」をクリックしてダイアログ・ボックスが閉じられ、選択した接続が「サーバーへのログイン」ダイアログ・ボックスのテーブルに追加されます。

Analytic Server に接続中

複数の Analytic Server が使用できる場合は、「Analytic Server 接続」ダイアログを使用して、IBM SPSS Modeler 内での使用目的で複数のサーバーを定義できます。管理者が、デフォルトの Analytic Server を <Modeler_install_path>/config/options.cfg ファイルにセットアップしている場合があります。しかし、その他のサーバーを定義した後に、それらの有効なサーバーを使用することもできます。例えば、Analytic Server 入力ノードおよびエクスポート・ノードを使用している場合、あるストリームのそれぞれのブランチ内に異なる Analytic Server 接続を使用することで、各ブランチの実行時にブランチがそれ自体の Analytic Server を使用して IBM SPSS Modeler Server にデータが引き出されないようにすることができます。1 つのブランチに複数の Analytic Server 接続が含まれている場合は、Analytic Server から IBM SPSS Modeler Server にデータが引き出されることに注意してください。制限などの詳細については、48 ページの『Analytic Server のストリームのプロパティ』を参照してください。

新しい Analytic Server 接続を作成するには、「ツール」 > 「Analytic Server 接続」に移動し、ダイアログの以下の各セクションに必要な情報を指定します。

接続

「URL」。Analytic Server の URL を `https://hostname:port/contextroot` の形式で入力します。ここで、hostname は Analytic Server の IP アドレスまたはホスト名、port はそのポート番号、contextroot は Analytic Server のコンテキスト・ルートです。

「テナント」。IBM SPSS Modeler Server がメンバーであるテナントの名前を入力します。テナントが不明な場合は、管理者に連絡してください。

認証

モード: 以下の認証モードから選択します。

- 「ユーザー名およびパスワード」の場合、ユーザー名およびパスワードの入力が必要です。
- 「保管されている資格情報」の場合、IBM SPSS Collaboration and Deployment Services Repository から資格情報を選択する必要があります。
- 「Kerberos」の場合、サービス プリンシパル名および構成ファイル パスの入力が必要です。この情報が不明な場合は、管理者に連絡してください。

「ユーザー名」。Analytic Server ユーザー名を入力します。

「パスワード」。Analytic Server のパスワードを入力します。

「接続」。「接続」をクリックして、新しい接続をテストします。

接続

上記の情報を指定して「接続」をクリックすると、この「接続」テーブルに接続が追加されます。接続を削除する必要がある場合は、その接続を選択し、「削除」をクリックします。

管理者がデフォルトの Analytic Server 接続を `options.cfg` ファイルに定義した場合は、「デフォルト接続の追加 (Add default connection)」をクリックして、使用可能な接続にその接続を追加することもできます。ユーザー名およびパスワードの入力を要求するプロンプトが出されます。

一時ディレクトリの変更

IBM SPSS Modeler Server が行う処理や操作の中には、一時ファイルを作成する必要があるものもあります。IBM SPSS Modeler のデフォルトでは、システムの一時ファイル用ディレクトリーに一時ファイルが作成されます。一時ディレクトリーの場所を変更するには、次の手順に従ってください。

1. 新規ディレクトリー `spss` およびそのサブディレクトリー `servertemp` を作成します。
2. IBM SPSS Modeler のインストール・ディレクトリー中の `/config` にある、`options.cfg` を編集します。次のファイルの `temp_directory` パラメーターを編集して読み込みます。`temp_directory, "C:/spss/servertemp"`
3. この作業を行った後は、IBM SPSS Modeler Server サーバー・サービスを再起動する必要があります。再起動するには、Windows コントロール・パネルの「管理ツール」にある「サービス」タブをクリックしてください。サービスを停止した後、再び開始すると変更内容が有効になります。また、マシンを再起動しても、サービスが再開されます。

これで、新しいディレクトリーに一時ファイルが作成されるようになります。

注:

- スラッシュを使用する必要があります。
- IBM SPSS Collaboration and Deployment Services ジョブ経由で評価ストリームを実行する場合、temp_directory 設定は適用されません。そのようなジョブを実行するときは、一時ファイルが作成されます。デフォルトでは、このファイルは IBM SPSS Modeler Server のインストール・ディレクトリーに保存されます。IBM SPSS Modeler で IBM SPSS Modeler Server 接続を作成するときは、一時ファイルを保存するデフォルトのデータ・フォルダーを変更できます。

複数の IBM SPSS Modeler セッションの起動

複数の IBM SPSS Modeler を一度に起動する必要がある場合、IBM SPSS Modeler および Windows の設定を変更する必要があります。例えば、2 つの個別のサーバー・ライセンスを持ち、同じクライアント・マシンから 2 つの異なるサーバーに対して 2 つのストリームを実行する場合に、変更を行う必要があります。

複数の IBM SPSS Modeler セッションを有効化する手順は、次のとおりです。

1. 以下のメニューをクリックします。

「スタート」 > 「(すべての) プログラム」 > 「IBM SPSS Modeler」

2. IBM SPSS Modeler 18.1 のショートカット (アイコンで表示) で、右クリックして「プロパティー」を選択します。
3. 「対象」 テキスト・ボックスで、文字列の終わりに `-noshare` を追加します。
4. Windows の Explorer で、次の項目を選択します。

「ツール」 > 「フォルダー・オプション...」

5. 「ファイル タイプ」 タブで、「IBM SPSS Modeler ストリーム」 オプションを選択し、「詳細」 をクリックします。
6. 「ファイル タイプの編集」 ダイアログ・ボックスで、「IBM SPSS Modeler で開く」 を選択し、「編集」 をクリックします。
7. 「アクションの実行に使用するアプリケーション」 テキスト・ボックスで、`-stream` 引数の前に `-noshare` を追加します。

IBM SPSS Modeler インターフェースについて

IBM SPSS Modeler の使いやすいインターフェースでは、データ・マイニング処理の各ポイントで、特定ビジネスの専門知識が必要となります。予測、分類、セグメント化、関連性検出などのモデリング・アルゴリズムによって、強力かつ正確なモデルが作成されます。モデルの結果は簡単に展開でき、データベース、IBM SPSS Statistics、およびさまざまなアプリケーションで使用することができます。

IBM SPSS Modeler での作業は、3 ステップのデータの処理で行われます。

- まず、データを IBM SPSS Modeler に読み込みます。
- そして一連の操作によってデータを実行します。
- 最後に、データを宛て先に送信します。

各操作により、入力から各レコード単位でデータが流れて処理されて、最後に終点 (モデルまたはデータ出力タイプ) に到達するため、この操作の流れはデータ・ストリームと呼ばれます。



図 2. 単純なストリーム

IBM SPSS Modeler ストリーム領域

ストリーム領域は、IBM SPSS Modeler ウィンドウ内で最も広い領域を占めています。ここで、データ・ストリームの作成と操作を行います。

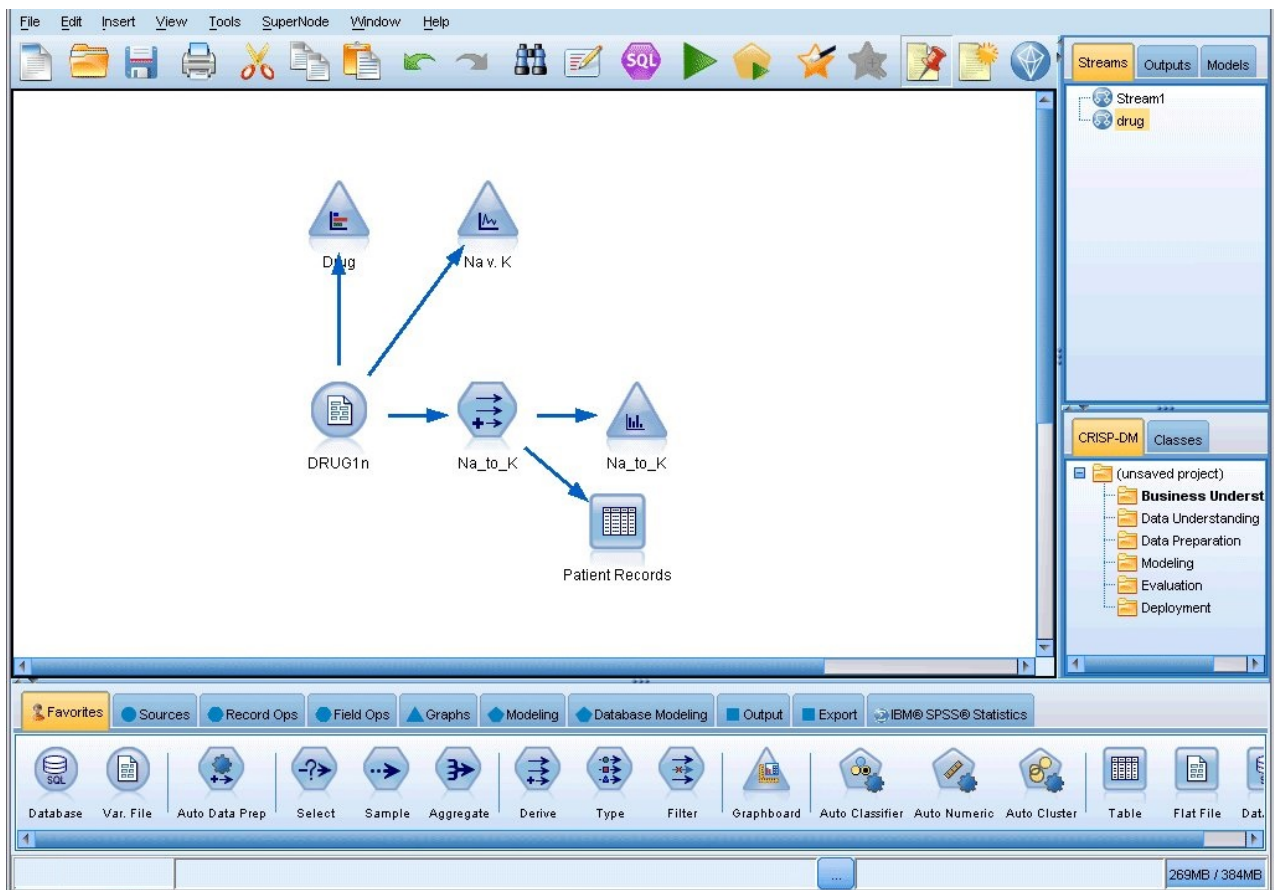


図 3. IBM SPSS Modeler 作業領域 (デフォルトのビュー)

ストリームは、業務に関連するデータ操作のダイアグラムをインターフェースのメイン領域に描画することによって作成します。各操作はノードと呼ばれるアイコンで表されます。ノードは、データの流れと操作を表すストリーム中に配置され、相互に接続されます。

IBM SPSS Modeler では、同じストリーム領域内で、または別のストリーム領域を開いて、複数のストリームに関する作業を同時に行えます。セッションの間、ストリームは IBM SPSS Modeler ウィンドウの右上にあるストリーム・マネージャに格納されます。

注: 組み込みトラックパッドの「強めのクリックと触覚フィードバック」設定を有効に指定して MacBook を使用している場合、ノード・パレットからストリーム領域へのノードのドラッグ・アンド・ドロップを行うと、その領域に重複ノードが追加されることがあります。この問題を回避するために、「強めのクリックと触覚フィードバック」トラックパッド・システム設定を無効にすることをお勧めします。

ノード・パレット

SPSS Modeler のほとんどのデータおよびモデル作成ツールは、ストリーム領域下のウィンドウの下部にあるノード・パレット から使用できます。

例えば、「レコード操作」パレット・タブには、選択、結合、追加など、データ・レコードの操作を実行するために使用できるノードが含まれています。

ストリーム領域にノードを追加するには、ノード・パレットのアイコンをダブルクリックするか、アイコンをストリーム領域にドラッグします。次にそれらのノードを接続し、データの流れを表すストリームを作成していきます。

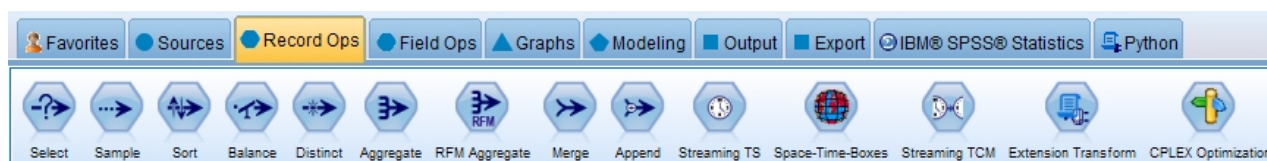


図 4. ノード・パレットの「レコード設定」タブ

それぞれのパレット・タブには、次のようなストリーム操作の各段階で使われる関連ノード群が用意されています。

- 「入力」ノードでは、SPSS Modeler にデータを入力します。
- 「レコード操作」ノードでは、選択、結合、追加など、データ・レコード の操作を実行します。
- 「フィールド操作」ノードでは、フィルタリング、新規フィールドの派生、特定フィールドの測定 of 尺度の特定など、データ・フィールド の操作を実行します。
- 「グラフ作成」ノードでは、モデルの作成前および作成後にデータをグラフィカルに表示します。グラフには、散布図、ヒストグラム、Web グラフ・ノード、および評価グラフなどがあります。
- 「モデル作成」ノードでは、ニューラル・ネットワーク、ディジション・ツリー、クラスタリング・アルゴリズム、データのシーケンス化など、SPSS Modeler で利用できるモデル作成アルゴリズムを使用します。
- 「DB モデリング」ノードでは、Microsoft SQL Server、IBM DB2、Oracle データベース、および Netezza データベースで利用できるモデル作成アルゴリズムを使用します。
- 「出力」ノードでは、SPSS Modeler で表示できるデータ、グラフ、モデルの結果のさまざまな出力を作成します。
- 「エクスポート」ノードでは、IBM SPSS Data Collection や Excel などの外部アプリケーションで表示できるさまざまな出力を作成します。
- 「IBM SPSS Statistics」ノードでは、IBM SPSS Statistics プロシーチャーを実行するほか、IBM SPSS Statistics との間でデータのインポートまたはエクスポートを行います。
- 「Python」ノードは、Python アルゴリズムの実行に使用できます。

SPSS Modeler に慣れてきたら、パレットの内容を使いやすくカスタマイズすることができます。

「ノード パレット」の左側で、「監視」、「アソシエーション」、または「セグメンテーション」を選択して、表示するノードをフィルタリングできます。

ノード・パレットの下部にあるレポート・パネルには、データ・ストリームにデータを読み込む場合など、各種操作の進行状況に関するフィードバック情報が表示されます。ステータス・パネルもノード・パレットの下にあります。このウィンドウには、アプリケーションの現在の処理状況や、ユーザーへのフィードバックが必要な場合の指示などが表示されます。

注: 組み込みトラックパッドの「強めのクリックと触覚フィードバック」設定を有効に指定して MacBook を使用している場合、ノード・パレットからストリーム領域へのノードのドラッグ・アンド・ドロップを行うと、その領域に重複ノードが追加されることがあります。この問題を回避するために、「強めのクリックと触覚フィードバック」トラックパッド・システム設定を無効にすることをお勧めします。

IBM SPSS Modeler マネージャ

ウィンドウの右上がマネージャー・パネルです。3 つのタブがあり、ストリーム、出力、モデルの管理を行います。

「ストリーム」タブでは、セッション中に作成されたストリームを開いたり、名前を変更したり、保存したり、削除することができます。



図 5. 「ストリーム」タブ



図 6. 「出力」タブ

「出力」タブには、IBM SPSS Modeler のストリーム操作で作成されたグラフおよびテーブルなどのさまざまなファイルが表示されます。ここに記載されているテーブル、グラフ、およびレポートを表示したり、

名前を変更したり、または閉じることができます。



図 7. モデル ナゲットを含む [モデル] タブ

「モデル」タブは、最も強力なマネージャ タブです。このタブにはすべてのモデル ナゲットが表示されます。それらは現在のセッションの IBM SPSS Modeler で生成されるモデルです。これらのモデルは、「モデル」タブから直接参照することも、領域内のストリームに追加することもできます。

IBM SPSS Modeler プロジェクト

ウィンドウの右下には、データ・マイニング・プロジェクト (データ・マイニング・タスクに関連するファイルのグループ) の作成と管理に使用するプロジェクト パネルがあります。IBM SPSS Modeler で作成したプロジェクトを表示するには、クラス・ビューと CRISP-DM ビューの 2 つの方法が使用できます。

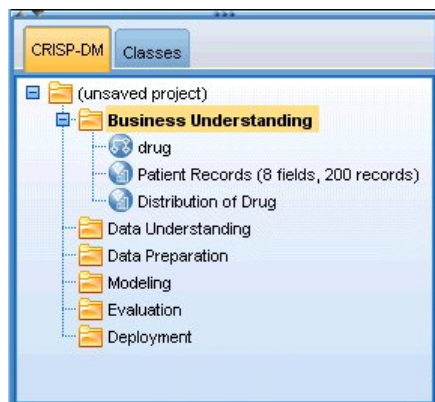


図 8. CRISP-DM ビュー

「CRISP-DM」タブでは、世界的に知られている方法論で実績のある CRISP-DM (Cross-Industry Standard Process for Data Mining) に基づいてプロジェクトを編成することができます。データ・マイニングを熟知している方でも、初めてデータ・マイニングを行う方でも、CRISP-DM ツールを使用することにより、より円滑にプロジェクトを編成し、最良の結果を得ることができます。

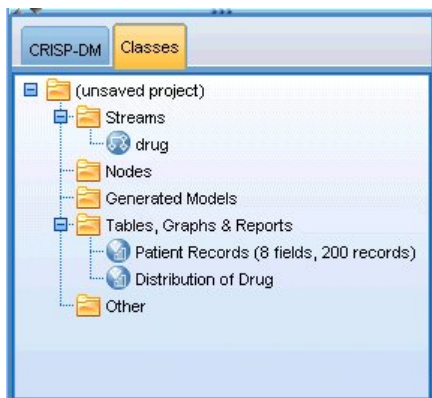


図 9. クラス・ビュー

「クラス」タブでは、IBM SPSS Modeler で行った作業内容を、作成したオブジェクトの種類に応じて編集することができます。このビューは、データ、ストリーム、およびモデルの現状を把握するときに便利です。

IBM SPSS Modeler ツールバー

IBM SPSS Modeler ウィンドウの上部には、アイコンが配置されたツールバーがあります。このツールバーは、役に立つさまざまな機能を提供しています。ツールバー・ボタンとその機能を次に示します。



新規ストリームの作成



ストリームを開く



ストリームの保存



現在のストリームの印刷



切り取ってクリップボードに移動



クリップボードにコピー



貼り付けの選択



最後の操作を元に戻す



やり直し



ノードの検索



ストリームのプロパティを編集



SQL 生成をプレビュー

	ストリームを実行		選択したストリームを実行
	ストリームの中止 (ストリームの実行中にだけ利用可能)		スーパーノードの追加
	ズーム・イン (スーパーノード専用)		ズーム・アウト (スーパーノード専用)
	ストリームにマークアップがありません		コメントの挿入
	ストリーム・マークアップの非表示 (ある場合)		非表示のストリーム・マークアップを表示
	IBM SPSS Modeler Advantage でストリームを開く		

ストリーム・マークアップはストリームのコメント、モデル・リンク、およびスコアリングブランチの表示で構成されています。

モデルのリンクについては、『IBM SPSS モデル作成ノード ガイド』を参照してください。

ツールバーのカスタマイズ

ツールバーは、次のようにさまざまな観点から変更できます。

- 表示するかどうか
- アイコンのツールヒントを使用できるかどうか
- 大きいアイコンまたは小さいアイコンのどちらを使用するか

ツールバー表示をオンまたはオフにするには

1. メイン・メニューで次の各項目をクリックします。

「表示」 > 「ツールバー」 > 「表示」

ツールヒントまたはアイコンのサイズ設定を変更するには

1. メイン・メニューで次の各項目をクリックします。

「表示」 > 「ツールバー」 > 「カスタマイズ」

必要に応じて、「ツールヒントを表示」 または 「大きいボタン」 をクリックします。

IBM SPSS Modeler ウィンドウのカスタマイズ

SPSS Modeler インターフェースのさまざまな構成要素の間にあるディバイダを使用して、そのサイズを変更したり、ツールを閉じることができます。例えば、大きいストリームを作成して作業を行う場合などに、各ディバイダにある小さい矢印をクリックしてノード・パレット、マネージャー・パネル、およびプロジェクト・パネルを閉じることができます。これらのウィンドウを閉じることによってストリーム領域が広がるため、大きいストリームや複数のストリームで作業を行うために十分な領域を確保することができます。

または、「表示」メニューで、「ノード パレット」、「マネージャー」 または 「プロジェクト」 をクリックして、これらの項目の表示をオンまたはオフにします。

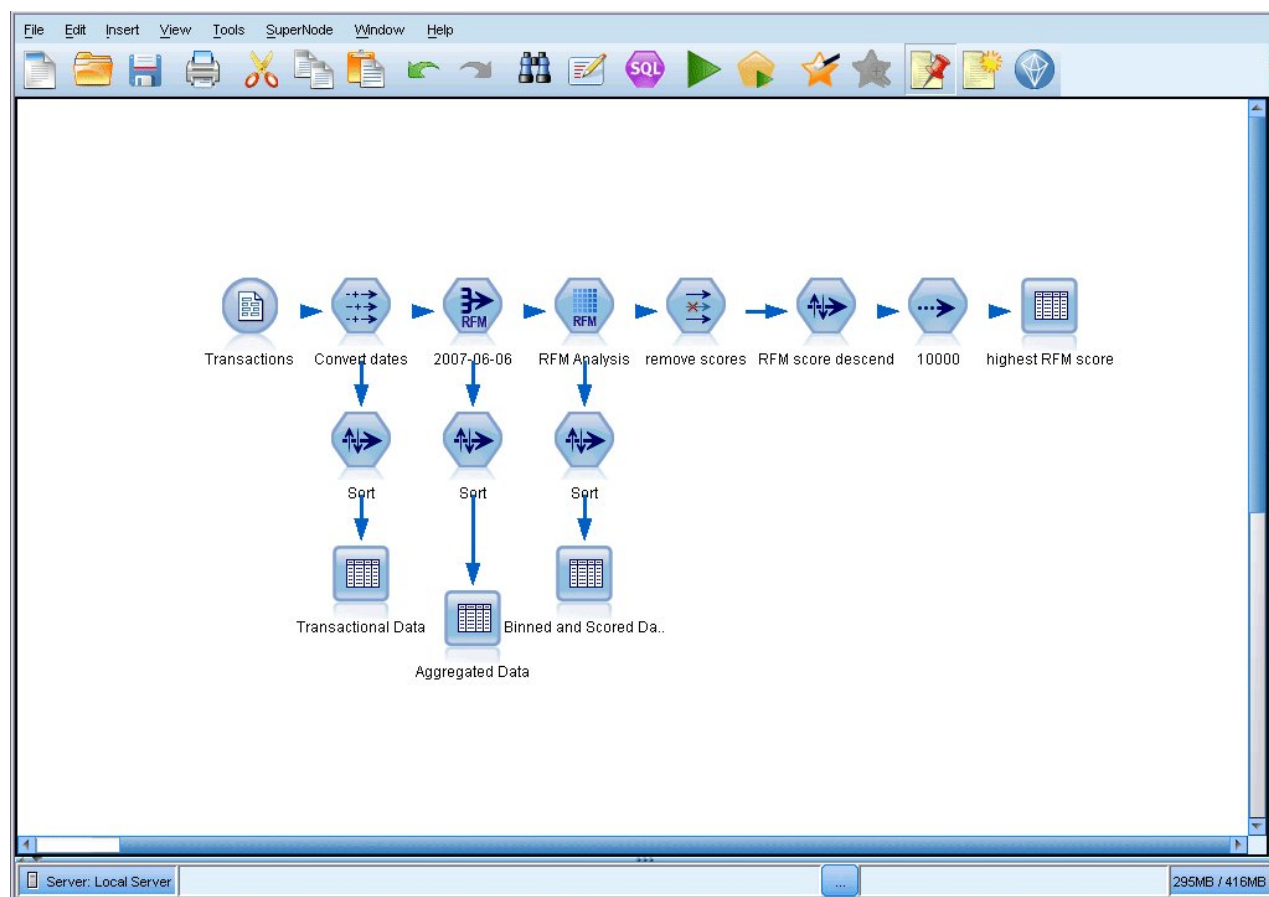


図 10. 最大化されたストリーム キャンバス

ノード・パレット、マネージャー・パネル、およびプロジェクト・パネルを閉じる代わりに、SPSS Modeler ウィンドウの横と下にある青いスクロール・バーを動かして、ストリーム領域をスクロールすることもできます。

画面マークアップの表示を制御できます。画面マークアップはストリームのコメント、モデル・リンク、およびスコアリングブランチの表示で構成されています。この表示をオンまたはオフにするには、次をクリックします。

「表示」 > 「ストリーム マークアップ」

ストリームのアイコン・サイズの変更

以下の方法で、ストリームのアイコンのサイズを変更できます。

- ストリーム・プロパティの設定から
- ストリームのポップアップ・メニューから
- キーボードの使用

ストリーム ビュー全体を標準アイコンサイズの 8% ~ 200% のいずれかのサイズに設定します。

ストリーム全体のサイズ変更 (ストリーム・プロパティから)

1. メイン・メニューから次の各項目を選択します。

「ツール」 > 「ストリームのプロパティ」 > 「オプション」 > 「レイアウト」

2. 「アイコンのサイズ」メニューからサイズを選択します。
3. 結果を表示するには「適用」をクリックします。
4. 「OK」をクリックして、変更を保存します。

ストリーム全体のサイズ変更 (メニューから)

1. 領域のストリームの背景を右クリックします。
2. 「アイコン サイズ」を選択し、該当するサイズを選択します。

ストリーム全体のサイズ変更 (キーボードを使用)

1. Ctrl + [-] キーを押すと、ズームアウトします。
2. Ctrl + [+] キーを押すと、ズーム・インします。

この機能は、複雑なストリームの全体のビューを表示するのに適しています。ストリームの印刷に必要なページ数を最小限にすることもできます。

IBM SPSS Modeler でのマウスの使用方法

IBM SPSS Modeler でよく使われるマウス操作を次に示します。

- シングルクリック。マウスの右または左ボタンを使用して、メニューからオプションを選択したり、ポップアップ・メニューを開いたり、さまざまな標準のコントロールやオプションにアクセスすることができます。ノードをクリックしたままマウスを動かしてドラッグすれば、ノードを移動できます。
- ダブルクリック。マウスの左ボタンを使用してダブルクリックすると、ノードをストリーム領域に置いて、既存のノードを編集できるようになります。
- 中央ボタンのクリック。マウスの中央ボタンをクリックして、カーソルをドラッグすることにより、ストリーム領域のノードを接続します。マウスの中央ボタンをダブルクリックすると、ノードの接続が解除されます。中央ボタンがないマウスを使用している場合は、代わりに Alt キーを押しながらクリックしたり、ドラッグしてください。

ショートカット・キーの使用

大部分の IBM SPSS Modeler ビジュアル・プログラミング操作には、対応するショートカット・キーが用意されています。例えば、ノードをクリックしてキーボードの Del キーを押すと、ノードを削除することができます。同様に、Ctrl キーを押したまま S キーを押すと、ストリームを素早く保存できます。このようなコントロール・コマンドは、Ctrl+S のように、Ctrl と他のキー名で示されています。

Ctrl+X (切り取り) のように、標準の Windows 操作で使われているショートカット・キーも数多くあります。IBM SPSS Modeler では、後述するアプリケーション独自のショートカット・キーのほかに、これらの標準のショートカット・キーを利用することもできます。

注: 一部、IBM SPSS Modeler で使われていた古いショートカット・キーが Windows 標準のショートカット・キーと重複している場合があります。これらの古いショートカット・キーを使用する場合は、Alt キーも一緒に押しながら使用してください。例えば、Ctrl+Alt+C キーを使用して、キャッシュをオンとオフに切り換えることができます。

表 1. サポートしているショートカット・キー

ショートカット・キー	関数
Ctrl+A	すべてを選択する
Ctrl+X	切り取り
Ctrl+N	新規ストリーム
Ctrl+O	ストリームを開く
Ctrl+P	印刷
Ctrl+C	コピー
Ctrl+V	貼り付け
Ctrl+Z	元に戻す
Ctrl+Q	選択したノードの下流にあるすべてのノードを選択
Ctrl+W	下流のすべてのノードの選択を解除 (Ctrl+Q と切り替わる)
Ctrl+E	選択したノードから実行
Ctrl+S	現在のストリームを保存
Alt+矢印キー	選択したストリーム領域上のノードを矢印の方向に移動
Shift+F10	選択したノードのポップアップ・メニューを表示

表 2. 古いホットキーに対応するショートカット・キー

ショートカット・キー	関数
Ctrl+Alt+D	ノードの複製
Ctrl+Alt+L	ノードのロード
Ctrl+Alt+R	ノード名の変更
Ctrl+Alt+U	ユーザー入力ノードの生成
Ctrl+Alt+C	キャッシュのオン/オフの切り替え
Ctrl+Alt+F	キャッシュの取消
Ctrl+Alt+X	スーパーノードの展開
Ctrl+Alt+Z	ズーム・イン/ズーム・アウト
Delete	ノードまたは接続の削除

印刷中

IBM SPSS Modeler では、次のオブジェクトを印刷できます。

- ストリームのダイアグラム
- グラフ作成
- テーブル
- レポート (レポート・ノードおよびプロジェクト・レポートから)
- スクリプト (「ストリームのプロパティ」、「スタンドアロン・スクリプト」、または「スーパーノード スクリプト」ダイアログ・ボックスから)
- モデル (モデル・ブラウザー、現在フォーカスのあるダイアログ・ボックスのタブ、ツリー・ビューア)
- 注釈 (出力の「注釈」タブを使用して)

オブジェクトを印刷するには、次のようにします。

- プレビューを行わないでオブジェクトを印刷するには、ツールバーの「印刷」ボタンをクリックします。
- 印刷前に印刷の設定を行うには、「ファイル」メニューの「ページ設定」を選択します。
- 印刷前にプレビューを行う場合は、「ファイル」メニューの「印刷プレビュー」を選択します。
- 標準の印刷ダイアログ・ボックスに、選択されているプリンタのオプションを表示して、さまざまなオプションを設定するには、「ファイル」メニューの「印刷」を選択します。

IBM SPSS Modeler の自動化

高度なデータ・マイニング作業は複雑で長期間になることもあるため、IBM SPSS Modeler にはさまざまな種類のコーディングや自動化のサポート機能が用意されています。

- **Control Language for Expression Manipulation (CLEM)** は、IBM SPSS Modeler ストリーム中を流れるデータの分析と操作を行うための強力な言語です。CLEM を使用すれば、経費と収入データから利益を算出するような簡単な操作から、Web ログ・データを有益な情報を含む一連のフィールドやレコードに変換するような複雑な操作まで、さまざまなストリーム操作を行うことができます。
- スクリプトは、ユーザー・インターフェースのプロセスを自動化する強力なツールです。スクリプトは、マウスやキーボードを使用して実行するのと同じような操作を実行できます。また、出力を指定して生成されたモデルを操作することができます。

第 4 章 データ・マイニングについて

データ・マイニングの概要

データ・マイニングでは、さまざまな手法を使用して、データ本体に関する貴重な情報を探し出します。データ・マイニングでは、意思決定支援、予測、推定などの領域で使われるさまざまな手法を利用して情報を取り出します。データ量は膨大ですが、利用価値の低いものやそのままの形では直接役に立たないものがあります。価値があるのは、データに隠されている貴重な情報です。

データ・マイニングは、データに関するユーザー（またはユーザー側の専門家）の知識と、高度で有効な分析手法が結びついたときに大きな成果を上げます。分析手法とは、データの基本的な関係と特性を、コンピュータを使用して識別する手法です。データ・マイニングの過程では、履歴データからモデルが作成されます。このモデルは、工程の予測やパターン検出などさまざまな目的で使用されます。これは、マシン学習またはモデル作成と呼ばれる、モデルを構築するための手法です。

モデル作成手法

IBM SPSS Modeler には、数多くのマシン学習およびモデル作成手法が含まれており、それらは、解くことができる問題の種類に応じておおまかに分類できます。

- 予測モデルリング手法には、ディジション・ツリー、ニューラル・ネットワーク、および統計モデルが含まれます。
- クラスタリング・モデルは、類似したレコードのグループを識別し、そのグループに従ってレコードにラベルを付けます。クラスタリング手法には、Kohonen、k-means、および TwoStep が含まれます。
- アソシエーション・ルールは、特定の結果（特定の製品の購入）と条件セット（複数の他の製品の購入）を関連付けます。
- スクリーニング・モデルを使用することにより、データ・スクリーニングを行って、モデル作成で関心の対象となる可能性が最高に高いフィールドとレコードを探し出したり、既知のパターンに適合しない可能性がある外れ値を識別したりできます。ここで使用できる方法には、機能選択と異常性検出が含まれます。

データの操作と発見

IBM SPSS Modeler はまた、さまざまな機能を使用して自分の専門知識をデータに適用できます。

- データの操作：既存のデータから新しいデータ項目を構成したり、データを意味のあるサブセットに分割します。さまざまなソースからのデータを結合したり、フィルタリングすることができます。
- 参照と視覚化：データ検査ノードを使用してデータのさまざまな側面を表示し、グラフや統計量などを含めた初期検査を行います。視覚化されたインタラクティブ グラフィックを利用したり、グラフィックをエクスポートしてプロジェクト・レポートに入れることができます。
- 統計：データ内の変数の間の予測された関係を確認します。IBM SPSS Modeler では、IBM SPSS Statistics からの統計量を使用することもできます。
- 仮説の検定：データの動作をモデル化し、これらのモデルを検査します。

通常、これらの機能は、データ中の有効な一連の属性を識別するために用いられます。次にこれらの属性を、基本的なルールと関係を識別するモデル作成手法に適用します。

典型的なアプリケーション

データ・マイニング・テクニックの典型的なアプリケーションには、次の例が含まれます。

ダイレクト・メール: どの人口統計的なグループが、最も高い応答率を持っているのかを識別する。この情報は、将来のメール送付の際に、応答を最大にするために使用できます。

信用情報のスコアリング: 個人の信用履歴を、ローン審査に使用する。

人事: 過去の雇用業務を理解して、雇用プロセスを効率化するディシジョン・ルールを作成する。

医学研究: 医学的な証拠に基づいて適当な処置を提案するディシジョン・ルールを作成する。

市場分析: 地域、価格、および顧客の特長などのどの変数が売り上げに関連しているかを識別する。

品質管理: 製品の製造行程からのデータを分析して、製品の欠陥を識別する変数を決定する。

政策研究: 最も重要な変数を選択するディシジョン・ルールを使用した政策決定で、調査データを使用する。

健康管理: ユーザー調査および臨床データを組み合わせて、健康に寄与している変数を発見する。

用語

属性、フィールド、および変数という用語は、検討中の全ケースに共通する一つのデータ項目を示します。特定のケースを示す属性値の集合は、レコード、例、またはケースといいます。

データへの評価

データ・マイニングでは、使用するデータが一定の基準を満たしていない限り、成果を期待できません。次の項で、データの側面をいくつか示し、検討すべきアプリケーションを示します。

データが使用できることを確認してください。

これは漠然としているように思われますが、データを使用できても、簡単に利用できる形式にはなっていない場合があることに注意してください。IBM SPSS Modeler では、データをデータベース (ODBC を介して) またはファイルからインポートできます。ただし、直接アクセスできないマシン上で、データが別の形式で保持されている場合があります。そのようなデータは、ダウンロードするか、適切な形式でダンプしないと、使用できません。データが異なるデータベースやソースに散在しており、それらを同時に取得する必要がある場合もあります。データは、オンライン上にあるとは限りません。データが紙にしか記されていない場合は、データ・マイニングを始める前に、データを入力する必要があります。

データが関連する属性をカバーしているかどうかの確認

データ・マイニングの目的は関連する属性を識別することなので、ここの確認などは奇妙に思われるかもしれませんが、使用できるデータを調べ、記録されていない関連因子を見つけることは、非常に役に立ちます。例えば、アイスクリームの販売数を予測するときに、小売店や過去の販売記録については大量の情報があっても、大きな影響を与えると思われる天候や気温については情報がないという場合があります。属性が欠けていても、データ・マイニングで有効な結果を出すことはできますが、結果として出された予測の精度が落ちる場合があります。

このような状況を評価するために、データの総合的な検査を行うことができます。まず始めに、データ・ソースにデータ検査ノードを接続して実行し、レポートを生成、調査することも検討してください。

ノイズを含んだデータの注意

データには、誤ったデータやさまざまに判断できる主観的なデータが含まれていることがよくあります。これらは一括してノイズと呼ばれます。データ内のノイズが正常な場合もあります。基本ルールがある場合もありますが、すべてのケースにあてはまるとは限りません。

通常、データにノイズが多くなると、それだけ正確な結果を得ることが難しくなります。しかし、IBM SPSS Modeler のマシン学習手法はノイズ データを処理することができ、50% 近くのノイズを含むデータセットに対しても有効に使用されています。

データが十分にあることを確認してください。

データ・マイニングでは、データ・セットのサイズは必ずしも重要ではありません。データ・セットの標本性、可能性のある結果がデータ・セットの標本範囲内に含まれるかどうか、および変数の組み合わせの方がはるかに重要です。

通常、考慮する属性が多くなると、有効な標本範囲を満たすのに必要なレコード数も多くなります。

データが標本として有効で、共通する基本ルールがある場合は、数千 (数百でも可) レコードのデータ サンプルでも、百万単位のデータ サンプルの場合と同様に、正確な結果を素早く得ることができます。

データのエキスパートの探索

多くの場合、ユーザーは自分のデータを処理するので、データの内容と意味はよく理解しています。しかし、別の部門や顧客のデータを処理する場合は、データについてよく知っている専門家の力を借りた方がはるかに効率が上がります。専門家は関連する属性の識別方法を指導したり、真に貴重な情報の集まりと、データ・セットの異常などの理由により生じた無益な情報を区別し、データマイニングの結果を解釈するための支援を行います。

データ・マイニングの戦略的方法

日常の企業努力と同様に、データ・マイニングは、計画的かつ体系的な方法で行えば、より大きい効果をあげることができます。IBM SPSS Modeler などの最新鋭のデータ・マイニング・ツールを使用したときでも、データ・マイニングの多くの作業で、熟練したビジネス・アナリストが注意深く処理を監視する必要があります。よりよい計画を策定するためには、次の事項を確認してください。

- 解決すべき重要な問題は何か?
- どのようなデータ・ソースを使用できるか、データのどの部分が現在の問題に関連しているか?
- データ・マイニングを開始する前に、どのような事前処理とデータの事前処理が必要か?
- どのようなデータ・マイニング手法を使用できるか?
- データ・マイニング分析の結果をどのように評価するか?
- データ・マイニングから最大限の情報を得るにはどうするか?

通常、データ・マイニング・プロセスは、すぐに複雑になっていく傾向があります。複雑なビジネス問題、複数のデータ・ソース、データ・ソースのさまざまなデータ品質、データ・マイニング手法の配列、データ・マイニングの結果を測定するさまざまな方法など、追跡する項目は多岐にわたります。

追跡を継続するには、データ・マイニングのために明確に定義されたプロセス・モデルがあると役に立ちます。プロセス・モデルを使用して、この項で前述した質問に回答し、重要なポイントが処理されていることを確認することができます。プロセス・モデルは、複雑なデータを処理しているときにユーザーが迷子にならないように、データ・マイニングのロード・マップとして機能します。

SPSS Modeler で使用を推奨しているデータ・マイニング・プロセス・モデルは、CRISP-DM (Cross-Industry Standard Process for Data Mining) です。名前からわかるように、このモデルは広範囲な産業やビジネス問題に適用できる一般モデルとして設計されています。

CRISP-DM プロセス・モデル

一般 CRISP-DM プロセス・モデルには 6 つのフェーズがあり、データ・マイニングの主要な課題を解決します。6 つのフェーズは、データ・マイニングをさらに大きなビジネスの実践に組み込むために設計された、循環するプロセス内に一緒に適合されます。

次に 6 つのフェーズを説明します。

- **ビジネスの理解**：データ・マイニングで最も重要なフェーズです。このフェーズには、ビジネス対象の決定、状況の評価、データ・マイニングの目標の決定、プロジェクト計画の作成があります。
- **データの理解**：データは、データ・マイニングの「生の素材」を提供します。このフェーズは、必要なデータ・リソースとそのリソースの特性を理解するためのフェーズです。このフェーズには、初期データの収集、データの説明、データの検査、データ品質の妥当性検証があります。「出力」パレットにあるデータ検査ノードは、データの理解に必要不可欠なツールです。
- **データの準備**：データ・リソースのカタログを作成したら、そのデータをデータ・マイニング用に準備する必要があります。準備作業には、データの選択、クリーニング、構成、構築、統合、およびフォーマットなどがあります。
- **モデリング**。これはデータ・マイニングの最重要部分で、複雑な分析手法を使用してデータから情報を取得します。このフェーズには、モデル化手法の選択、テスト設計の生成、モデルの構築、評価などがあります。
- **評価**：いったんモデルを選択したら、ビジネス目標に到達するために、データ・マイニングの結果がどのように役に立つかを評価することができます。このフェーズには、結果の評価、データ・マイニング・プロセスのレビュー、次のステップの決定があります。
- **展開**：これまでの努力が報われるとき、実りを収穫するときです。このフェーズでは、新しい知識を日常のビジネス処理に統合して、それまでのビジネス上の問題を解決します。このフェーズには、展開、モニターリング、メンテナンス、最終レポートの作成、プロジェクトのレビューがあります。

このプロセス・モデルには、重要なポイントがいくつかあります。まず、通常は前の段落で概要を示した順序で処理を進めますが、多くの場所で各フェーズは非線型的に相互に影響し合います。例えば、通常、データの準備はモデル作成の前に行われます。しかし、モデル作成フェーズで出された決定と収集された情報によっては、データの準備フェーズの再考が必要になる場合があります。データの準備フェーズを変更すると、次は新しいモデル作成を検討する必要があります。このように、2 つのフェーズは、適切に解決されるまで相互にフィードバックします。同様に、評価フェーズで誤った問いに答えようとしていたことがわかり、元のビジネスの理解から再評価しなければならない場合もあります。この時点でビジネスの理解を訂正しよりよい目標を設定してから、残りの処理を進めることができます。

2 番目の主要な点は、データ・マイニングの反復特性です。ユーザーの仕事が、単にデータ・マイニング・プロジェクトを計画し実行して、自分のデータを片付けてから家に帰るだけ、ということはほとんどありません。データ・マイニングを使用して顧客の要望を解決するというのが当面の目標です。データ・マイニングの 1 サイクルから知識を得ると、多くの場合は顧客の要望を判別してそれを満たすための新しい問い、新しい課題、新しい機会が生じます。これらの新しい問い、課題、機会は、データを再度マイニングすることで解決できます。このマイニングと新しい機会の識別プロセスは、ビジネスについての考え方の一部になり、ビジネス戦略全体の基礎になります。

ここでは、CRISP-DM プロセス・モデルについて簡単に概要だけ説明します。モデルの詳細は、次のリソースを参照してください。

- 『CRISP-DM ガイド』は、インストール・ディスクの ¥Documentation フォルダの他のドキュメントと共にアクセスできます。
- 「スタート」メニューから、または IBM SPSS Modeler 内で「ヘルプ」メニューから「CRISP-DM のヘルプ」をクリックして利用可能な、CRISP-DM ヘルプ・システム。

モデルの種類

IBM SPSS Modeler には、マシン学習、人工知能、および統計に基づいたさまざまなモデル作成方法が用意されています。「モデル作成」パレットを利用して、データから新しい情報を引き出したり、予測モデルを作成することができます。各手法によって、利点や適した問題の種類が異なります。

IBM SPSS Modeler アプリケーション ガイド では、これらの手法の多くの例が、モデル作成プロセスの概要とともに提供されています。このガイドは、オンライン・チュートリアル、および PDF 形式で使用できます。詳しくは、トピック 4 ページの『アプリケーションの例』を参照してください。

モデル作成方法は、次のカテゴリに分けられます。

- 監視
- アソシエーション
- セグメンテーション

監視モデル

監視モデルでは、1 つまたは複数の入力フィールドの値を使用し、1 つまたは複数の出力、または対象フィールドの値を予測します。これらの手法の例として、ディジション ツリー (C&R ツリー、QUEST、CHAID および C5.0 アルゴリズム)、回帰 (線型、ロジスティック、一般化線型、Cox 回帰アルゴリズム)、ニューラル・ネットワーク、サポート・ベクター・マシン、Bayesian ネットワークがあります。

組織は監視モデルを活用して、既知の結果に基づく予測に役立てることができます。例えば、顧客が購入するか立ち去るか、または特定の取引が既知の詐欺パターンに当てはまるかどうか、です。モデル作成手法には、マシン学習、ルール算出、サブグループ識別、統計的手法、および多重モデル生成が含まれます。

監視ノード



自動分類ノードは、2種類の結果 (yes/no、churn/don't churn など) を生じる多くの異なるモデルを作成および比較し、与えられた分析への最善のアプローチを選ぶことができますようになります。多くのモデル作成アルゴリズムに対応し、希望する方法、各特定のオプション、そして結果を比較するための基準を選択することができます。このノードで、指定されたオプションに基づいてモデルのセットが生成され、指定された基準に基づいて最善の候補がランク付けされます。



自動数値ノードでは、多くのさまざまな方法を使用し、連続する数値範囲の結果を求めてモデルを推定し比較します。このノードは、自動分類ノードと同じ方法で動作し、1 回のモデル作成のパスで、複数の組み合わせのオプションを使用し試すアルゴリズムを選択することができます。使用できるアルゴリズムには、ニューラル・ネットワーク、C&R Tree、CHAID、線型回帰、一般化線型回帰、サポート・ベクトル・マシン (SVM) が含まれています。モデルは、相関、相対エラー、または使用された変数の数に基づいて比較できます。



C&R Tree (分類と回帰ツリー) ノードは、ディジション・ツリーを生成し、将来の観測値を予測または分類できるようにします。この方法は再帰的なデータ区分を使用して学習レコードを複数のセグメントに分割し、各ステップで不純性を最小限に抑えます。ツリーのノードが「純粋」であると考えられるのは、ノード中にあるケースの 100% が、対象フィールドのある特定のカテゴリーに分類される場合です。対象フィールドおよび入力フィールドは、数値範囲またはカテゴリー (名義型、順序型、フラグ) が使用できます。すべての分岐は 2 分割です (2 つのサブグループのみ)。



QUEST ノードには、ディジション・ツリーの構築用に 2 分岐の方法が用意されています。これは、大規模な C&R ツリー分析が必要とする処理時間を短縮すると同時に、より多くの分割を可能にする入力値が優先される分類ツリー内の傾向を低減するように設計されています。入力フィールドは、数値範囲 (連続型) にできますが、目標変数はカテゴリーでなければなりません。すべての分割は 2 分岐です。



CHAID ノードはディジション・ツリーを生成し、カイ二乗統計値を使用して最適な分割を識別します。C&R ツリーおよび QUEST ノードと違って、CHAID は、非 2 分岐ツリーを生成できます。これは、ある分岐が 3 個以上のブランチを持つことを意味します。対象フィールドおよび入力フィールドは、数値範囲 (連続型) またはカテゴリーとなります。Exhaustive CHAID は CHAID の修正版で、可能性のある分割すべてを調べることで、よりよい結果を得られますが、計算時間も長くなります。



C5.0 ノードは、ディジション・ツリーとルール・セットのどちらかを構築します。このモデルは、各レベルで最大の情報の対応をもたらすフィールドに基づいてサンプルを分割します。対象フィールドは、カテゴリーでなければなりません。複数の分割を 2 つ以上のサブグループに分割できます。



ディジション・リスト・ノードは、母集団に関連する与えられた 2 値の結果の高いもしくは低い尤度を示すサブグループまたはセグメントを識別します。例えば、離れる可能性の少ないもしくはキャンペーンに好意的に答える可能性のある顧客を探すことができます。顧客区分を追加し、結果を比較するために他のモデルを並べて表示することによって、ビジネスに関する知識をモデルに導入することができます。ディジション・リスト・モデルは、ルール・リストから構成され、各ルールには条件と結果が含まれます。ルールは順番に適用され、一致する最初のルールで、結果が決まります。



線型回帰モデルは、対象と 1 つまたは複数の予測値との線型の関係に基づいて連続型対象を予測します。



因子分析ノードには、データの複雑性を整理する強力なデータ分解手法が 2 種類あります。主成分分析 (PCA) : 入力フィールドの線型結合が検出されます。成分が互いに直交する (直角に交わる) 場合に、フィールドのセット全体の分散を把握するのに役立ちます。因子分析 : 一連の観測フィールド内の相関パターンを説明する基本因子が識別されます。どちらの手法でも、元のフィールド・セットの情報を効果的に要約する少数の派生フィールドの検出が目標です。



フィールド選択ノードで、（欠損値の割合などの）諸基準に基づいて入力フィールドをスクリーニングして削除にかけ、指定した目標に相対的な残りの入力フィールドの重要度をランク付けします。例えば、数百の潜在的入力フィールドを含むデータセットがあるとして、患者予後のモデリングにはどれが役に立つのでしょうか？



判別分析によって、ロジスティック回帰より厳密な仮説を立てることができますが、これらの仮説が一致した場合、ロジスティック回帰分析に対する様々な代替あるいは補足になります。



ロジスティック回帰は、入力フィールドの値に基づいてレコードを分類する統計手法です。線型回帰と似ていますが、数値範囲ではなくカテゴリ対象フィールドを使用します。



一般化線型モデルは、指定したリンク関数によって従属変数が因子および共変量と線型関係になるよう、一般線型モデルを拡張したものです。さらにこのモデルでは、非正規分布の従属変数を使用することができます。線型回帰、ロジスティック回帰、カウント・データに関するログ線型モデル、そして区間打ち切り生存モデルなど、統計モデルの機能性が数多く含まれています。



一般化線型混合モデル (GLMM) は線型モデルを拡張したため、対象が非正規分布となる場合があり、指定されたリンク関数を介して因子および共変量に線形に関連し、観測が相関できるようにしました。一般化線型混合モデルには、単純な線型回帰から、非正規分布の縦断的データを取り扱う複雑なマルチレベル・モデルまで、さまざまなモデルがあります。



Cox 回帰ノードを使用すると、打ち切りレコードの存在下でイベントまでの時間のデータの生存モデルを構築します。モデルは、対象のイベントが入力変数の指定の値で指定の時間 (t) に発生する確率を予測する生存関数を作成します。



サポート・ベクター・マシン (SVM) ノードを使用すると、オーバーフィットすることなく、データを 2 つのグループのいずれかに分類することができます。SVM は、非常に多数の入力フィールドを含むデータセットなど、広範なデータセットを処理することができます。



ベイズ・ネットワーク・ノードを使用すると、観測された情報および記録された情報を実際の知識を組み合わせることによって確率モデルを作成し、発生の尤度を確立できます。ノードは主に分類に使用される Tree Augmented Naïve Bayes (TAN) および Markov Blanket ネットワークに焦点を当てています。



SLRM (自己学習応答モデル) ノードを使用するとモデルを構築でき、単一または少数の新しいケースを使用して全データを使用するモデルの保持をすることなく、モデルの再見積もりを行うことができます。



時系列ノードは、時系列から指数平滑法、1 変量の自己回帰型統合移動平均法 (ARIMA)、および多変量 ARIMA (または伝達関数) モデルを推測し、将来のパフォーマンスの予測を作成します。この時系列ノードは、SPSS Modeler バージョン 18 で使用されなくなった以前の時系列ノードと類似しています。ただし、この新しい時系列ノードは、IBM SPSS Analytic Server の機能を活用してビッグ データを処理するように設計されており、結果モデルは SPSS Modeler バージョン 17 で追加された出力ビューアーに表示されます。



k が整数である場合、 k 最近傍 (KNN) ノードは、新しいケースを、予測領域の新しいケースに最も近い k 個のオブジェクトのカテゴリまたは値と関連付けます。類似したケースはお互いに近く、類似していないケースはお互いに離れています。



時空間予測 (STP) ノードは、ロケーション・データ、予測用の入力フィールド (予測値)、時間フィールド、および対象フィールドを使用します。各ロケーションには、それぞれの測定時の各予測値を表すデータの行が多数あります。データを分析すると、そのデータを使用して、分析で使用される形状データ内の任意のロケーションの対象値を予測できます。

アソシエーション・モデル

アソシエーション・モデルでは、イベント、購入、属性など、1 つまたは複数のエンティティが 1 つまたは複数のその他のエンティティと関連するデータ内のパターンを検出します。モデルは、これらの関係性を定義するルール・セットを構築します。データ内のフィールドは、入力および対象のいずれのフィールドとしても機能します。これらのアソシエーションは手動で検出できますが、アソシエーション・ルール・アルゴリズムはより迅速に検出が可能で、より複雑なパターンも検証できます。Apriori および Carma モデルが、アソシエーション・アルゴリズムの使用例です。他にアソシエーション・モデルの 1 つとしてシーケンス検出モデルがあり、時間構造データのシーケンス・パターンを検索します。

アソシエーション モデルは、例えば、商品 X を購入した顧客は Y と Z も購入するなど、複数の結果を予測する場合に最も役立ちます。アソシエーション・ルールのアルゴリズムは、どのような属性の間にも関連を成立させることができるという点で、より一般的なディジション・ツリーのアルゴリズム (C5.0 や C&R Trees など) より勝っています。ディジション・ツリーのアルゴリズムは、一つの結果にいたるルールを構築するのに対し、アソシエーション・ルールのアルゴリズムは、それぞれが異なる結果にいたる多数のルールを見つけようとします。

アソシエーション・ノード



Apriori ノードで、データからルール・セットを抽出し、情報内容が最も充実したルールを引き出します。Apriori には、5 種類のルール選択方法があり、高度なインデックス作成方法を使用して、大きなデータ・セットが効率的に処理されます。大きな問題の場合は、一般に、Apriori の方が高速に学習できます。保持できるルール数に特に制限はありません。また、最大 32 の前提条件を持つルールを処理できます。Apriori では、入力フィールドと出力フィールドのすべてがカテゴリであることが必要ですが、この種類のデータに合わせて最適化されているので、よりよいパフォーマンスを実現します。



CARMA モデルは、入力または対象フィールドを指定しなくても、データからルール・セットを抽出します。Apriori とは対照的に、CARMA ノードでは、前提条件サポートだけではなく、ルール・サポート (前提条件と結果の両方のサポート) を対象とした構築の設定が可能です。これは、生成されたルールをさまざまなアプリケーションで活用できることを意味します。例えば、この休暇シーズンに販売促進する項目を結果とする、商品またはサービス (前提条件) のリストを調べることができます。



シーケンス・ノードで、シーケンシャルな、または時間経過が伴うデータ内のアソシエーション・ルールを検出します。予測可能な順序で起こる傾向にあるアイテム・セットのリストを、シーケンスと呼びます。例えば、顧客がひげそりとアフター・シェーブローションを購入した場合、その顧客は次の購入時にシェービングクリームを購入する可能性があります。シーケンス・ノードは CARMA アソシエーション・ルール・アルゴリズムに基づいているため、効率的な 2 段階通過法でシーケンスが検出されます。



アソシエーション ルール ノードは Apriori ノードに似ていますが、Apriori とは異なり、アソシエーション ルール ノードはリスト・データを処理できます。さらに、アソシエーション ルール ノードを IBM SPSS Analytic Server と共に使用すると、ビッグデータの処理や高速な並列処理の利用が可能になります。

セグメント化モデル

セグメント化モデルでは、データを入力フィールドの類似したパターンを持つレコードのセグメント、またはクラスターに分割します。入力フィールドにのみ関心があるため、セグメント・モデルには出力フィールドまたは対象フィールドの概念はありません。セグメント・モデルの例として、Kohonen ネットワーク、K-Means クラスタリング、TwoStep クラスタリングおよび異常値検出があります。

「クラスタリング・モデル」とも呼ばれるセグメント化モデルは、特定の結果が不明である場合に適しています (例えば、詐欺の新しいパターンを識別する場合や、既存の顧客ベースから関心の対象となるグループを識別する場合です)。クラスタリング・モデルは、類似したレコードのグループを識別し、そのグループに従ってレコードにラベルを付けます。この作業は、各グループとそれぞれの特性に関する事前の知識を活用せずに実行されます。これは、クラスタリング・モデルと他のモデル作成技法との違いであり、クラスタリング・モデルには、モデルが予測する定義済みの出力フィールドや対象フィールドはありません。これらのモデルには、正、誤という回答はありません。モデルの価値は、データのグループ構成を把握し、それらのグループについて役に立つ説明を提供できるかどうかで決まります。クラスタリング・モデルは、クラスターやセグメントを作成するためによく利用されます。このクラスターやセグメントは、後の分析で入力として使用されます (例えば、潜在的な顧客を、等質のサブグループに分類する方法です)。

セグメント化ノード



自動クラスター・ノードは、同様の特性を持つレコードのグループを識別するクラスタリング・モデルを推定し、比較します。ノードは他の自動化モデル作成ノードと同じように動作し、複数の組み合わせのオプションを単一のモデル作成の実行で検証できます。モデルは、クラスター・モデルの有用性をフィルタリングおよびランク付けする基本的な指標を使用して比較し、特定のフィールドの重要度に基づいて指標を提供します。



K-Means ノードで、データ・セットが異なるグループ (つまりクラスター) へ、クラスターリングされます。この方法で、固定数のクラスターを定義し、クラスターにレコードを繰り返し割り当てて、これ以上調整してもモデルが改善されなくなるまで、クラスターの中心を調整します。K-means では、結果を予測するのではなく、入力フィールドのセット内のパターンを明らかにするために、「非監視学習」として知られるプロセスが使用されます。



Kohonen ノードは、ニューラル・ネットワークの一種であり、データ・セットをクラスター化して異なるグループを形成する目的で使用できます。ネットワークの学習が完了すると、類似のレコードは出力マップで互いに近くに表示され、違いの大きいレコードほど離れたところに表示されます。強度の高いユニットを識別するために生成されたモデル内で、各ユニットが獲得した観察の数値を調べることができます。これは、適切なクラスター数についてのヒントになる場合があります。



TwoStep ノードで、2 段階のクラスター化手法が使用されます。最初のステップでは、データを 1 度通過させて、未処理の入力データを管理可能な一連のサブクラスターに圧縮します。2 番目のステップでは、階層クラスター化手法を使用して、サブクラスターをより大きなクラスターに結合させていきます。**TwoStep** には、学習データに最適なクラスター数を自動的に推定するという利点があります。また、フィールド・タイプの混在や大規模データ・セットも効率よく処理できます。



異常値検出ノードで、「正常な」データのパターンに合致しない異常ケースや外れ値を識別します。このノードで、外れ値が既知のパターンに当てはまらなかったり、何を探しているのかははっきりしなかったりする場合でも、外れ値を識別できます。

データベース内マイニング・モデル

IBM SPSS Modeler は、データベース・ベンダーから入手できる、Oracle Data Miner や Microsoft Analysis Services などのデータ・マイニングおよびモデル作成のツールとの統合をサポートしています。データベース内および IBM SPSS Modeler アプリケーション内のモデル、スコアおよびストア モデルすべての作成が可能です。詳細は、『IBM SPSS Modeler データベース内マイニング・ガイド』を参照してください。

IBM SPSS Statistics モデル

コンピューターに IBM SPSS Statistics をインストールしライセンスが付与されている場合、IBM SPSS Modeler 内から特定の IBM SPSS Statistics ルーチンにアクセスおよび実行して、モデルを作成およびスコアリングできます。

データ・マイニングの例

データ・マイニングについて学習する現実的で最良の方法は、最初に例を使うことです。「IBM SPSS Modeler アプリケーション・ガイド」には、多数のアプリケーション例が記載されています。これらの例により、具体的なモデルの作成方法と技術について、目的別に簡潔に説明しています。詳しくは、トピック 4 ページの『アプリケーションの例』を参照してください。

第 5 章 ストリームの構築

ストリームの構築の概要

IBM SPSS Modeler を使用したデータ・マイニングでは、ストリームと呼ばれる、一連のノードを通過するデータの流れを処理する過程に注目します。この一連のノードは、データに関して行う処理や操作を表し、ノード間のリンクはデータの流れ (データ・フロー) の方向を示します。通常は、データ・ストリームを使用して IBM SPSS Modeler にデータを読み込み、そのデータを一連の処理や操作を経由して、テーブルやビューアなどに出力します。

例えば、データ・ソースを開いて新しいフィールドを追加し、そのフィールドの値に基づいてレコードを選択して、結果をテーブルに表示する場合を考えてみましょう。この場合、データ・ストリームは次の 4 つのノードで構成されます。



可変長ノード：データをデータ・ソースから読み込むために使用します。



フィールド作成ノード：データ・セットに新しく計算されたフィールドを追加するために使用します。



条件抽出ノード：データ・ストリームから除外するレコードの選択基準を設定するために使用します。



テーブル・ノード：操作結果を画面に表示するために使用します。

データ・ストリームの作成

SPSS Modeler 特有のインターフェースにより、データ・ストリームのダイアグラムを使用して、データ・マイニング作業を視覚的に行うことができます。基本的には、データ・ストリームは次の手順で作成していきます。

- ストリーム領域にノードを追加する。
- ノードを接続してストリームを形成する。
- 任意のノードまたはストリーム・オプションを指定する。
- ストリームを実行する。

このセクションでは、ノードを使ったより複雑なデータ・ストリームの作成方法を詳細に説明していきます。また、ノードやストリームのオプションや設定方法も取り上げていきます。SPSS Modeler に付属のデ

モ・データ（プログラム インストール・ディレクトリーの Demos フォルダー）を使ったストリーム構築の例は、4 ページの『アプリケーションの例』を参照してください。

ノードの作業

ノードは、IBM SPSS Modeler でデータを探索するために用いられます。作業領域内のさまざまなノードが、それぞれさまざまなオブジェクトや処理を表しています。IBM SPSS Modeler ウィンドウの下部にあるパレットには、ストリームの構築に使われるすべてのノードが用意されています。

スーパーノードには、いくつかの種類があります。入力ノードはデータをストリームに投入し、ノード・パレットの「入力」タブにあります。プロセス ノードは、各データ・レコードおよびフィールドで操作を行い、パレットの「レコード設定」タブおよび「フィールド設定」タブに表示されます。出力ノードは、さまざまなデータの出力、グラフ、モデル結果を生成し、ノード・パレットの「グラフ」タブ、「出力」タブ、「エクスポート」タブに表示されます。モデル作成ノードは統計アルゴリズムを使用してモデル ナゲットを作成し、ノード・パレットの「モデル作成」タブ、(有効化されている場合は)「DB モデリング」タブに表示されます。詳しくは、トピック 16 ページの『ノード・パレット』を参照してください。

ノード間を接続してストリームを形成し、それを実行することによって、関係を視覚化して、結論を導き出します。ストリームはスクリプトに似ています。ストリームは保存して、さまざまなデータ・ファイルで再利用できます。

ストリーム・データを処理する実行可能なノードは、ターミナル・ノードと呼ばれます。モデル作成ノードまたは出力ノードがストリームまたはストリームブランチの最後に表示される場合、ターミナル・ノードとなります。高度なノードをターミナル・ノードに接続できません。

注: ノード・パレットはユーザー設定できます。詳しくは、トピック 204 ページの『ノード・パレットのカスタマイズ』を参照してください。

ストリームへのノードの追加

ノード・パレットからストリームにノードを追加するには、次の複数の方法があります。

- パレットのノードをダブルクリックする。注: ノードをダブルクリックすると、現在のストリームとノードが自動的に接続されます。詳しくは、トピック 37 ページの『ストリーム内のノードの接続』を参照してください。
- パレットからストリーム領域にノードをドラッグ・アンド・ドロップする。
- パレット上のノードをクリックして、次にストリーム領域をクリックする。
- IBM SPSS Modeler の「挿入」メニューから適切なオプションを選択する。

ノードをストリーム領域に追加したら、そのノードをダブルクリックすると、ノードに関するダイアログ・ボックスが表示されます。使用できるオプションは、追加するノードの種類によって異なります。ダイアログ・ボックス内の特定のコントロールについての情報を表示するには、「ヘルプ」ボタンをクリックします。

ノードの削除

データ・ストリームからノードを削除するには、ノードをクリックして Del キーを押すか、または、右クリックしてメニューから「削除」を選択します。

ストリーム内のノードの接続

ストリーム領域に追加した各ノードは、接続しないとデータ・ストリームにはなりません。ノード間の接続は、ある操作から別の操作へとデータの流れる方向を表しています。ノード間を接続してデータ・ストリームを作成するには、さまざまな方法があります。ダブルクリック、マウスの中央ボタンの使用、または手作業

ダブルクリックでノードを追加、接続するには

データ・ストリームを形成するもっとも簡単な方法は、パレット上のノードをダブルクリックすることです。この操作を行うと、ストリーム領域上で選択されているノードに、ダブルクリックしたノードが自動的に接続されます。例えば、ストリーム領域にデータベース・ノードがある場合、このノードを選択した後に、フィールド作成ノードなどのパレット上のノードをダブルクリックします。これによって、既存のデータベース・ノードにフィールド作成ノードが自動的に接続されます。この手順を、ヒストグラムやテーブル・ノードなどのターミナル・ノードに到達するまで繰り返すことができます。ターミナル・ノードでは、任意のノードが最後の非ターミナル・ノードに接続されます。

マウスの中央ボタンを使用してノードを接続するには

ストリーム領域で、あるノードをマウスの中央ボタンを使用してクリックした後、別のノードにドラッグします(マウスの中央ボタンがない場合は、Alt キーを押しながらクリック、ドラッグしてください)。

ノードを手作業で接続するには

マウスの中央ボタンがなく、手作業でノードを接続したい場合は、ノードのポップアップ・メニューを使用して、ストリーム領域上の他のノードに接続することができます。

1. 接続を開始するノードを右クリックします。右クリックすると、ノード・メニューが表示されます。
2. メニューで「接続」をクリックします。
3. 始点ノードとカーソルの両方に接続アイコンが表示されます。領域上の目的のノードをクリックすると、2つのノードが接続されます。

ノードを接続する際には、注意しなければいけない事柄があります。次のような接続を作成しようとすると、エラー・メッセージが表示されます。

- 入力ノードへの接続
- ターミナル・ノードからの接続
- ノードの接続数が最大入力接続数を超える場合
- すでに接続されている2つのノードの接続
- データがすでに通過したノードに戻る場合 (循環性)

ストリーム内でのノードの迂回

データ・ストリーム内でノードを迂回する場合、そのノードのすべての入力接続と出力接続が、そのノードの入力ノードから出力ノードに直接接続されます。ノードに入力接続または出力接続がない場合、ノードの接続はすべて、経路は変更されずに破棄されます。

例えば、新規フィールドを作成し、フィールドのフィルタリングを行った後、結果をヒストグラムとテーブルに表示するストリームを考えてみましょう。フィールドをフィルタリングする前にもデータのグラフとテーブルを参照したい場合、ストリームに新規にヒストグラム・ノードとテーブル・ノードを追加するか、またはフィルター・ノードを迂回することができます。フィルター・ノードを迂回すると、グラフとテーブルはフィールド作成ノードから直接接続されます。フィルター・ノードの接続は解除されます。

ノードを迂回するには

1. ストリーム領域で、迂回するノードをマウスの中央ボタンでダブルクリックします。または、Alt キーを押しながらダブルクリックすることもできます。

注：この操作を元に戻すには、「編集」メニューの「元に戻す」をクリックするか、または Ctrl+Z キーを押してください。

ストリーム内でのノードを無効化

ストリーム内の 1 回の入力によるプロセス・ノードを無効化できます。その結果として、ストリームの実行時にノードが無視されます。これにより、ノードを削除またはバイパスする必要がなくなり、残りのノードを接続したままにできます。ノード設定を開いて編集することもできますが、ノードを再度有効化した後で変更が有効となります。

例えば、複数のフィールドをフィルタリングするストリームがあり、減らされたデータ・セットによりモデルを作成するとします。それとは別に、フィールドをフィルタリングせずに同じモデルを作成して、モデルの結果が改善されるかどうかを確認する場合は、フィルター・ノードを無効化できます。フィルター・ノードを無効化すると、モデル作成ノードへの接続は、フィールド作成ノードからデータ型ノードに直接通過します。

ノードを無効化するには

1. ストリーム領域で、無効にするノードを右クリックします。
2. ポップアップ・メニューの「ノードを無効化」をクリックします。

または、「編集」メニューの「ノード」 > 「ノードを無効化」を選択することもできます。ストリームにノードを戻す場合、同じようにして「ノードを有効化」を選択します。

注：この操作を元に戻すには、「編集」メニューの「元に戻す」をクリックするか、または Ctrl+Z キーを押してください。

既存の接続へのノードの追加

2 つのノードを接続する矢印をドラッグすることにより、接続された 2 つのノードの間に新しいノードを追加できます。

1. マウスの中央ボタンを使用して、接続の矢印をクリックし、ノードの挿入先までドラッグします。中央ボタンがない場合は、Alt キーを押しながらクリック・アンド・ドラッグすることもできます。
2. ボタンを押したまま挿入するノードまで接続をドラッグし、マウス・ボタンを離します。

注：ノードを迂回すれば、ノードから新しい接続を削除して、元の接続を復元できます。

ノード間の接続の削除

2 つのノード間の接続を削除する手順は、次のとおりです。

1. 接続の矢印を右クリックします。
2. メニューで「接続の削除」をクリックします。

ノードとのすべての接続を削除するには、次のいずれかを実行します。

- ノードを選択して、F3 を押します。
- ノードを選択して、メイン・メニューで次の各項目をクリックします。

「編集」 > 「ノード」 > 「接続の解除」

ノードのオプションの設定

ノードを作成して接続したら、ノードをカスタマイズするためのさまざまなオプションがあります。ノードを右クリックして、次のいずれかのオプションを選択します。

- 選択したノードのダイアログ・ボックスを開くには、「編集」をクリックします。
- あるノードから別のノードに手作業で接続するには、「接続」をクリックします。
- 選択したノードに対するすべてのリンクを削除するには、「接続の解除」をクリックします。
- 「編集」ダイアログ・ボックスの「注釈」タブを表示するには、「名前の変更と注釈」をクリックします。
- 「新規コメント」をクリックして、ノードに関連するコメントを追加します。詳しくは、トピック 58 ページの『ノードおよびストリームへのコメントおよび注釈の追加』を参照してください。
- 「ノードを無効化」をクリックして、処理時にノードを「非表示」にします。処理する場合にノードを再度表示するには、「ノードを有効化」をクリックします。詳しくは、トピック 38 ページの『ストリーム内でのノードを無効化』を参照してください。
- ストリーム領域からノードを削除するには、「切り取り」または「削除」をクリックします。注：「切り取り」をクリックした場合、そのノードを貼り付けることができます。ただし、「削除」を選択した場合は、貼り付けることはできません。
- 接続を持たないノードのコピーを作成するには、「ノードをコピー」をクリックします。コピーしたノードは、既存のストリームに追加することも、新しいストリームに追加することもできます。
- 以前に保存したノードを開いて、そのオプション設定を現在選択しているノードに取り込むには、「ノードをロード」をクリックします。この操作を行うには、ノードの種類が同じでなければなりません。
- 「ノードの取得」をクリックして、接続した IBM SPSS Collaboration and Deployment Services Repository からノードを取得します。
- ノードの詳細をファイルに保存するには、「ノードの保存」をクリックします。保存したノード定義は、同じ種類のノードにしかロードできません。
- 「ノードを格納」をクリックして、接続した IBM SPSS Collaboration and Deployment Services Repository の選択したノードを格納します。
- 選択したノードのキャッシュに関するオプションを設定するには、「キャッシュ」をクリックしてメニューを表示します。
- データの新規ソースへのマッピングや必須フィールドの指定などのオプションを設定するには、「データ マッピング」をクリックしてメニューを表示します。
- 現在のストリームにスーパーノードを作成するには、「スーパーノードの作成」をクリックしてメニューを表示します。
- 選択したノードを置き換えるには、「ユーザー入力ノードの生成」をクリックします。このノードは、元のノードと同じフィールドを持ちます。
- 選択したノードの下流にあるすべてのターミナル・ノードを実行するには、「ここから実行」をクリックします。

ノードのキャッシュ・オプション

ストリームの実行を最適化するために、ターミナル・ノード以外の任意のノードにキャッシュを設定することができます。ノードにキャッシュを設定すると、次のデータ・ストリームの実行時にそのノードを通過するデータがキャッシュされます。データがキャッシュされたら、以降のデータはデータ・ソースの代わりに、キャッシュ (ディスクの一時ディレクトリーに格納) から読み込まれます。

キャッシュは、ソート、結合、集計など、時間のかかる操作に最も有用です。例えば、データベースから販売データを読み込む入力ノードと、販売地域ごとに集計するレコード集計ノードがあるとします。この場合、データ・セット全体ではなく集計済みデータをキャッシュしたいので、入力ノードではなくレコード集計ノードにキャッシュを設定します。

注: 入力ノードでキャッシュする場合、IBM SPSS Modeler に読み込まれるように元のデータのコピーを保存するため、多くの環境でパフォーマンスは改善されません。

キャッシュを有効にしたノードの右上には、小さい文書アイコンが表示されます。データがキャッシュされると、この文書アイコンが緑に変わります。

キャッシュを有効にするには

1. ストリーム領域でノードを右クリックし、メニューの「キャッシュ」をクリックします。
2. キャッシュ設定用のサブメニューから、「使用する」を選択します。
3. キャッシュを解除するには、ノードを右クリックして、キャッシュ設定用のサブメニューから「使用しない」を選択します。

データベース内でのノードのキャッシュ

データベース内で実行されるストリームの場合、ファイル・システムでなくデータベース内の一時テーブルへ、データを中流でキャッシュできます。この機能を SQL 最適化と組み合わせると、パフォーマンスが著しく向上する可能性があります。例えば、データ・マイニング・ビューを作成するために複数のテーブルを併合するストリームからの出力をキャッシュし、必要に応じて再使用できます。すべての下流ノードで自動的に SQL を生成するようにすると、パフォーマンスがさらに向上する可能性があります。

データベースのキャッシングを利用するには、SQL 最適化とデータベース・キャッシングの両方を有効にする必要があります。Client の設定は、Server の最適化設定によって無効になるので注意してください。詳しくは、トピック 45 ページの『ストリームの最適化オプションの設定』を参照してください。

データベースのキャッシングが有効にされると、任意の非ターミナル・ノードを右クリックするだけでその場所のデータがキャッシュされ、次にストリームが実行されるときに自動的に、データベース内にキャッシュが直接作成されます。データベース・キャッシングまたは SQL 最適化が有効にされていないと、キャッシュは、代わりにファイル・システムへ書き出されます。

注: 次のデータベースは、キャッシュのために一時テーブルをサポートします。

DB2、Netezza、Oracle、SQL Server および Teradata。その他のデータベースでは、データベース・キャッシングに通常のテーブルが使用されます。SQL コードを特定のデータベース向けにカスタマイズできません。サービスに連絡してください。

キャッシュを取り消すには

ノードに表示されている文書アイコンが白い場合、キャッシュが空であることを表しています。キャッシュがいっぱいになると、文書アイコンは濃い緑色になります。キャッシュの内容を置き換える場合は、最初にキャッシュを取り消してからデータ・ストリームを再実行して、キャッシュにデータを格納しなおす必要があります。

1. ストリーム領域でノードを右クリックし、メニューの「キャッシュ」をクリックします。
2. キャッシュ設定用のサブメニューから、「取り消し」を選択します。

キャッシュを保存するには

キャッシュの内容は、IBM SPSS Statistics データ・ファイル (*.sav) として保存できます。その後、ファイルをキャッシュとして再ロードしたり、キャッシュ・ファイルをデータ・ソースとして使用するノードを設定することができます。さらに、別のプロジェクトで保存したキャッシュをロードすることもできます。

1. ストリーム領域でノードを右クリックし、メニューの「キャッシュ」をクリックします。
2. キャッシュ設定用のサブメニューから、「キャッシュの保存」を選択します。
3. 「キャッシュの保存」ダイアログ・ボックスで、キャッシュ・ファイルを保存する場所を指定します。
4. 「ファイル名」テキスト・ボックスに名前を入力します。
5. 「ファイルの種類」リストで「*.sav」が選択されていることを確認し、「保存」をクリックします。

キャッシュをロードするには

キャッシュ・ファイルをノードから削除する前に保存した場合、それを再ロードすることができます。

1. ストリーム領域でノードを右クリックし、メニューの「キャッシュ」をクリックします。
2. キャッシュ設定用のサブメニューから、「キャッシュのロード」を選択します。
3. 「キャッシュのロード」ダイアログ・ボックスで、目的のキャッシュ・ファイルを選択し、「ロード」をクリックします。

ノードのデータをプレビューする

ストリーム構築時に予測していたようにデータが変更されているか確認するために、各重要なステップごとにテーブル・ノードを使用してデータを実行する場合があります。このような手間を省くために、ノード適用時に作成されるデータのサンプルを表示する各ノードからプレビューを生成し、各ノードの構築時間を短縮できます。

モデル ナゲットのノード上流の場合、プレビューでは入力フィールドを表示します。モデルナゲット、またはナゲットのノードの下流 (ターミナル・ノードを除く) の場合、プレビューでは入力フィールドと生成されたフィールドが表示されます。

表示される行のデフォルト数は 10 行です。ただし、ストリームのプロパティーで変更できます。詳しくは、トピック 43 ページの『ストリームの一般的なオプションの設定』を参照してください。

「生成」メニューから、複数の種類のノードを作成できます。

注: このノードによって生成されたデータをプレビューすると、すべてのプロパティー変更がこのノードに適用され (「適用」をクリックした場合と同じ動作)、これをキャンセルすることはできません。

ノードのロック

その他のユーザーがストリームの 1 つまたは複数のノード設定を修正できないようにするために、スーパーノードという特別な種類のノードにノードをカプセル化して、パスワード保護を適用してノードをロックすることができます。

ストリームの作業

ストリーム領域上の入力ノード、プロセス・ノード、およびターミナル・ノードを接続したら、データ・ストリームの作成は完了です。ノードの集合となるストリームは、保存したり、注釈を付けたり、プロジェクトに追加することができます。また、最適化、日付と時刻の設定、パラメーター、およびスクリプトなどのさまざまなオプションを設定することができます。これらのプロパティーの詳細は後述します。

IBM SPSS Modeler では、同じ IBM SPSS Modeler セッションの複数のデータ・ストリームを使用および変更することができます。メイン・ウィンドウの右側には、現在表示しているストリーム、出力、およびモデルを把握するために役立つマネージャ (管理) パネルが用意されています。マネージャー・パネルが表示されない場合、「表示」メニューの「マネージャー」をクリックし、「ストリーム」タブをクリックします。

このタブから、次のような作業を行うことができます。

- ストリームへのアクセス。
- ストリームの保存。
- 現在のプロジェクトへのストリームの保存。
- ストリームを閉じる。
- 新規ストリームを開く。
- IBM SPSS Collaboration and Deployment Services リポジトリでストリームの保存と検索を行う (ユーザーのサイトで利用可能な場合)。詳しくは、トピック 163 ページの『IBM SPSS Collaboration and Deployment Services Repository について』を参照してください。

これらの設定を利用するには、「ストリーム」タブでストリームを右クリックしてください。

ストリームのオプションの設定

現在のストリームに適用するさまざまなオプションを指定することができます。また、これらのオプションをすべてのストリームにデフォルトとして適用するよう保存することができます。オプションは次の通りです。

- **全般:** ストリームで使用する記号やテキストのエンコードなどのその他のオプション。詳しくは、トピック 43 ページの『ストリームの一般的なオプションの設定』を参照してください。
- **日時:** 日付と時刻の表現形式に関連するオプション。詳しくは、トピック 44 ページの『ストリームの日時オプションの設定』を参照してください。
- **数値書式:** 数値表現の形式を制御するオプション。詳しくは、トピック 45 ページの『ストリームの数値形式オプションの設定』を参照してください。
- **最適化:** ストリームのパフォーマンスを最適化するためのオプション。詳しくは、トピック 45 ページの『ストリームの最適化オプションの設定』を参照してください。
- **ロギングとステータス:** SQLのロギングおよび記録の状態を制御するオプション。詳しくは、トピック 47 ページの『ストリームの SQL ロギングおよびレコード ステータスオプションの設定』を参照してください。
- **レイアウト:** 領域上のストリームのレイアウトに関連するオプション。詳しくは、トピック 48 ページの『ストリームのレイアウト・オプションの設定』を参照してください。
- **Analytic Server.** SPSS Modeler での Analytic Server の使用に関連するオプション。詳しくは、トピック 48 ページの『Analytic Server のストリームのプロパティ』を参照してください。
- **地理空間.** ストリームで使用する地理空間データの書式設定に関連するオプション。詳しくは、トピック 49 ページの『ストリームの地理空間オプションの設定』を参照してください。

ストリーム・オプションを設定するには

1. 「ファイル」メニューから、「ストリームのプロパティ」を選択します (またはマネージャー・パネルの「ストリーム」タブからストリームを選択し、ポップアップ・メニューを右クリックして「ストリームのプロパティ」をクリックします)。
2. 「オプション」タブをクリックします。

また、「ツール」メニューから次の各項目を選択することもできます。

「ストリームのプロパティ」 > 「オプション」

ストリームの一般的なオプションの設定：一般的なオプションは、現在のストリームのさまざまな面に適用されるその他のオプションのセットです。

「基本」セクションには以下の基本オプションがあります。

- 小数点記号：桁区切り記号として、コンマ (,)、またはピリオド (.) を選択します。
- グループ化記号：数字の表示形式で、値をグループ化するのに使用する記号を選択します (例: 3,000.00 のコンマ)。オプションには、なし、ピリオド、コンマ、スペース、およびロケール定義 (現在のロケールがデフォルトとして使用されている場合)。
- エンコード：テキスト・エンコードについて、ストリームのデフォルトの方法を指定します。(注：可変長ファイル・ソース・ノードとフラット・ファイル・エクスポート・ノードだけに適用されます。多くのデータ・ファイルには埋め込み文字コード情報があります)。システム・デフォルトまたは UTF-8 のいずれかを選択できます。システム・デフォルトは、Windows のコントロール・パネル、または分散モードで動作している場合はサーバー・コンピューターから指定できます。詳しくは、トピック 227 ページの『IBM SPSS Modeler 内での Unicode のサポート』を参照してください。
- ルールセットの評価：ルールセット・モデルの評価方法を決定します。デフォルトでは、ルールセットは「票決」を使用して個別のルールからの予測を組み合わせ、最終的な予測を決定します。ルールセットが該当する最初のルールを使用するように設定するには、「最初のヒット」を選択します。このオプションは、ディジション・リスト・モデルには適用されません。このモデルでは、アルゴリズムに定義されたように、最初のヒットが常に使用されます。

データ プレビューに表示する最大行数：データのプレビューがノードに必要な場合に表示する行の最大数。詳しくは、トピック 41 ページの『ノードのデータをプレビューする』を参照してください。

名義型フィールドの最大メンバー：名義型 (セット型) フィールドの最大メンバー数を指定する場合に選択します。これ以降のフィールドのデータ型は、「データ型不明」になります。このオプションは、特に大きい名義型フィールドで作業を行うような場合に役立ちます。注：測定の尺度が「データ型不明」に設定されると、役割が自動的に「なし」に設定されます。つまり、このフィールドをモデリングには利用できません。

Kohonen および K-Means モデリングのセット・サイズを制限：Kohonen ネットワークおよび K-Means モデリングに使用される名義型フィールドの最大メンバー数を指定する場合に選択します。デフォルトは 20 です。以降のフィールドは無視され、フィールドに問題があることを示す警告が発せられます。

互換性を維持するために、このオプションは IBM SPSS Modeler のバージョン 14 で置き換えられた旧式のニューラル・ネットワーク・ノードにも適用されること、古いストリームには現在でもこのノードが含まれている場合があることに注意してください。

実行時に入力ノードをリフレッシュ：現在のストリームの実行時に、すべての入力ノードを自動的にリフレッシュする場合に指定します。この操作は、現在のストリーム上にあるすべての入力ノード (ユーザー入力ノードを除く) が自動的にリフレッシュされることを除いては、入力ノードで「リフレッシュ」ボタンをクリックするのと同じ働きをします。

注：このオプションを選択すると、データが変更されていない場合でも、下流にあるノードのキャッシュが消去されます。ツールバーから「現在のストリームを実行」オプションを使用すると、キャッシュの消去はストリームの実行ごとに一度だけ行われます。つまり、1 回の実行において下流のキャッシュを一時ストレージとして使用することができます。例えば、複雑なフィールド作成操作の後に中流にキャッシュを作成

し、このフィールド作成ノードの下流に複数のグラフやレポートを追加した場合を考えてみましょう。ストリームの実行時に、フィールド作成ノードのキャッシュは消去され、その後最初のグラフまたはレポートがキャッシュに格納されます。以降のターミナル・ノードは、そのデータをフィールド作成ノードのキャッシュから読み込みます。「現在のストリームを実行」オプションを使用せず、各ターミナル・ノードを個別に実行すると（ターミナル・ノードが複数存在する場合）、キャッシュの消去はターミナル・ノードを実行するたびに行われることに注意してください。

出力中のフィールドと値ラベルを表示する：テーブル中のフィールド、値ラベル、グラフ、およびその他の出力を表示します。ラベルが存在しない場合、フィールド名とデータ値が代わりに表示されます。デフォルトでは、ラベルはオフになっていますが、IBM SPSS Modeler の他の場所で必要になった場合には、個々のラベルごとにオンとオフを切り替えることができます。ツールバーから利用できるトグル ボタンを使用して、出力ウィンドウにラベルを表示させることもできます。



図 11. フィールドラベルと値ラベルの切り替えに使われるツールバー・アイコン

実行時間を表示：ストリームを実行した後、「実行時間」タブにストリーム・ノードの各実行時間を表示します。詳しくは、トピック 50 ページの『ノード実行時間の表示』を参照してください。

「自動ノード作成」セクションには、各ストリームで自動的にノードを作成するための以下のオプションがあります。これらのオプションでは、新規ナゲットの生成時にストリーム領域上にモデル作成ナゲットを挿入するかどうかを制御します。デフォルトでは、これらのオプションはバージョン 16 以降で作成されたストリームにのみ適用されます。IBM SPSS Modeler 16 以降で、バージョン 15 以前で作成されたストリームを開き、モデル作成ノードを実行すると、ナゲットは前のリリースのときのようにストリーム領域に置かれなくなります。IBM SPSS Modeler 16 以降を使用して新規ストリームを作成し、モデル作成ノードを実行すると、生成されたナゲットはストリーム領域に置かれます。これは設計どおりの動作です。こうした設計になった理由は、例えば、「新規モデル出力のモデル適用ノードの作成」オプションは、バッチで、IBM SPSS Collaboration and Deployment Services で、および、IBM SPSS Modeler Server クライアントのユーザー インターフェイスが存在しないその他の環境で実行されているバージョン 16 より前のストリームを中断させる可能性があるからです。

- 「新規モデル出力のモデル適用ノードの作成」。新規モデル出力のモデル適用ノードを自動的に作成します。このオプションを選択した場合は、「モデル更新リンクの作成」から、リンクを有効にするか無効にするか、あるいはリンクを作成しないかを選択できます。

新規モデル アプライヤまたは入力ノードを作成する際は、ドロップダウンのリンク オプションによって、ビルダー ノードと新規ノードとの間の更新リンクを作成するかどうかを制御し、さらに、作成する場合はそのリンクのモードを制御します。リンクを作成した場合、おそらくそれらのリンクを有効にすることになりますが、これらのオプションによってユーザーはすべてを制御できるようになります。

- 「ソース ビルダーから入力ノードを作成」。入力ノードをソース ビルダーから自動的に作成します。前のオプションと同様に、このオプションを選択すると、「ソース リフレッシュ リンクの作成」ドロップダウンから、リフレッシュ リンクを有効にするか無効にするか、あるいはリンクを作成しないかを選択できます。

「デフォルトとして保存」。指定したオプションは、現在のストリームのみにも適用されます。このボタンをクリックすると、これらのオプションをすべてのストリームのデフォルトとして設定します。

ストリームの日時オプションの設定：これらのオプションは、現在のストリーム内のさまざまな日付と時刻の表現に使用する書式を指定します。

日付/時刻のインポート形式: 日付/時刻フィールドに日付/時刻ストレージを使用するか、または日付/時刻を文字列変数としてインポートするかを選択します。

日付のフォーマット: 日付ストレージ フィールド用に使用され、また文字列が CLEM 日付関数によって日付として解釈された場合に使用される日付の形式を選択します。

時間のフォーマット: 時間ストレージ フィールド用に使用され、また文字列が CLEM 時間関数によって時間として解釈された場合に使用される時間の形式を選択します。

日/分をロールオーバー: 時間のフォーマットで、マイナスの時差を、前の日または時間を表すものとして解釈するかどうかを選択します。

基準日付 (1月1日): CLEM 日付関数で単一日付を扱う場合に使用する基準の年 (月日は常に 1 月 1 日) を選択します。

次世紀までのロールオーバーの閾値: 2 桁で年号を表す場合に、前の世紀の年号かどうかを判断するための、閾値とする年を指定します。例えば、閾値の年として 1930 を指定すると 05/11/02 が 2002 年であると想定されます。同じ設定で 30 以降の日付に 20 世紀を使用するため、05/11/73 は 1973 年であると想定されます。

時間帯: `datetime_now` CLEM 式で使用する時間帯の選択方法を選択します。

- 「サーバー」を選択した場合、時間帯は以下の項目に依存します。
 - 現行ストリームが Analytic Server データ ソースを使用している場合、`datetime_now` 式は Analytic Server の時刻を使用します。デフォルトでは、サーバーは協定世界時を使用します。
 - 現行ストリームがデータベース ソース ノードを使用している場合、サポートされるデータベースは SQL プッシュバックを使用し、`datetime_now` 式ではデータベースの時刻が使用されます。
 - その他のすべてのストリームでは、時間帯に SPSS Modeler Server の時刻が使用されます。
- 「**Modeler Client**」を選択した場合、時間帯には SPSS Modeler がインストールされているマシンの時間帯の詳細が反映されます。
- 時間帯に任意の協定世界時の値を選択することもできます。

デフォルトとして保存: 指定したオプションは現行ストリームのみ適用されます。このボタンをクリックすると、これらのオプションをすべてのストリームのデフォルトとして設定します。

ストリームの数値形式オプションの設定: これらのオプションは、現在のストリーム内のさまざまな数値の表現に使用する書式を指定します。

数字の表示形式 :標準 (#####.###)、科学的 (#.###E+###)、または通貨表記形式 (\$###.##) から選択できます。

小数点以下の表示 (標準、科学的、通貨) : 数値の表示形式で、実数を表示または印刷するときに使用する、小数点以下の桁数を指定します。このオプションは、各表示形式ごとに別々に指定します。

計算単位: CLEM 式で使用する三角関数の角度単位として、「ラジアン」または「度数」を選択します。詳しくは、トピック 139 ページの『三角関数』を参照してください。

「デフォルトとして保存」。指定したオプションは、現在のストリームのみ適用されます。このボタンをクリックすると、これらのオプションをすべてのストリームのデフォルトとして設定します。

ストリームの最適化オプションの設定: 最適化設定を使用して、ストリームのパフォーマンスを最適化することができます。IBM SPSS Modeler Server が使用されている場合のパフォーマンスと最適化の設定

は、クライアントの設定より優先されます。これらの設定がサーバーで無効になっている場合、クライアントが有効に設定することはできません。ただし、これらの設定がサーバーで有効になっている場合は、クライアントが無効に設定することは可能です。

注：データベース・モデル作成および SQL 最適化では、IBM SPSS Modeler Server 接続が IBM SPSS Modeler コンピューター上で可能でなければなりません。この設定を有効にすると、データベース・アルゴリズムにアクセスし、IBM SPSS Modeler から SQL を直接プッシュバック、IBM SPSS Modeler Server にアクセスできます。現在のライセンスステータスを確認するには、IBM SPSS Modeler メニューから次を選択します。

「ヘルプ」 > 「バージョン情報」 > 「その他の詳細」

接続が有効な場合、「ライセンス ステータス」タブにオプション「サーバーの有効化」が表示されます。

詳しくは、トピック 10 ページの『IBM SPSS Modeler Server に接続中』を参照してください。

注: SQL プッシュバックおよび最適化がサポートされているかどうかは、使用しているデータベースの種類によって異なります。IBM SPSS Modeler での使用がサポートおよびテストされているデータベースおよび ODBC ドライバーの最新情報については、当社サポート・サイト (<http://www.ibm.com/support>) を参照してください。

ストリームのリライトを有効にする：このオプションを選択すると、IBM SPSS Modeler のストリームのリライトが有効になります。リライトには 4 種類あり、それらの 1 つ以上を選択することができます。ストリームのリライトは、ストリーム内のノードをバックグラウンドで並べ替え、ストリームの意味論を変更することなくストリームをより効率的に操作できるようにする処理です。

- **SQL 生成の最適化**：このオプションを選択すると、より多くのストリーム操作がデータベースにプッシュバックされて SQL 生成を使用して実行されるように、ストリーム内のノードが並べ替えられます。SQL へ変換できないノードが検出されると、SQL へ変換可能な下流にあるかどうか、またそれを、ストリームの意味論に影響を与えずに安全に問題のノードの前に移動できるかどうかを調べることができます。データベースの方が IBM SPSS Modeler より効率的に操作を実行できるだけでなく、このようなプッシュバックによって IBM SPSS Modeler で処理するために返されるデータ・セットのサイズも小さくすることができます。この結果、ネットワークトラフィックを減らし、ストリーム操作を高速化することができます。SQL 最適化を有効にするには、「SQL 生成」チェックボックスをオンにする必要があります。
- **CLEM 式の最適化**：このオプションを選択すると、最適化プログラムは、ストリームの実行前に事前処理できる CLEM 式を検索して処理速度を高めることができます。単純な例として、 $\log(\text{salary})$ という式が指定されると、最適化プログラムは、実際の salary 値を計算してそれを処理のために引き渡します。このようにすると、SQL プッシュバックと IBM SPSS Modeler Server の両方のパフォーマンスを向上させることができます。
- **シンタックスの実行を最適化**：この方法のストリームのリライトでは、IBM SPSS Statistics シンタックスを含む複数のノードを結合する操作を効率化します。各操作それぞれを実行する代わりに、シンタックス・コマンドを 1 つの操作に結合することによって最適化が行われます。
- **その他の実行を最適化**：この方法のストリームのリライトでは、データベースで処理できない操作を効率化します。最適化は、ストリームのデータ量をできるだけ早く減らすことによって実現されます。データの整合性を保ちながら、操作をデータ・ソースの近くに置いて、下流で結合などの時間のかかる操作に使用されるデータの量を減らすように、ストリームがリライトされます。

並行処理を有効にする：マルチ プロセッサのコンピューターで実行する場合、このオプションでシステムがプロセッサ間の負荷のバランスがとれるようになり、より速いパフォーマンスとなります。C5.0、レコード結合 (キーによる)、ソート、分割 (順位付け法分位法)、およびレコード集計 (1 つ以上のキー・フィールドを使用) の複数または個別のノードを使用すると、並行処理の恩恵を受けることができます。

SQL 生成：SQL 生成を有効にするには、このオプションを選択します。実行プロセスを生成する SQL コードを使用することで、ストリーム操作をデータベースにプッシュバックできるようにし、それがパフォーマンスを向上させる可能性があります。さらにパフォーマンスを向上させるには、データベースにプッシュバックする操作の数を最大限にできるように、「SQL 生成の最適化」も選択できます。ノードの操作がデータベースにプッシュバックされると、そのノードはストリームの実行時に紫色で強調表示されます。

- **データベース・キャッシング**：SQL を生成するストリームをデータベース内で実行するために、ファイル・システムでなくデータベース内の一時テーブルへ、データを中流でキャッシュできます。この機能を SQL 最適化と組み合わせると、パフォーマンスが著しく向上する可能性があります。例えば、データ・マイニング・ビューを作成するために複数のテーブルを併合するストリームからの出力をキャッシュし、必要に応じて再使用できます。データベースのキャッシングが有効化されると、任意の非ターミナル・ノードを右クリックするだけでその場所のデータがキャッシュされ、次にストリームが実行されると自動的に、データベース内にキャッシュが直接作成されます。これにより SQL が下流のノード用に生成されるようになり、パフォーマンスをさらに改善します。代わりに、例えばポリシーまたは許可によってデータベースに書き込まれたデータが損なわれる場合、このオプションを必要に応じて無効にすることができます。データベース・キャッシングまたは SQL 最適化が有効にされていないと、キャッシュは、代わりにファイル・システムへ書き出されます。詳しくは、トピック 39 ページの『ノードのキャッシュ・オプション』を参照してください。
- **緩和された変換を使用**：単一の形式で保存されている場合、文字列から数値、または数値から文字列へのデータの変換を可能にします。例えば、データが文字列としてデータベースに保存されているが、実際に意味のある数値が含まれている場合、プッシュバックが発生したときに使用するようデータを変換することができます。

注：SQL 実装の些細な違いが原因で、データベース内で実行されるストリームは、IBM SPSS Modeler 内で実行されるときと若干異なる結果を返すことがあります。これらの違いは、似たような理由で、データベース・ベンダーによっても異なる可能性があります。

「デフォルトとして保存」。指定したオプションは、現在のストリームのみにも適用されます。このボタンをクリックすると、これらのオプションをすべてのストリームのデフォルトとして設定します。

ストリームの **SQL** ロギングおよびレコード ステータスオプションの設定：これらの設定には、ストリームによって生成される SQL 文の表示、およびストリームで処理されたレコード数の表示を制御するさまざまなオプションが含まれています。

ストリーム実行中に **SQL** を メッセージ・ログに表示する：ストリームの実行中に、生成された SQL をメッセージ・ログに渡すかどうかを指定します。

ストリームの準備中に、**SQL** 生成の詳細をメッセージ ログに表示する：ストリームのプレビュー中に、生成される SQL のプレビューをメッセージログに渡すかどうかを指定します。

SQL の表示：ログに出力する SQL に、ネイティブの SQL 関数を含めるか、SPSS Modeler によって生成される {fn FUNC(...)} 形式の標準 ODBC 関数を含めるかを指定します。前者は、実装されていない可能性がある ODBC ドライバーの機能に依存しています。

SQL を再フォーマットして読みやすくする : ログに表示された **SQL** を読みやすくするためにフォーマットするかどうかを指定します。

レコードのステータスを表示 : レコードがターミナル・ノードに達したときに、そのレコードを報告する時期を指定します。N レコードごとにステータスの更新を行うことを、数字で指定します。

「デフォルトとして保存」。指定したオプションは、現在のストリームのみ適用されます。このボタンをクリックすると、これらのオプションをすべてのストリームのデフォルトとして設定します。

ストリームのレイアウト・オプションの設定: これらの設定は、ストリーム領域の表示と使用方法に関連する多くのオプションを提供します。

ストリーム領域の最小幅 : ストリーム領域の最小幅をピクセル単位で指定します。

ストリーム領域の最小高 : ストリーム領域の最小の高さをピクセル単位で指定します。

ストリームのスクロール率: ストリーム領域のスクロール率を指定して、ノードが領域上のある場所から別の場所へドラッグされる場合に、ストリーム領域ウィンドウのスクロールの速度を制御します。大きい数字を指定すると、スクロール率が高速になります。

アイコン名の文字数 : ストリーム領域上のノード名の文字数制限を指定します。

アイコン サイズ : ストリーム ビュー全体を標準アイコンサイズの 8% ~ 200% のいずれかのサイズに設定します。

グリッド セル サイズ : リストからグリッド・セル・サイズを選択します。ここに指定した数値は、ストリーム領域上のノードを、目に見えないグリッドを使用して配置するために用いられます。デフォルトは 0.25 です。

グリッドに合わせる : アイコンを目に見えないグリッド・パターンに沿って並べる場合に選択します (デフォルト)。

生成されたアイコンの位置 : モデル ナゲットから生成するノードのアイコンを配置するキャンバスの場所を選択します。デフォルトは左上です。

「デフォルトとして保存」。指定したオプションは、現在のストリームのみ適用されます。このボタンをクリックすると、これらのオプションをすべてのストリームのデフォルトとして設定します。

Analytic Server のストリームのプロパティ: これらの設定は、**Analytic Server** を操作するための多くのオプションを提供します。

Analytic Server の外部で処理するレコードの最大数

Analytic Server のデータ・ソースから **SPSS Modeler Server** にインポートするレコードの最大数を指定します。

Analytic Server でノードを処理できない場合の通知

この設定により、**Analytic Server** に送信されるストリームに **Analytic Server** で処理できないノードが含まれている場合の動作が決まります。警告を出してストリームの処理を続行するか、またはエラーをスローして処理を停止するかを指定します。

分割モデル・ストレージ設定

モデルが次のサイズ (MB) を超過した場合に参照により **Analytic Server** 内に分割モデルを保管 (**Store split models by reference on Analytic Server when model size (MB) exceeds**)

通常、モデル ナゲットはストリームの一部として保管されます。多くの分割を含む分割モデルでは大きなナゲットが生成されることがあり、そのナゲットをストリームと **Analytic Server** の間で往復させるとパフォーマンスに影響を及ぼす可能性があります。解決策として、分割モデルが指定のサイズを超過した場合はそのモデルを **Analytic Server** に保管し、**SPSS Modeler** 内のナゲットにそのモデルへの参照を含めます。

実行の完了後、参照によりモデルを保管する **Analytic Server** 内のデフォルトのフォルダー 分割モデルを保管する **Analytic Server** 内の場所のデフォルト・パスを指定します。このパスは、有効な **Analytic Server** のプロジェクト名で開始する必要があります。

プロモートされたモデルを保管するフォルダー

「プロモートされた」モデルを保管する場所のデフォルト・パスを指定します。プロモートされたモデルは、**SPSS Modeler** セッションの終了時にクリーンアップされません。

ストリームの地理空間オプションの設定: すべての地理空間フィールドは、それが形状、座標、単一の軸の値 (x や y、または緯度と経度) のどれであろうとも、座標系が関連付けられています。この座標系は、起点 (0,0) や、値に関連付けられている単位などの属性を設定します。

座標系は複数あり、地理および投影の 2 つのタイプがあります。 **SPSS Modeler** のすべての空間処理関数は、投影座標系でのみ使用できます。

座標系の性質により、2 つの別個の地理空間データ ソースのデータを結合または追加するには、それらのソースが同じ座標系を使用していることが求められます。このため、ストリームで使用されている地理空間データの座標設定を指定する必要があります。

以下の状況では、選択したストリーム座標系を使用するためにデータが自動的に再投影されます。

- 空間処理関数 (area、closeto、within など) では、関数に渡されたパラメータが自動的に再投影されます。ただし、元の生データは変更されません。
- 時空間予測 (STP) でビルド ノードまたはスコアリング (ナゲット) ノードのいずれかを使用すると、場所フィールドが自動的に再投影されます。スコアリング時に、ナゲットから得られる場所が元の場所です。
- マップ視覚化ノードの使用時。

「ストリーム座標系」。このチェック ボックスを選択している場合にのみ使用可能です。使用可能な投影座標系のリストを表示し、現行ストリームに使用する座標系を選択するには、「変更」をクリックします。

「デフォルトとして保存」。選択した座標系は、現行ストリームにのみ適用されます。座標系をすべてのストリームのデフォルトとして選択するには、このボタンをクリックします。

地理空間座標系の選択: **SPSS Modeler** のすべての空間処理関数は投影座標系でのみ使用できます。

「ストリームの座標系の選択 (Select Stream Coordinate System)」ダイアログ ボックスには、ストリームで使用される地理空間データ用に選択できるすべての投影座標系のリストが含まれています。

各座標系について、次の情報がリストされます。

- **WKID** 各座標系に固有のウェルノウン ID。
- 名前 座標系の名前。
- ユニット 座標系に関連する測定の単位。

すべての座標系のリストに加え、このダイアログ ボックスには「フィルタリング」コントロールがあります。必要な座標系の名前の全体または一部がわかっている場合は、このダイアログ ボックスの下部にある「名前」フィールドにそれを入力します。選択対象の座標系のリストが自動的にフィルタリングされ、入力したテキストを含む名前を持つシステムのみが表示されるようになります。

ストリーム操作に関するメッセージの表示

ストリームの実行、最適化、およびモデル構築および評価の時間などのストリーム操作に関するメッセージは、「ストリームのプロパティ」ダイアログ・ボックスの「メッセージ」タブから簡単に参照することができます。このテーブルには、エラー・メッセージも報告されます。

ストリーム メッセージを表示するには

1. 「ファイル」メニューから、「ストリームのプロパティ」を選択します (またはマネージャー・パネルの「ストリーム」タブからストリームを選択し、ポップアップ・メニューを右クリックして「ストリームのプロパティ」をクリックします)。
2. 「メッセージ」タブをクリックします。

また、「ツール」メニューから次の各項目を選択することもできます。

「ストリームのプロパティ」 > 「メッセージ」

ここには、ストリーム操作に関するメッセージだけでなく、エラー・メッセージも報告されます。ストリームの実行がエラーにより中断された場合、このダイアログ・ボックスの「メッセージ」タブにそのエラー・メッセージが表示されます。また、ストリーム領域上でエラーの発生したノードは赤で強調表示されます。

「ユーザー オプション」ダイアログ・ボックスで、SQL 最適化とログ記録オプションが有効な場合、生成された SQL についての情報も表示されます。詳しくは、トピック 45 ページの『ストリームの最適化オプションの設定』を参照してください。

ここに表示されたメッセージを保存するには、ドロップダウン・リスト (「メッセージ」タブのすぐ下、左側) から、「メッセージの保存」をクリックします。

あるストリームのメッセージをすべて消去するには、ドロップダウンから「すべてのメッセージを消去」を選択します。

CPU 時間は、サーバー プロセスで CPU が使用されている時間であることに注意してください。経過時間は実行開始から実行終了までの合計時間であるため、ファイルの転送や出力の表示などの時間も含まれます。1 つのストリームで複数の CPU が活用されている場合 (並列実行)、CPU 時間が経過時間を超える場合があります。データ ソースとして使用されているデータベースにストリームが完全にプッシュ バックされると、CPU 時間はゼロになります。

ノード実行時間の表示

「メッセージ」タブで、実行時間を表示するように選択することができます。その場合、ストリーム内の、IBM SPSS Modeler Server で実行されたすべてのノードの個別の実行時間を表示することができます。他の領域 (R または Analytic Server) で実行されたストリームについては、実行時間が正確でない場合があります。

注: この機能を使用するには、「オプション」タブの「全般」設定の「実行時間を表示」チェック ボックスをオンにします。

ノード実行時間の表では、列は次のようになります。列見出しをクリックし、エントリーを昇順または降順にソートします (例えば、最も実行時間の長いノードを表示)。

ターミナル・ノード: ノードが属するブランチの識別子。識別子は、ブランチの末端のターミナル・ノードの名前です。

ノード・ラベル: 実行時間が参照するノードの名前。

ノード ID: 実行時間が参照するノードの一意の識別子。この識別子は、ノードが作成されるシステムによって生成されます。

実行時間: このノードの実行にかかる時間 (秒)。

ストリームとセッション・パラメーターの設定

パラメーターは、CLEM 式とスクリプトで使用するために定義できます。実際のところ、パラメーターはユーザー定義の変数であり、保存されて、現在のストリーム、セッション、またはスーパーノードで継続します。さらに、スクリプトを使用する場合と同様に、ユーザー・インターフェースからもアクセスできます。例えば、ストリームを保存すると、そのストリームに設定されているパラメーターも保存されます (これは、ローカル・スクリプト変数と異なる点です。ローカル・スクリプト変数は、宣言されたスクリプト内でのみ使用できます)。多くの場合、パラメーターはスクリプトで使用され、スクリプト内でハードコーディングする必要のないフィールドや値の情報を提供することにより、スクリプトの動作を制御します。

パラメーターの有効範囲は、それがどこで設定されたかによって異なります。

- ストリーム・パラメーターは、ストリーム・スクリプト内またはストリーム・プロパティのダイアログ・ボックス内で設定でき、ストリーム内のすべてのノードで使用できます。Clem 式ビルダーの「パラメーター」リストに表示されます。
- セッション・パラメーターは、スタンドアロン・スクリプト内または「セッション・パラメーター」ダイアログ・ボックス内で設定できます。セッション・パラメーターは、現在のセッションのすべてのストリーム (「マネージャ」ウィンドウの「ストリーム」タブに表示されているすべてのストリーム) で利用できます。

パラメーターは、スーパーノード用にも設定できます。この場合、スーパーノード内にカプセル化されたノードでだけ表示できます。

ユーザー・インターフェースからストリームとセッションのパラメーターを設定するには

1. ストリーム・パラメーターを設定するには、次のメニュー項目を選択してください。

「ツール」 > 「ストリームのプロパティ」 > 「パラメーター」

2. セッション・パラメーターを設定するには、「ツール」メニューの「セッション パラメーターの設定」を選択します。

プロンプト?: ランタイムにこのパラメーターに値を入力するよう、プロンプトを表示したい場合は、このボックスをオンにします。

「名前」。ここにはパラメーター名が表示されます。新しくパラメーターを作成するには、このフィールドに名前を入力します。例えば、最低気温を表すパラメーターを作成する場合に、minvalue と入力することができます。CLEM 式内でパラメーターを示す接頭辞の \$P- を付けないようにしてください。ここで指定した名前は、CLEM 式ビルダーにも表示されます。

ロング ネーム : 作成したパラメーターを説明する名前が表示されます。

ストレージ。 リストからストレージ・タイプを選択します。ストレージで、データ値がパラメーター内のように格納されるかを示します。例えば、「008」のように先頭に 0 がある値を扱う場合に、その 0 を保持する必要があるならば、ストレージ・タイプとして「文字列」を選択する必要があります。選択しないと、値から 0 が除去されます。ストレージ・タイプとしては、文字列、整数、実数、時間、日付、またはタイムスタンプを利用できます。日付のパラメーターには、次の段落で示す ISO 規格の表記を使用して値を指定する必要があります。

値。 各パラメーターの現在の値が表示されます。必要に応じてパラメーターを調整してください。日付のパラメーターには、ISO 規格の表記（つまり、YYYY-MM-DD）を使用して値を指定する必要があります。他の形式で指定された日付は受け入れられません。

データ型 (オプション)： ストリームを外部アプリケーションに展開する場合は、使用する測定の尺度をリストから選択します。それ以外の場合は、データ型の欄はそのままにしておくことをお勧めします。数値範囲の上限および下限など、パラメーターに値の制約を指定したい場合、リストから「指定」を選択します。

ロング ネーム、ストレージ、およびデータ型のオプションは、ユーザー・インターフェースを通じてだけ、パラメーターに設定できます。これらのオプションは、スクリプトを使用して設定できません。

右にある矢印をクリックして、選択したパラメーターを使用可能なパラメーターのリストの上または下に移動することができます。選択したパラメーターを削除するには、削除ボタン (X マーク) を使用します。

パラメーター値のライタイム・プロンプトの指定

さまざまな機会に、同じパラメーターに異なる値を入力する必要があるストリームがある場合、1 つまたは複数のストリームまたはセッション・パラメーター値にランタイム・プロンプトを指定できます。

パラメーター: (オプション) パラメーターの値を入力、またはデフォルト値がある場合はデフォルト値のままにします。

これらのプロンプトの無効化 ストリーム実行時にこれらのプロンプトを表示しない場合は、このボックスをオンにします。パラメーターが定義されているストリーム・プロパティまたはセッション・プロパティのダイアログ・ボックスの「プロンプト?」チェック・ボックスをオンにすると、プロンプトを再度表示することができます。詳しくは、トピック 51 ページの『ストリームとセッション・パラメーターの設定』を参照してください。

パラメーター・タイプの値の制約の指定

データ・モデリング・ストリームを読み込む外部アプリケーションにストリームを展開する際に、パラメーターの値の制約を行うことができます。このダイアログ・ボックスでは、ストリームを実行する外部ユーザーが利用できる値を指定することができます。データ型によって、ダイアログ・ボックスで指定できる値の範囲は動的に変わります。ここに表示されるオプションは、データ型ノードからの値に対して利用できるオプションと同じものです。

データ型: 現在選択されている測定の尺度を表示します。IBM SPSS Modeler でのパラメーターの利用目的に応じて、値を変更することができます。

ストレージ。 わかっている場合に、ストレージ・タイプを表示します。ストレージ・タイプは、IBM SPSS Modeler における作業で選択する測定の尺度 (連続型、名義型、フラグ型) には影響されません。ストレージ・タイプは、「パラメーター」タブで変更することができます。

このダイアログ・ボックスの下半分は、「データ型」で選択した測定の尺度によって異なります。

連続型測定 of 尺度

下限 :パラメーター値の下限を指定します。

上限 :パラメーター値の上限を指定します。

ラベル: 集計範囲フィールドの任意の値のラベルを指定できます。ラベル ボタンをクリックして、値ラベルを指定する個別のダイアログボックスを開きます。

名義型測定 of 尺度

値。名義型フィールドとして使用されるパラメーターの値を指定することができます。この値は、IBM SPSS Modeler ストリーム中では強制されませんが、外部展開アプリケーションのドロップダウン・リストで使用されます。矢印および削除ボタンを使用して、既存の値を変更したり、値の並び替えや削除などの作業を行えます。

フラグ型測定 of 尺度

真 (**true**) :条件を満たす場合のパラメーターのフラグ値を指定します。

偽 (**false**) :条件を満たさない場合のパラメーターのフラグ値を指定します。

ラベル: フラグ型フィールドの値のラベルを指定できます。

ストリーム展開のオプション

「ストリームのプロパティ」ダイアログ・ボックスの「展開」タブで、モデル・リフレッシュ、自動ジョブ・スケジュール、または IBM Analytical Decision Management のさらなる活用を目的に、IBM SPSS Collaboration and Deployment Services 内のストリームを展開するためのオプションを指定できます。すべてのストリームは、スコアリング・ブランチを指定しないと、展開できません。詳しくは、トピック 164 ページの『リポジトリ・オブジェクトの保存と展開』を参照してください。

ストリームのループ実行

「ストリームのプロパティ」ダイアログ・ボックスの「実行」タブを使用すると、ループ条件を設定して、現在のストリームの繰り返しタスクを自動化することができます。

これらの条件を設定すると、それをスクリプトへの導入として使用できます。これは、より良いスクリプトを作成するためのベースとして使用する目的などで変更可能なストリームの基本スクリプトがスクリプト・ウィンドウに取り込まれるためです。詳しくは、トピック 158 ページの『グローバル関数』を参照してください。

ストリームのループの設定

1. 「ファイル」メニューから、「ストリームのプロパティ」を選択します (またはマネージャー・パネルの「ストリーム」タブからストリームを選択し、ポップアップ・メニューを右クリックして「ストリームのプロパティ」をクリックします)。
2. 「実行」タブをクリックします。
3. 「ループ / 条件式実行」実行モードを選択します。
4. 「ループ」タブをクリックします。

また、「ツール」メニューから次の各項目を選択することもできます。

「ストリームのプロパティ」 > 「実行」

さらに別の方法としては、ノードを右クリックして、コンテキスト・メニューから以下をクリックします。

「ループ / 条件式実行」 > 「ループ設定の編集」

反復。この行番号の値は編集できませんが、テーブルの右側にあるボタンを使用することで反復の追加や削除を行ったり、反復を上下に移動させたりすることができます。

テーブル・ヘッダー。テーブル・ヘッダーには、反復キーと、ループの設定時に作成したすべての反復変数が反映されます。

ストリームのグローバル値の表示

「ストリームのプロパティ」ダイアログ・ボックスの「グローバル」タブでは、現在のストリームに設定されているグローバル値を表示することができます。グローバル値は、グローバル・ノードを使用して作成され、選択したフィールドの平均、合計、または標準偏差などの統計量を判断するために用いられます。

グローバル・ノードが実行されると、ストリーム操作内でこれらの値をさまざまに利用することができます。詳しくは、トピック 158 ページの『グローバル関数』を参照してください。

ストリームのグローバル値を表示するには

1. 「ファイル」メニューから、「ストリームのプロパティ」を選択します (またはマネージャー・パネルの「ストリーム」タブからストリームを選択し、ポップアップ・メニューを右クリックして「ストリームのプロパティ」をクリックします)。
2. 「グローバル」タブをクリックします。

また、「ツール」メニューから次の各項目を選択することもできます。

「ストリームのプロパティ」 > 「グローバル」

利用できるグローバル値 : このテーブルには、利用できるグローバル値が表示されています。ここでグローバル値を編集することはできません。ただし、テーブルの右にある「すべての値の消去」ボタンを使用して、ストリームのグローバル値をすべて消去することができます。

ストリーム内のノード検索

ノード名、カテゴリ、識別子などの検索基準を指定してストリーム内のノードを検索できます。この機能は、多くのノードを含む複雑なストリームに役立ちます。

ストリーム内のノードを検索するには

1. 「ファイル」メニューから、「ストリームのプロパティ」を選択します (またはマネージャー・パネルの「ストリーム」タブからストリームを選択し、ポップアップ・メニューを右クリックして「ストリームのプロパティ」をクリックします)。
2. 「検索」タブをクリックします。

また、「ツール」メニューから次の各項目を選択することもできます。

「ストリームのプロパティ」 > 「検索」

その他のオプションを除外してノード ID によって検索する (「次の ID に一致」フィールドを使用) という場合を除き、複数のオプションを指定して検索を絞り込みます。

以下のノード・ラベルを含む：このボックスをチェックして、ノード・ラベルのすべてまたは一部を入力して特定のノードを検索します。検索では大文字と小文字の区別を行わず、複数の単語をテキストの単一部分として扱います。

ノード カテゴリ：このボックスをチェックして、特定の種類のノードを検索するリストのカテゴリをクリックします。「プロセス ノード」は、ノード・パレットの「レコード設定」タブまたは「フィールド設定 タブ」のノードを指し、「モデル ノードの適用」は、モデル ナゲットを指します。

次のキーワードを含む：このボックスをチェックして、1 つまたは複数のキーワードを入力して、ノード・ダイアログ・ボックスの「注釈」タブの「キーワード」フィールドに入力されたテキストを含むノードを検索します。入力したキーワードは完全に一致する必要があります。あるいは、複数のキーワードで検索する場合、キーワードをセミコロンで区切ります (例えば、proton;neutron と入力すると、これらのキーワードのいずれかを含むすべてのノードが検出されます)。詳しくは、トピック 63 ページの『注釈』を参照してください。

次の注釈を含む：このボックスをチェックして、1 つまたは複数のキーワードを入力して、ノード・ダイアログ・ボックスの「注釈」タブのメイン・テキスト領域にこのテキストを含むノードを検索します。検索では大文字と小文字の区別を行わず、複数の単語をテキストの単一部分として扱います。詳しくは、トピック 63 ページの『注釈』を参照してください。

生成するフィールド名：このボックスをチェックして、生成するフィールドの名前を入力します (例：\$C-Drug)。このオプションを使用して、特定のフィールド名を生成するモデル作成ノードを検索できます。フィールド名は 1 つだけ入力してください。完全に一致する必要があります。

次の ID と一致：このボックスをチェックしてノード ID を入力し、この ID を持つ特定のノードを検索します (このオプションを選択すると、前述のオプションはすべて無効になります。)。ノード ID はノード作成時にシステムによって割り当てられ、ID を使用してスクリプトまたは自動化の目的でノードを参照できます。ノード ID は 1 つだけ入力してください。完全に一致する必要があります。詳しくは、トピック 63 ページの『注釈』を参照してください。

スーパーノードで検索：このボックスはデフォルトで選択されています。そのため、スーパーノード内外のノードに検索が実行されます。ストリームの最上位で、スーパーノード外のノードのみを検索する場合は、このボックスの選択を解除します。

検索：使用するすべてのオプションを指定し、このボタンをクリックすると検索が開始されます。

指定されたオプションに一致するノードは、ダイアログ・ボックスの下部に表示されます。リスト内のノードを選択すると、ストリーム領域内で強調表示されます。

ストリーム名の変更

「ストリームのプロパティ」ダイアログ・ボックスの「注釈」タブでは、ストリームを説明する注釈を追加したり、ストリームの名前を作成することができます。これらのオプションは特に、プロジェクト・ペインに追加されたストリームのレポートを生成する際に役立ちます。詳しくは、トピック 63 ページの『注釈』を参照してください。

ストリームの説明

作成する各ストリームについて、IBM SPSS Modeler は、ストリームの内容に関する情報を含むストリームの説明を作成します。ストリームが何を行っているかを確認したいが IBM SPSS Modeler をインストールしていない場合、例えば IBM SPSS Collaboration and Deployment Services を使用してストリームにアクセスする場合に役に立ちます。

ストリームの説明が、さまざまなセクションで構成されている HTML 文書の形式で表示されます。

ストリームの一般情報

ここでは、ストリーム名、いつストリームが作成されたか、最後に保存されたのはいつかについての詳細が表示されます。

説明とコメント

このセクションには次が表示されます。

- ストリームの注釈 (63 ページの『注釈』を参照)
- 特定のノードに接続していないコメント
- ストリームのモデル作成ブランチおよびスコアリングブランチのノードに接続するコメント

スコアリング情報

ここでは、ストリームのスコアリングブランチに関連するさまざまな見出しの情報が表示されます。

- コメント。スコアリングブランチのノードにのみリンクするコメントが表示されます。
- 入力: ストレージ タイプ (**string**、**integer**、**real** など) とともに、入力フィールドが表示されます。
- 出力: ストレージ・タイプとともに、モデル作成ノードで生成された追加フィールドなど、出力フィールドが表示されます。
- パラメーター: ストリームのスコアリングブランチに関連するパラメーターを表示し、モデルをスコアリングするごとに表示または編集できます。「ストリームのプロパティ」ダイアログで「展開」タブの「スコアリング パラメーター」 ボタンをクリックすると、これらのパラメーターが識別されます。
- モデル ノード: モデル名と種類を表示します (ニューラル・ネットワーク、C&R Tree など)。「ストリームのプロパティ」ダイアログ・ボックスで「展開」タブの「モデル ノード」 フィールドに選択されたモデル ナゲットです。
- モデルの詳細。前の見出しで識別されたモデル ナゲットの詳細を表示します。可能な場合、モデルの予測値の重要度および評価グラフが表示されます。

モデル作成情報

ストリームのモデル作成ブランチに関連する情報が表示されます。

- コメント。モデル作成ブランチのノードに接続するコメントまたは注釈がリストされます。
- 入力: モデル作成枝の役割とともに入力フィールドが表示されます (たとえば入力、出力、分割など、フィールド方向の値の形式)。
- パラメーター: ストリームのモデル作成ブランチに関連するパラメーターを表示し、モデルを更新するごとに表示または編集できます。「ストリームのプロパティ」ダイアログ・ボックスで「展開」タブの「モデル構築のパラメーター」 ボタンをクリックすると、これらのパラメーターが識別されます。
- モデル作成ノード: モデルを生成または更新するために使用されるモデル作成ノードの名前と種類を表示します。

ストリームの説明のプレビュー

「ストリームのプロパティ」ダイアログ・ボックスでオプションをクリックして、Web ブラウザーでストリームの説明の内容を表示することができます。説明の内容は、ダイアログ・ボックスの「展開」タブで指定するオプションによって異なります。詳しくは、トピック 177 ページの『ストリーム展開のオプション』を参照してください。

ストリームの説明を表示するには

1. IBM SPSS Modeler のメイン・メニューで次の各項目をクリックします。

「ツール」 > 「ストリームのプロパティ」 > 「展開」

2. 展開タイプ、指定のスコアリング・ノードおよびスコアリング・パラメーターを設定します。
3. 展開タイプが「モデルのリフレッシュ」の場合、オプションで次を選択できます。
 - モデル作成ノードとモデル構築パラメーター
 - ストリームのスコアリングブランチのモデル ナゲット
4. 「ストリーム記述のプレビュー」 ボタンをクリックします。

ストリームの説明のエクスポート

ストリームの説明の内容を HTML ファイルにエクスポートできます。

ストリームの説明をエクスポートするには

1. メイン・メニューで次の各項目をクリックします。

「ファイル」 > 「ストリームの記述のエクスポート」

2. HTML ファイルの名前を入力して、「保存」をクリックします。

ストリームの実行

ストリームに必須オプションを指定し、必要なノードを接続したら、それらのノードを使用してストリーム内でデータを実行することによって、ストリームを実行できます。IBM SPSS Modeler でストリームを実行するには、さまざまな方法があります。以下を行うことができます。

- 「ツール」メニューから「実行」を選択します。
- ツールバーの「実行...」 ボタンをクリックします。これらのボタンを利用すれば、ストリーム全体を実行することも、単に選択したターミナル・ノードだけを実行することもできます。詳しくは、トピック 19 ページの『IBM SPSS Modeler ツールバー』を参照してください。
- ターミナル・ノードを右クリックし、ポップアップ・メニューから「実行」をクリックして、1 つのデータ・ストリームを実行します。
- 非ターミナル・ノードを右クリックし、ポップアップ・メニューから「ここから実行」をクリックして、1 つのデータ・ストリームの一部を実行します。これにより、選択したノードの後のこれらの操作のみを実行できます。

実行中のストリームを中止するには、ツールバーにある赤い中止ボタンをクリックするか、または「ツール」メニューの「実行を中止」をクリックします。

何らかのストリームの実行が 3 秒以上かかる場合は、「実行のフィードバック」ダイアログ・ボックスが表示され、進行状況を示します。

一部のノードでは、ストリーム実行に関して追加情報を提供するためにさらなる表示を行うことがあります。これらの表示は、ダイアログ・ボックスの対応する行を選択することで表示されます。最初の行は自動的に選択されます。

モデルの処理

ストリームにモデル作成ノード (ノード・パレットの「モデル作成」タブまたは「DB モデリング」タブのモデル作成ノード) が含まれている場合、ストリーム実行時にモデル ナゲットが作成されます。モデル ナゲットは、モデルのコンテナです。つまり、入力データに対して予測を生成できる一連のルール、式または方程式で、予測分析の中心となるものです。



図 12. モデル ナゲット

モデル作成ノードを正常に実行すると、対応するモデル ナゲットがストリーム領域に置かれ、金色のダイヤモンドの形のアイコンで表示されます (名前は「ナゲット」)。ナゲットを開いて内容を参照し、モデルに関する詳細を表示できます。予測を表示するには、1 つまたは複数のターミナル・ノードを接続および実行、出力は読み取り可能な形式で予測を表示します。

一般的なモデル作成ストリームは、2 つのブランチで構成されています。モデル作成ブランチには、モデル作成ノード、そしてそれに先行する入力ノードおよび処理ノードが含まれています。スコアリングブランチは、モデル作成ノードを実行すると作成され、予測の表示に使用するモデル ナゲットおよびターミナル・ノードが含まれています。

詳細については、『IBM SPSS Modeler モデル作成ノード ガイド』を参照してください。

ノードおよびストリームへのコメントおよび注釈の追加

ストリームについて、組織内の他の部門に説明が必要な場合があります。説明を行うために、ストリーム、ノード、モデル ナゲットに説明のコメントを追加することができます。

他のユーザーがコメントを画面上で表示したり、コメントを含むストリームの画像を印刷することができます。

ストリームまたはスーパーノードのすべてのコメントを一覧表示し、リスト内のコメントの順序を変更、コメント テキストを編集、コメントの前景または背景の色を変更できます。詳しくは、トピック 62 ページの『ストリームのコメントの一覧表示』を参照してください。

「ストリームのプロパティ」ダイアログ・ボックス、ノードのダイアログ・ボックス、またはモデル ナゲットのウィンドウの「注釈」タブを使用して、説明をテキスト注釈の形式でストリーム、ノード、ナゲットに追加することもできます。これらの説明は、「注釈」タブが開いているときのみ表示できますが、ストリームの注釈を画面上のコメントとして表示することもできます。詳しくは、トピック 63 ページの『注釈』を参照してください。

コメント






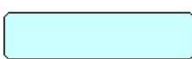
コメントはテキスト・ボックスの形式で、任意の文字数のテキストを入力でき、必要に応じてコメントを追加できます。コメントは独立 (ストリーム・オブジェクトに接続していない) していたり、ストリーム内のノードまたはモデル ナゲットに接続している場合があります。独立しているコメントは、通常ストリーム

全体の目的を説明するために使用され、接続しているコメントは、接続しているノードまたはナゲットを説明します。ノードおよびナゲットは複数のコメントを接続し、ストリームは独立したコメントをいくつも持つことができます。

注：ストリームの注釈を画面上のコメントとして表示することもできますが、これらをノードまたはナゲットに接続することはできません。詳しくは、トピック 62 ページの『注釈のコメントへの変換』を参照してください。

テキスト・ボックスの外観は、次の表が示すように、コメント（または画面上のコメントとして表示される注釈）の現在のモードを示すよう変更されます。

表 3. コメントおよび注釈のテキスト・ボックスのモード

「コメント」テキスト・ボックス	「注釈」テキスト・ボックス	モード	指定内容	取得方法
		編集	コメントが開き、編集できます。	新しいコメントまたは注釈を作成するか、既存のコメントをダブルクリックします。
		最後の選択	コメントを移動、サイズ変更または削除できます。	編集後、ストリームの背景をクリックするか、既存のコメントまたは注釈をシングルクリックします。
		ビュー	編集が完了しています。	編集後、別のノード、コメントまたは注釈をクリックします。

新しい独立したコメントを作成すると、コメントは最初ストリーム領域の左上の隅に表示されます。

コメントをノードまたはナゲットに接続している場合、コメントは最初接続されているストリーム・オブジェクトの上に表示されます。

テキスト・ボックスは白色の場合が、テキストが入力できることを示します。テキストを入力した場合、テキスト・ボックスの外側をクリックします。コメントの背景が黄色に変化し、テキスト入力完了を示します。コメントは選択されたままになっていて、移動、サイズ変更、または削除できます。

もう一度クリックすると、境界線が実線に変わり、編集が完了していることを示します。

コメントをダブルクリックすると、テキスト・ボックスの背景色が白に変わって編集モードとなり、コメントテキストを編集できます。

スーパーノードにコメントを追加することもできます。

コメントに関する操作： コメントに多くの操作を行うことができます。以下を行うことができます。

- 独立したコメントの追加
- コメントのノードまたはナゲットへの接続
- コメントの編集
- コメントのサイズ変更
- コメントの移動
- コメントの接続解除

- コメントの削除
- ストリームのすべてのコメントの表示または非表示

独立したコメントを追加するには

1. ストリームに何も選択されていないことを確認します。
2. 次のいずれかを実行します。
 - メイン・メニューで次の各項目をクリックします。

「挿入」 > 「新規コメント」

- ストリームの背景を右クリックし、ポップアップ・メニューで「新規コメント」をクリックします。
 - ツールバーにある「新規コメント」ボタンをクリックします。
3. コメント テキストを入力 (またはクリップ ボードからテキストを貼り付け) ます。
 4. ストリームのノードをクリックして、コメントを保存します。

コメントをノードまたはナゲットに接続するには

1. ストリーム領域で 1 つまたは複数のノードまたはナゲットを選択します。
2. 次のいずれかを実行します。
 - メイン・メニューで次の各項目をクリックします。

「挿入」 > 「新規コメント」

- ストリームの背景を右クリックし、ポップアップ・メニューで「新規コメント」をクリックします。
 - ツールバーにある「新規コメント」ボタンをクリックします。
3. コメント テキストを入力します。
 4. ストリームの別のノードをクリックして、コメントを保存します。

または、次を行うことができます。

5. 独立したコメントの挿入 (前項を参照)
6. 次のいずれかを実行します。
 - コメントを選択して F2 キーを押し、ノードまたはナゲットを選択します。
 - ノードまたはナゲットを選択して F2 キーを押し、コメントを選択します。
 - (3 ボタンのマウスのみ) マウス・ポインタをコメントに移動し、中央のボタンを押したままマウス・ポインタをノードまたはナゲットに移動して、マウス・ボタンを離します。

コメントを追加のノードまたはナゲットに接続するには

コメントがすでにノードまたはナゲットに接続されている、または現在ストリーム・レベルにあり、追加のノードまたはナゲットにコメントを接続する場合は、次のいずれかを実行します。

- コメントを選択して F2 キーを押し、ノードまたはナゲットを選択します。
- ノードまたはナゲットを選択して F2 キーを押し、コメントを選択します。
- (3 ボタンのマウスのみ) マウス・ポインタをコメントに移動し、中央のボタンを押したままマウス・ポインタをノードまたはナゲットに移動して、マウス・ボタンを離します。

既存のコメントを編集するには

1. 次のいずれかを実行します。
 - コメント テキスト・ボックスをダブルクリックします。
 - テキスト・ボックスを選択して Enter キーを押します。
 - テキスト・ボックスを右クリックして、メニューを表示し、「編集」をクリックします。
2. コメント テキストを編集します。編集時、Ctrl + C キーを使用してテキストをコピーするなど、標準的な Windows ショートカット・キーを使用できます。編集中にそのほかのオプションが、コメントのポップアップ・メニューに表示されます。
3. テキスト・ボックスの外部を 1 回クリックするとサイズ変更のコントロールが表示され、もう 1 回クリックするとコメントが完了します。

コメント テキスト・ボックスをサイズ変更するには

1. コメントを選択してサイズ変更のコントロールを表示します。
2. コントロールをクリック・アンド・ドラッグしてボックスのサイズを変更します。
3. テキスト・ボックスの外部をクリックして、変更を保存します。

既存のコメントを移動するには

接続したオブジェクト (あれば) を移動せずにコメントを移動する場合、次のいずれかを実行します。

- コメントにマウス・ポインタを移動し、左のマウス・ボタンを押したままコメントを新しい位置にドラッグします。
- コメントを選択し、Alt キーを押したまま矢印キーを使用してコメントを移動します。

コメントを接続したノードまたはナゲットとともに移動する場合は、次のようにします。

1. 移動するすべてのオブジェクトを選択します。
2. 次のいずれかを実行します。
 - オブジェクトのいずれかにマウス・ポインタを移動し、左のマウス・ボタンを押したままオブジェクトを新しい位置にドラッグします。
 - オブジェクトのいずれかを選択し、Alt キーを押したまま矢印キーを使用してオブジェクトを移動します。

コメントとノードまたはナゲットの接続を解除するには

1. 接続を解除するコメントを選択します。
2. 次のいずれかを実行します。
 - F3 キーを押します。
 - 選択したコメントを右クリックして、メニューの「接続の解除」をクリックします。

コメントを削除するには

1. 削除するコメントを選択します。
2. 次のいずれかを実行します。
 - Del キーを押します。
 - 選択したコメントを右クリックして、メニューの「削除」をクリックします。

コメントがノードまたはナゲットに接続している場合、接続も同時に削除されます。

コメントが本来、独立したコメントに変換されたストリームまたはスーパーノードの注釈である場合、コメントは領域から削除されますが、テキストはストリームまたはスーパーノードの「注釈」タブに保持されます。

ストリームのコメントの表示または非表示

1. 次のいずれかを実行します。

- メイン・メニューで次の各項目をクリックします。

「表示」 > 「コメント」

- ツールバーにある「コメントの表示/非表示」 ボタンをクリックします。

ストリームのコメントの一覧表示: 特定のストリームまたはスーパーノードに作成されたすべてのコメントのリストを表示できます。

このリスト上で、次のことができます。

- コメントの順序の変更
- コメント テキストを編集
- コメントの前景色または背景色の変更

コメントの一覧表示

ストリームに作成されたコメントを一覧表示するには、次のいずれかを実行します。

- メイン・メニューで次の各項目をクリックします。

「ツール」 > 「ストリームのプロパティ」 > 「コメント」

- マネージャー・パネルでストリームを右クリックし、「ストリームのプロパティ」 を選択して「コメント」 をクリックします。
- ストリーム領域上でストリームの背景を右クリックし、「ストリームのプロパティ」 を選択して「コメント」 をクリックします。

テキスト。コメントのテキスト。テキストをダブルクリックして、フィールドを編集可能なテキスト・ボックスに変更します。

リンク。 コメントが接続されるノードの名前。このフィールドが空の場合、コメントがストリームに適用されます。

配置ボタン : リスト内の選択したコメントを上下に移動させます。

コメントの色 : コメントの前景色または背景色を変更するには、コメントを選択し、「カスタム カラー」チェック・ボックスを選択し、「背景」 または 「前景」 リスト (または両方) から色を選択します。「適用」 をクリックし、ストリームの背景をクリックして、変更の内容を表示します。「OK」 をクリックして、変更を保存します。

注釈のコメントへの変換: ストリームまたはスーパーノードに作成された注釈は、コメントに変換できません。

ストリームの場合、注釈はストリーム領域の独立したコメントに変換されます (つまり、どのノードにも接続されません)。

スーパーノードの注釈をコメントに変換する場合、コメントはストリーム領域のスーパーノードに接続されませんが、スーパーノードにズーム・インすると確認することができます。

ストリームの注釈をコメントに変換するには

1. 「ツール」メニューから「ストリームのプロパティー」を選択します。(代わりに、マネージャー・パネル内でストリームを右クリックして、「ストリームのプロパティー」を選択することもできます。)
2. 「注釈」タブをクリックします。
3. 「注釈をコメントとして表示」チェック・ボックスを選択します。
4. 「OK」をクリックします。

スーパーノードの注釈をコメントに変換するには

1. 領域の「スーパーノード」アイコンをダブルクリックします。
2. 「注釈」タブをクリックします。
3. 「注釈をコメントとして表示」チェック・ボックスを選択します。
4. 「OK」をクリックします。

注釈

ノード、ストリーム、およびモデルには、さまざまな方法で注釈を付けることができます。詳細な注釈を追加したり、新しい名前を指定することができます。これらのオプションは特に、プロジェクト・ペインに追加されたストリームのレポートを生成する際に役立ちます。ノードおよびモデル ナゲットには、ストリーム領域にある、よく似たノードと区別するのに役立つように、ツールヒント テキストを追加することもできます。

注釈の追加

ノードまたはモデル ナゲットを編集すると、ダイアログ・ボックスが表示されます。このダイアログ・ボックスの「注釈」タブでは、注釈に関するさまざまなオプションを設定できます。このタブは、次の方法で表示することもできます。

1. ノードまたはナゲットに注釈を付けるには、ストリーム領域にあるノードまたはナゲットを右クリックし、「名前の変更と注釈」をクリックします。「編集」ダイアログ・ボックスの「注釈」タブが表示されます。
2. ストリームに注釈を付けるには、「ツール」メニューから「ストリームのプロパティー」をクリックします。(代わりに、マネージャー・パネル内でストリームを右クリックして、「ストリームのプロパティー」を選択することもできます。)「注釈」タブをクリックします。

「名前」。ストリーム領域に表示されるノードについて、自動生成された名前を調整するか、固有の名前を作成するには、「カスタム」を選択します。

ツールヒント テキスト：(ノードおよびモデル ナゲットに対してのみ) ストリーム領域のツールヒントとして使用されるテキストを入力します。特に、似たようなノードが多数あるような作業に役立ちます。

キーワード：プロジェクト・レポート内で使用したりストリーム内のノードを検索する場合に、リポジトリに保存されているオブジェクトの検索または追跡時に使用したりするキーワードを指定します (163 ページの『IBM SPSS Collaboration and Deployment Services Repository について』を参照)。複数のキーワードをセミコロンで区切って指定できます income; crop type; claim value各キーワードの前後にある

空白文字は削除されます。例えば、`income ; crop type` は、`income;crop type` と同じ結果になります。(ただし、キーワード内の空白文字は削除されません。例えば、スペースが 1 個の `crop type` とスペースが 2 個の `crop type` は同一ではありません。)

メイン・テキスト領域には、ノードの操作やノード作成時の決定事項などの操作に関する注釈を記入します。例えば、ストリームを共有、再利用する場合、「空白が多いフィールドをフィルター・ノードを使用して破棄する。」のような情報を記載しておけば、大変役に立ちます。ノードに注釈を付けると、ノードに情報が保存されます。これらの注釈を、プロジェクト・パネルを使用して作成したプロジェクト・レポートに表示するように設定することもできます。詳しくは、トピック 187 ページの『プロジェクトの概要』を参照してください。

注釈をコメントとして表示：(ストリームおよびスーパーノードの注釈についてのみ) このボックスをチェックすると、注釈をストリーム領域上に表示される独立したコメントに変換します。詳しくは、トピック 58 ページの『ノードおよびストリームへのコメントおよび注釈の追加』を参照してください。

ID: スクリプトまたはオートメーションの目的でノードを参照するために使用できる一意の ID を表示します。この値はノード作成時に自動的に生成され、変更することはありません。「O」の文字との混乱を避けるため、ノード ID にはゼロを使用しません。必要に応じて、右側にあるコピー・ボタンを使用して ID をスクリプトなどの場所にコピーして貼り付けます。

データ・ストリームの保存

ストリームを作成したら、今後再利用するためにストリームを保存できます。

ストリームを保存するには、以下を実行します。

1. 「ファイル」メニューの「ストリームを上書き保存」または「名前を付けてストリームを保存」をクリックします。
2. 「保存」ダイアログ・ボックスで、ストリーム・ファイルを保存するフォルダーを指定します。
3. 「ファイル名」テキスト・ボックスにストリームの名前を入力します。
4. 保存したストリームを現在のプロジェクトに追加する場合は、「プロジェクトに追加」を選択します。

「保存」をクリックすると、指定したディレクトリーにストリームが、ファイル拡張子 `*.str` で保存されます。

自動バックアップ・ファイル：毎回、ストリームが保存されると、そのファイルの以前に保存されたバージョンは、自動的にバックアップとして保存されます。バックアップ・ファイルはのファイル名には、ハイフンが追加されます (例: `mystream.str-`)。バックアップ バージョンを復元するには、ハイフンを削除してそのファイルを再び開くだけです。

ステートの保存

ストリームのほかに、ステートを保存することもできます。ステートには、現在表示されているストリーム・ダイアグラムとモデル ナゲットが含まれています (マネージャー・パネルの「モデル」タブに記載)。

ステートを保存するには

1. 「ファイル」メニューで次の各項目をクリックします。
「ステート」 > 「ステートの保存」または「名前を付けてステートを保存」
2. 「保存」ダイアログ・ボックスで、ステートを保存するフォルダーを指定します。

「保存」 をクリックすると、指定したディレクトリーにステートが、ファイル拡張子 **.cst* で保存されます。

ノードの保存

ストリーム領域でノードを右クリックし、ポップアップ・メニューから 「ノードの保存」 をクリックして、個々のノードを保存することもできます。ファイルの拡張子は、**.nod* を使用します。

複数のストリーム・オブジェクトの保存

ストリーム、プロジェクト、またはモデル ナゲットなどの複数のオブジェクトが保存されていない状態で IBM SPSS Modeler を終了する場合、ソフトウェアを終了する前にそれらのオブジェクトを保存するよう指示するメッセージが表示されます。これらのオブジェクトを保存することを選択した場合、各オブジェクトを保存するためのオプションがダイアログ・ボックスに表示されます。

1. 保存するオブジェクトに対応するチェック・ボックスを選択してください。
2. 各オブジェクトを指定した場所に保存するには、「OK」 をクリックします。

各オブジェクトに対応する標準の「保存」ダイアログ・ボックスが表示されます。保存が完了すると、アプリケーションが終了します。

出力の保存

IBM SPSS Modeler 出力ノードから生成されたテーブル、グラフ、およびレポートは、出力オブジェクト (**.cou*) 形式で保存できます。

1. 保存したい出力を表示するときは、出力ウィンドウのメニューで、次をクリックします。

「ファイル」 > 「保存」

2. 出力ファイルの名前と場所を指定します。
3. 必要に応じて、現在のプロジェクトにファイルを追加するには、「保存」ダイアログ・ボックスにある「ファイルをプロジェクトに追加」を選択します。詳しくは、トピック 187 ページの『プロジェクトの概要』を参照してください。

また、「マネージャー」パネルにリストされている任意の出力オブジェクトを右クリックして、ポップアップ・メニューから 「保存」 を選択することもできます。

暗号化および復号化の情報

ストリーム、ノード、プロジェクト、出力ファイル、またはモデル ナゲットを保存するときは、無許可の使用を防止するために暗号化することができます。こうするには、保存するときに追加のオプションを選択して、保存するアイテムにパスワードを追加します。この暗号化は、保存するあらゆるアイテムに設定して、セキュリティーを追加できます。これは、IBM SPSS Modeler と IBM SPSS Modeler Server の間でファイルを受け渡す場合に使用する SSL 暗号とは異なります。

暗号化されたアイテムを開くときは、パスワードを入力するように促すプロンプトが表示されます。正しいパスワードを入力すると、アイテムが自動的に復号化され、通常通りに開きます。

項目を暗号化するには

1. 「保存」ダイアログ・ボックスで、暗号化するアイテムについて、「オプション」 をクリックします。「暗号化オプション」ダイアログ・ボックスが開きます。
2. 「このファイルを暗号化」 を選択します。

3. オプションで、さらにセキュリティーを高めるために、「キーワードのマスク」を選択します。こうすることで、入力したものがすべてドットの列として表示されます。
4. パスワードを入力します。警告 :パスワードを忘れた場合は、ファイルまたはモデルは開くことができなくなります。
5. 「キーワードのマスク」を選択した場合は、パスワードを再入力して、正しくパスワードを入力したことを確認します。
6. 「OK」をクリックして、「保存」ダイアログ・ボックスに戻ります。

注：暗号化保護されたアイテムのコピーを保存する場合は、「暗号化オプション」ダイアログ・ボックスで設定を変更しない限り、新しいアイテムは元のパスワードを使用して自動的に暗号化形式で保存されます。

ファイルのロード

IBM SPSS Modeler で保存された、さまざまなオブジェクトを再ロードすることができます。

- ストリーム (.str)
- ステート (.cst)
- モデル (.gm)
- モデル・パレット (.gen)
- ノード (.nod)
- 出力 (.cou)
- プロジェクト (.cpj)

新規ファイルを開く

ストリームは、「ファイル」メニューから直接ロードすることができます。

- 「ファイル」メニューの「ストリームを開く」をクリックします。

その他の種類のファイルは、「ファイル」メニューのサブメニューを使用して開くことができます。例えば、モデルをロードするには、「ファイル」メニューから次の各項目を選択します。

「モデル」 > 「モデルを開く」または「モデル・パレットをロード」

最近使用したファイルを開く

最近使用したファイルを素早くロードするには、「ファイル」メニューの下の部分にあるオプションを使用します。

「最近使用したストリーム」、「最近使用したプロジェクト」、または「最近使用したステート」を選択して、最近使用したファイルのリストを開きます。

データ・ストリームのマッピング

マッピング・ツールを使用して、新しいデータ・ソースを既存のストリームに接続できます。マッピング・ツールは接続を設定するだけでなく、既存のストリームのフィールドを新しいソースのフィールドで置換する方法を指定する場合にも役立ちます。新しいデータ・ソース用にデータ・ストリーム全体を作成しなくても、単に既存のストリームに接続するだけで済みます。

データ・マッピング・ツールを利用すれば、2つのストリーム・フラグメントを結合し、すべての(必須)フィールド名を正しく一致させることができます。基本的に、データのマッピングを行うと、名前を変更して適切なフィールドを一致させる新しいフィルター・ノードが単純に作成されます。

データのマッピングを行うには、次の2種類の方法があります。

置換ノードの選択：この方法は、置換されるノードを選択することから始まります。まず置換するノードを右クリックします。ポップアップ・メニューの「データ マッピング」>「置換ノードの選択」オプションを使用して、置換するノードを選択します。

マップ先：この方法は、ストリームに導入するノードから始まります。まず導入するノードを右クリックします。ポップアップ・メニューの「データ マッピング」>「マップ先」オプションを使用して、結合するノードを選択します。この方法は、ターミナル・ノードにマッピングする場合などに適しています。注：レコード結合ノードまたはレコード追加ノードにマップすることはできません。この場合は、普段と同じように、ストリームをレコード結合ノードに接続してください。

データ・マッピングは、ストリーム構築に密に統合されています。既に接続があるノードに接続する場合、接続を置き換える、またはそのノードにマッピングするオプションが提供されます。

データのテンプレートへのマッピング

テンプレート ストリームのデータ・ソースを、IBM SPSS Modeler に独自のデータを取り込む新しい入力ノードで置換するには、ポップアップ・メニューの「データ マッピング」の「置換ノードの選択」オプションを使用する必要があります。このオプションは、レコード結合ノード、レコード集計ノード、およびターミナル・ノードを除く、他のすべてのノードで利用することができます。データ・マッピング・ツールを使用してこの操作を行うことにより、既存のストリーム操作と新規データ・ソースのフィールドが正しく一致することが保証されます。データ・マッピング作業の概要を次に示します。

ステップ 1：元の入力ノードの必須フィールドを指定します。ストリーム操作が適切に行われるためには、必須フィールドを指定する必要があります。詳しくは、トピック 68 ページの『必須フィールドの指定』を参照してください。

ステップ 2：新しいデータ・ソースをストリーム領域に追加します。入力ノードを使用して、新しい置換データを取り込みます。

ステップ 3：テンプレート入力ノードを置換します。テンプレート入力ノードのポップアップ・メニューの「データ マッピング」オプションを使用して、「置換ノードの選択」をクリックし、置換データの入力ノードを選択します。

ステップ 4：マップされたフィールドを確認します。表示されているダイアログ・ボックスで、置換データ・ソースからストリームへ、フィールドが正しくマッピングされているかどうかを確認します。マップされていない必須フィールドは赤色で表示されています。これらのフィールドはストリーム操作で使用されるので、下流の操作が正しく行われるようにするためには、新しいデータ・ソース中の適切なフィールドと置き換える必要があります。詳しくは、トピック 68 ページの『マップされたフィールドの調査』を参照してください。

ダイアログ・ボックスですべての必須フィールドが正しくマップされていることを確認したら、マップと呼ばれるフィルター・ノードを使い、古いデータ・ソースの接続を解除して、新しいデータ・ソースをストリームに接続します。このフィルター・ノードは、ストリーム中のフィールドの実際のマッピングを指示します。ストリーム領域には、マップ解除フィルター・ノードも含まれています。マップ解除フィルター・ノードをストリームに追加して、フィールド名のマッピングを元に戻すことができます。これにより、フィール

ドのマッピングが解除されます。ただし、フィールドを再選択してオーバーレイするには、下流のターミナル・ノードを編集する必要があることに注意してください。

ストリーム間のマッピング

ノードの接続と同様に、この方法のデータ・マッピングでは、あらかじめ必須フィールドを設定しておく必要はありません。この方法では、データ・マッピングのポップアップ・メニューから「マップ先」を使用して、あるストリームを別のストリームに単に接続します。このデータ・マッピング手法は、ターミナル・ノードへのマッピングや、ストリーム間のコピーと貼り付けに役立ちます。注：「マップ先」オプションを使用して、レコード結合ノード、レコード追加ノード、およびすべての種類の入力ノードにマップすることはできません。

ストリーム間でデータをマップするには

1. 新規ストリームへの接続に使用するノードを右クリックします。
2. メニューで次の各項目をクリックします。

「データ マッピング」 > 「マップ先」

3. カーソルを使用して、対象ストリーム上の目的のノードを選択します。
4. 表示されたダイアログ・ボックスで、フィールドが正しく対応しているかどうかを確認して、「OK」をクリックします。

必須フィールドの指定

既存のストリームにマッピングする場合、通常はストリームの作成者により必須フィールドが指定されています。これらの必須フィールドは、特定のフィールドが下流の操作で使用されるかどうかを示します。例えば、既存のストリームで、*Churn* というフィールドを使用するモデルを作成している場合を考えてみましょう。フィールド *Churn* がないとモデルを作成できないため、このストリームでは *Churn* は必須フィールドになります。同様に、フィールド作成ノードなどの操作ノードで使われるフィールドは、新規フィールドを作成するために必要なフィールドになります。このようなフィールドを明示的に必須フィールドと設定すれば、新しい入力ノード中の適切なフィールドが確実にマップされます。必須フィールドが関連付けられていないと、エラー・メッセージが表示されます。ある操作ノードまたは出力ノードを不要と判断したら、それらのノードをストリームから削除し、該当するフィールドを「必須フィールド」リストから削除します。

必須フィールドを設定するには

1. 置換するテンプレート ストリームの入力ノードを右クリックします。
2. メニューで次の各項目をクリックします。

「データ マッピング」 > 「必須フィールドの指定」

3. フィールド・ピッカーを使用して、一覧にフィールドを追加したり削除することができます。フィールド・ピッカーを開くには、フィールド・リストの右にあるアイコンをクリックしてください。

マップされたフィールドの調査

あるデータ・ストリームまたはデータ・ソースを別のデータ・ストリームまたはデータ・ソースにマップする場所を選択すると、マッピングするフィールドを選択するかシステム・デフォルトのマッピングが正しいことを確認するためのダイアログ・ボックスが表示されます。ストリームまたはデータ・ソースの必須フィールドが設定されているけれども、それが一致していない場合、それらのフィールドは赤で表示されます。データ・ソースのマッピングされていないフィールドは、変更されずにフィルター・ノードを通過します。ただし、必須でないフィールドもマッピングできることに注意してください。

オリジナル: テンプレートまたは既存のストリーム中のすべてのフィールド、つまり下流に流れるすべてのフィールドが表示されます。新しいデータ・ソースのフィールドは、これらのフィールドにマップされません。

マップ済み: テンプレート・フィールドにマッピングするために選択したフィールドの一覧を表示します。これらのフィールドは、ストリーム操作で使われるオリジナルのフィールドと一致させるために、フィールド名を変更する必要がある場合があります。テーブル・セルをクリックすると、リストに利用可能なフィールドが表示されます。

マップするフィールドがわからない場合は、マップする前に入力データを調べることをお勧めします。例えば、入力ノードの「データ型」タブを使用して、入力データの概要を確認することができます。

ヒントとショートカット

ここで説明しているヒントやショートカットを覚えておくと、作業をより簡単かつ円滑に行うことができます。

- ストリームを素早く構築するにはダブルクリックを使用する: パレット上のノードを単にダブルクリックすれば、現在のストリーム上にそのノードが追加、接続されます。
- キーの組み合わせを使用して下流のノードを選択する: **Ctrl+Q** キーと **Ctrl+W** キーを使用して、下流のすべてのノード選択を切り替えます。
- ショートカット・キーを使用してノードの接続、接続解除を行う: ストリーム領域上のノードを選択している場合、**F2** キーを押した後、**Tab** キーを使用して接続先ノードを選択し、次に **Shift** キーを押しながらスペース・バーを押すことにより、元のノードと接続先ノードを接続することができます。選択したノードのすべての入力接続および出力接続を解除するには、**F3** キーを押します。
- 「ノード・パレット」タブによく使うノードを設定する: 「ツール」メニューの「パレット管理」をクリックすると、ダイアログ・ボックスが開きます。ここで、ノード・パレット上に表示されるノードを追加、削除、または移動します。
- ノードの名前の変更と追加に関するヒント: 各ノードのダイアログ・ボックスには「注釈」タブが用意されています。このタブでは、ノードのストリーム領域上での独自の名前を設定したり、ストリームを構築するために役立つツールヒントを追加することができます。また、進捗状況の追跡を行う目的で、または処理の説明、必要な作業、達成すべきビジネス上の目的などを記録するために、長い注釈を記入できる欄も用意されています。
- **CLEM** 式に値を自動的に入力する: さまざまなダイアログ・ボックス (フィールド作成ノードや置換ノードのダイアログ・ボックスなど) からアクセスできる式ビルダーを使用すると、**CLEM** 式にフィールド値を自動的に挿入できます。式ビルダーの値ボタンをクリックして、既存のフィールド値から選択してください。



図 13. 「値」ボタン

- ファイルを素早く参照する: 「開く」ダイアログ・ボックスでファイルを参照するときに、「ファイル」リストを使用して (ダイアログ・ボックスの上部で「検索対象」フィールドの横にある黄色いひし形のボタンをクリックします)、前に使用したディレクトリーや **IBM SPSS Modeler** のデフォルト・ディレクトリーにアクセスします。進むボタンおよび戻るボタンを使用して、過去に利用したディレクトリーを参照することができます。

- 不要な出力ウィンドウを最小化する：出力ウィンドウの右上角にある赤い X ボタンを使えば、出力ウィンドウを素早く閉じて削除することができます。こうすることにより、マネージャー・パネルの「出力」タブに、注目する結果や興味のある結果だけを保持することができます。

このソフトウェアでは、さまざまなキーボード・ショートカットを利用することができます。詳しくは、トピック 217 ページの『キーボード・アクセシビリティ』を参照してください。

ご存じですか

- マウスをドラッグして、ストリーム領域上の複数のノード グループを選択することができます。
- あるストリームから別のストリームにノードをコピーして貼り付けることができます。
- ダイアログ・ボックスや出力ウィンドウからヘルプを参照することができます。
- CRISP-DM (Cross Industry Standard Process for Data Mining) に関するヘルプを参照できます。
(「ヘルプ」メニューの「**CRISP-DM** ヘルプ」をクリックします。)

第 6 章 出力の処理

何らかのストリームを実行した場合、その結果はモデル ナゲットのノードの「モデル」タブまたは「詳細」タブによってビューアに表示されます。ビューアでは、表示したい出力に容易に移動できます。また、出力を操作し、目的の出力が正確に含まれている文書を作成することもできます。一部のグラフ出力でもビューアが使用されます。

ビューアは、IBM SPSS Modeler の以下の出力に使用されます。

- TCM モデル ナゲット
- STP モデル ナゲット
- TwoStep-AS クラスタのモデル ナゲット
- GSAR モデル ナゲット
- マップ視覚化グラフ作成ノード

ビューア

結果はビューアに表示されます。ビューアを使用すると、次の操作を実行できます。

- 結果を参照する
- 選択したテーブルおよびグラフの表示と非表示を切り替える
- 選択した項目を移動することによって、結果の表示順序を変更する
- ビューアと他のアプリケーションの間で項目を移動する

ビューアは、次の 2 つのペインに分割されています。

- 左のペインには、内容のアウトライン表示が含まれています。
- 右のペインには、統計テーブル、グラフ、およびテキスト出力が表示されます。

対応するテーブルまたはグラフに直接移動するには、アウトライン内の項目をクリックします。アウトライン・ペインの右罫線をクリックしてドラッグすると、アウトライン・ペインの幅を変更できます。

結果の表示と非表示

ビューアでは、プロシージャ全体から個別のテーブルまたは結果を選択的に表示することも、非表示にすることもできます。このプロセスは、コンテンツ・ペインに表示される出力の量を少なくする場合に有効です。

テーブルとグラフを非表示にするには

1. ビューアのアウトライン・ペインにある項目のブック・アイコンをダブルクリックします。

または

2. 項目をクリックして選択します。
3. メニューから次の項目を選択します。

「表示」 > 「非表示」

または

4. アウトラインのツールバーで、閉じたブックの (非表示) アイコンをクリックします。

開いたブック (表示) アイコンがアクティブ・アイコンになり、項目が現在非表示になっていることを示します。

プロシージャの結果を非表示にするには

1. アウトライン・ペインで、プロシージャ名の左にあるボックスをクリックします。

これにより、プロシージャの結果がすべて非表示になり、アウトライン表示が省略されます。

出力の移動、削除、およびコピー

項目を 1 つずつ、またはグループでコピー、移動、削除して、結果を再配置することができます。

ビューアで出力を移動するには

1. アウトライン・ペインまたはコンテンツ・ペインで項目を選択します。
2. 選択したアイテムを別の場所にドラッグ・アンド・ドロップします。

ビューアで出力を削除するには

1. アウトライン・ペインまたはコンテンツ・ペインで項目を選択します。
2. **Delete** キーを押します。

または

3. メニューから次の項目を選択します。

「編集」 > 「削除」

初期位置合わせの変更

デフォルトでは、すべての結果は最初に左揃えされます。新規出力項目の初期位置合わせを変更するには、次の手順に従います。

1. メニューから次の項目を選択します。

「編集」 > 「オプション」

2. 「ビューア」タブをクリックします。
3. 「初期出力状態」グループで、項目タイプ (例えば、ピボット・テーブル、グラフ、テキスト出力) を選択します。
4. 目的の位置合わせオプションを選択します。

出力項目の位置合わせの変更

1. アウトライン・ペインまたはコンテンツ・ペインで、位置合わせする項目を選択します。
2. メニューから次の項目を選択します。

「形式」 > 「左揃え」

または

「形式」 > 「中央揃え」

または

「形式」 > 「右揃え」

ビューアのアウトライン

アウトライン・ペインには、ビューア文書の目次が表示されます。アウトライン・ペインを使用すると、結果間の移動および表示の制御を行うことができます。アウトライン・ペインでのほとんどの操作により、コンテンツ・ペインはそれに対応する影響を受けます。

- アウトライン・ペインで項目を選択すると、コンテンツ・ペイン内に対応する項目が表示されます。
- アウトライン・ペインで項目を移動すると、コンテンツ・ペイン内の対応する項目が移動します。
- アウトライン表示を省略すると、省略したレベルのすべての項目の結果が非表示になります。

アウトライン表示の制御。アウトライン表示は次のように制御できます。

- アウトライン表示を展開したり省略したりする
- 選択した項目のアウトライン・レベルを変更する
- アウトライン表示の項目をサイズ変更する
- アウトライン表示で使用するフォントを変更する

アウトライン表示を省略したり展開したりするには

1. 省略するまたは展開するアウトライン項目の左にあるボックスをクリックします。

または

2. アウトラインの項目をクリックします。
3. メニューから次の項目を選択します。

「表示」 > 「閉じる」

または

「表示」 > 「展開」

アウトライン・レベルを変更するには

1. アウトライン・ペインの項目をクリックします。
2. メニューから次の項目を選択します。

「編集」 > 「アウトライン」 > 「レベルを上げる」

または

「編集」 > 「アウトライン」 > 「レベルを下げる」

アウトライン項目のサイズを変更するには

1. メニューから次の項目を選択します。

「表示」 > 「アウトラインのサイズ」

2. アウトライン・サイズを選択します（「小」、「中」、または「大」）。

アウトラインのフォントを変更するには

1. メニューから次の項目を選択します。

「表示」 > 「アウトライン・フォント...」

2. フォントを選択します。

ビューアへの項目の追加

ビューアには、表題、新規テキスト、グラフ、他のアプリケーションの要素などの項目を追加することができます。

タイトルまたはテキストを追加するには

テーブルまたはグラフに結合されていないテキスト項目は、ビューアに追加することができます。

1. タイトルまたはテキストを追加する位置の直前にあるテーブル、グラフ、またはその他のオブジェクトをクリックします。
2. メニューから次の項目を選択します。

「挿入」 > 「新しい表題」

または

「挿入」 > 「新しいテキスト」

3. 新規オブジェクトをダブルクリックします。
4. テキストを入力します。

テキスト・ファイルを追加するには

1. ビューアのアウトライン・ペインまたはコンテンツ・ペインで、テキストを追加する位置の直前にあるテーブル、グラフ、またはその他のオブジェクトをクリックします。
2. メニューから次の項目を選択します。

「挿入」 > 「テキスト・ファイル...」

3. テキスト・ファイルを選択します。

テキストを編集するには、そのテキストをダブルクリックします。

ビューアへのオブジェクトの貼り付け

ビューアには、他のアプリケーションのオブジェクトも貼り付けることができます。「後ろに貼り付け」または「形式を指定して貼り付け」のいずれかを使用できます。いずれの方法でも、ビューア内で現在選択されているオブジェクトの後に新規オブジェクトが貼り付けられます。貼り付けるオブジェクトの形式を選択する場合は、「形式を指定して貼り付け」を使用します。

ビューアの情報検索および置換

1. ビューアの情報検索または置換するには、メニューから次の項目を選択します。

「編集」 > 「検索」

または

「編集」 > 「置換」

検索と置換を使用して、次のことができます。

- 文書全体または選択した項目だけを検索。
- 現在の場所から下または上に検索。
- コンテンツ・ペインまたはアウトライン・ペインのいずれかに制限するか、両方のペインを検索。
- 非表示項目を検索。コンテンツ・ペインの非表示項目 (例えば、デフォルトでは非表示の記録テーブル) およびピボット・テーブルの非表示の行および列などの項目を含みます。
- 大文字と小文字を区別するように検索基準を制限。
- ピボット・テーブルの検索基準をすべてのセルの内容の一致に制限。
- ピボット・テーブルの検索基準を脚注マーカーのみに制限。このオプションは、ビューアでの選択にピボット・テーブル以外のものが含まれている場合、使用できません。

非表示の項目およびピボット・テーブル層

- 複合多次元のピボット・テーブルの現在表示されている層の下層は非表示とは見なされず、非表示の項目が検索に含まれない場合でも検索領域に含まれます。
- 非表示項目には、コンテンツ・ペインの非表示項目 (アウトライン・ペインで閉じた本のアイコンが付いた項目またはアウトライン・ペインの省略されたブロックに含まれる項目)、およびデフォルトで非表示になっている (例えば、空白の行と列はデフォルトでは非表示) か、テーブルを編集し、特定の行や列を選択して手動で非表示にしたピボット・テーブルの行および列が含まれます。非表示項目は、「非表示項目を含む」を明示的に選択している場合のみ、検索に含まれます。
- どちらの場合も、検索テキストまたは値を含む非表示または目に見えない要素は、検出されると表示されますが、その後項目は元の状態に戻ります。

ピボット・テーブルの値の範囲の検索

ピボット・テーブルで指定された値の範囲の値を検索するには

1. ビューアでピボット・テーブルをアクティブ化するか、1 つまたは複数のピボット・テーブルを選択します。ピボット・テーブルだけを選択してください。他のオブジェクトを選択すると、「範囲」オプションが使用できなくなります。
2. メニューから次の項目を選択します。

「編集」 > 「検索」

3. 「範囲」タブをクリックします。
4. 範囲の種類を「次の値の範囲内」、「次の値以上」、「次の値以下」から選択します。
5. 範囲を定義する値を選択します。
 - いずれかの値が非数値文字を含む場合、どちらの値も文字列として処理されます。
 - 両方の値が数値の場合、数値だけが検索されます。
 - 「範囲」タブを使用して値を置き換えることはできません。

この機能は、レガシー・テーブルには使用できません。詳しくは、トピック 102 ページの『レガシー・テーブル』を参照してください。

他のアプリケーションへの出力のコピー

出力オブジェクトをコピーして、ワード・プロセッサ・プログラムやスプレッドシートなどのその他のアプリケーションに貼り付けることができます。出力はさまざまな形式で貼り付けることができます。ターゲット・アプリケーションおよび選択した出力オブジェクトに応じて、次の形式の一部またはすべてが使用可能となります。

メタファイル。WMF および EMF メタファイル形式。これらの形式は、Windows オペレーティング・システムでのみ使用できます。

RTF (リッチ・テキスト・フォーマット)。選択した複数のオブジェクト、テキスト出力、ピボット・テーブルを RTF 形式でコピーして貼り付けることができます。ピボット・テーブルの場合、ほとんどのアプリケーションでは、後で他のアプリケーションで編集できるテーブルとしてテーブルが貼り付けられます。文書の幅に対して広すぎるピボット・テーブルは、ピボット・テーブルのオプション設定に応じて折り返されるか、文書の幅に合わせて縮小されるか、または変更しないまま残されます。詳しくは、トピック 103 ページの『ピボット・テーブル・オプション』を参照してください。

注: Microsoft Word では、極端に幅広いテーブルは適切に表示されないことがあります。

イメージ。JPG および PNG イメージ形式。

BIFF。ピボット・テーブルおよびテキスト出力は、BIFF 形式のスプレッドシートに貼り付けることができます。ピボット・テーブルの数値は、数値の精度を保持します。この形式は、Windows オペレーティング・システムでのみ使用できます。

テキスト。ピボット・テーブルおよびテキスト出力は、テキストとしてコピーし、貼り付けることができます。このプロセスは、電子メールなど、テキストだけを送受信できるアプリケーションで有用です。

Microsoft Office グラフィック オブジェクト。この形式をサポートするグラフは、Microsoft Office アプリケーションにコピーして、ネイティブの Microsoft Office グラフとしてこれらのアプリケーション内で編集できます。SPSS Statistics/SPSS Modeler のグラフと Microsoft Office のグラフの違いのため、SPSS Statistics/SPSS Modeler のグラフの機能によっては、コピーされたバージョンで維持されないものがあります。Microsoft Office グラフィック オブジェクト形式のグラフを複数選択してコピーする操作は、サポートされていません。

ターゲット・アプリケーションが複数の使用可能な形式をサポートしている場合は、形式を選択できる「形式を指定して貼り付け」メニュー項目が存在するか、または使用可能な形式のリストが自動的に表示されません。

複数の出力オブジェクトのコピーと貼り付け

複数の出力オブジェクトを他のアプリケーションに貼り付ける場合は、次の制限が適用されます。

- **RTF** 形式。ほとんどのアプリケーションでは、ピボット・テーブルはそのアプリケーションで編集可能なテーブルとして貼り付けられます。グラフ、ツリー、およびモデル・ビューは、イメージとして貼り付けられます。
- メタファイルおよびイメージ形式。選択したすべての出力オブジェクトは、他のアプリケーションの単一のオブジェクトとして貼り付けられます。
- **BIFF** 形式。グラフ、ツリー、およびモデル・ビューは除外されます。

形式を選択してコピー

大量の出力、特に非常に大きなピボット・テーブルをコピーして貼り付ける際には、「編集」 > 「特殊コピー」を使用して、クリップボードにコピーされる形式の数を制限することにより、操作の速度を向上させることができます。

選択した形式をデフォルトの形式セットとして保存し、クリップボードにコピーすることもできます。この設定は、セッションが変わっても保持されます。

インタラクティブ出力

インタラクティブ出力オブジェクトには、複数の関連する出力オブジェクトが含まれています。1 つのオブジェクトで選択を行うと、他のオブジェクトで表示または強調表示される内容を変更できます。例えば、テーブルでいずれかの行を選択すると、マップ内の 1 つのエリアが強調表示されたり、異なるカテゴリのグラフが表示されたりする場合があります。

インタラクティブ出力オブジェクトでは、編集機能 (例えば、テキスト、色、フォント、テーブルの罫線の変更など) はサポートされません。個別のオブジェクトをインタラクティブ オブジェクトからビューアにコピーできます。インタラクティブ出力からコピーされたテーブルは、ピボット テーブル エディタで編集できます。

インタラクティブ出力からのオブジェクトのコピー

「ファイル」 > 「ビューアにコピー (Copy to Viewer)」を使用して、個別の出力オブジェクトを「ビューア」ウィンドウにコピーします。

- 使用可能なオプションは、インタラクティブ出力の内容によって異なります。
- 「グラフ」および「マップ」では、グラフ オブジェクトを作成します。
- 「テーブル」ではピボット テーブルを作成します。ピボット テーブルはピボット テーブル エディタで編集可能です。
- 「スナップショット」では、現在のビューのイメージを作成します。
- 「モデル」では、現在のインタラクティブ出力オブジェクトのコピーを作成します。

「編集」 > 「オブジェクトをコピー」では、個別の出力オブジェクトをクリップボードにコピーします。

- コピーしたオブジェクトをビューアに貼り付ける操作は、「ファイル」 > 「ビューアにコピー (Copy to Viewer)」に相当します。
- オブジェクトを別のアプリケーションに貼り付けると、そのオブジェクトはイメージとして貼り付けられます。

ズームとパン

マップでは、「ビュー」 > 「ズーム」を使用してマップのビューをズームできます。ズームしたマップのビュー内で、「ビュー」 > 「パン (Pan)」を使用してビューを移動できます。

印刷設定

「ファイル」 > 「印刷設定 (Print Settings)」では、インタラクティブ オブジェクトの印刷方法を制御します。

- 「表示されているビューのみを印刷」。現在表示されているビューのみを印刷します。このオプションがデフォルトの設定値です。

- 「すべてのビューを印刷」。インタラクティブ出力に含まれるすべてのビューを印刷します。
- 選択したオプションは、出力オブジェクトをエクスポートするためのデフォルト アクションも決定します。

出力のエクスポート

「出力のエクスポート」では、ビューアの出力を HTML、テキスト、Word/RTF、Excel、PowerPoint (PowerPoint 97 以降が必要です)、および PDF の各形式で保存します。グラフは、さまざまなグラフィックス形式でもエクスポートできます。

注: PowerPoint へのエクスポートは Windows オペレーティング・システムでのみ利用できます。

出力をエクスポートするには

1. ビューアをアクティブ・ウィンドウにします (ウィンドウ内の任意の場所をクリックします)。
2. ツールバーで「エクスポート」ボタンをクリックするか、出力ウィンドウで右クリックして「エクスポート」を選択します。
3. ファイル名 (またはグラフの接頭辞) を入力し、エクスポート形式を選択します。

「エクスポートするオブジェクト」。ビューア内のすべてのオブジェクト、表示されているすべてのオブジェクト、または選択したオブジェクトのみをエクスポートすることができます。

「文書タイプ」。使用可能なオプションは次のとおりです。

- **Word/RTF**。ピボット テーブルは Word テーブルとしてそのままエクスポートされます。セルの罫線、フォント スタイル、背景色など、すべての書式属性も反映されます。テキスト出力は書式付きの RTF としてエクスポートされます。グラフ、ツリー図、モデル・ビューは PNG 形式で保存されます。Microsoft Word では、極端に幅広いテーブルは適切に表示されないことがあるので注意してください。
- **Excel**。ピボット テーブルの行、列、セルは Excel の行、列、セルとしてエクスポートされます。この場合、セルの罫線、フォント スタイル、背景色など、すべての書式属性もそのまま反映されます。テキスト出力はすべてのフォント属性をそのままにエクスポートされます。テキスト出力の各行は Excel ファイルの行になります。この場合、1 行の内容はすべて 1 つのセルに入れられます。グラフ、ツリー図、モデル・ビューは PNG 形式で保存されます。出力は Excel 97-2004 または Excel 2007 以降としてエクスポートできます。
- **HTML**。ピボット・テーブルは HTML テーブルとしてエクスポートされます。テキスト出力は、事前にフォーマットされた HTML としてエクスポートされます。グラフ、ツリー図、モデル・ビューは、選択したグラフィック形式で文書に埋め込まれます。HTML 形式でエクスポートした出力を表示するには、HTML 5 と互換性のあるブラウザが必要です。
- **Web レポート**。Web レポートは、ほとんどのブラウザと互換性のあるインタラクティブ文書です。ビューアで使用可能なピボット・テーブルのインタラクティブ機能は、その多くが Web レポートでも使用可能です。また、Web レポートを IBM Cognos® Active Report としてエクスポートすることもできます。
- **PDF 形式**。出力はすべて「印刷プレビュー」に表示されるとおりにエクスポートされます。すべての書式属性が維持されます。
- **PowerPoint ファイル**。ピボット・テーブルが Word テーブルとしてエクスポートされ、PowerPoint ファイルの個別のスライドに埋め込まれます。ピボット・テーブルにつき 1 つのスライドです。セルの罫線、フォント スタイル、背景色など、ピボット テーブルのあらゆる書式属性が保持されます。グラフ、ツリー図、モデル ビューは TIFF 形式でエクスポートされます。テキスト出力は含まれません。

PowerPoint へのエクスポートは Windows オペレーティング・システムでのみ利用できます。

- テキスト。テキスト出力形式はプレーン・テキスト、UTF-8、UTF-16 です。ピボット・テーブルは、タブ区切り形式またはスペース区切り形式でエクスポートできます。テキスト出力はすべてスペース区切り書式でエクスポートされます。グラフ、ツリー図、モデル・ビューの場合、グラフィックごとのテキスト・ファイルにイメージのファイル名を示す行が 1 行挿入されます。
- 「なし (グラフィックのみ)」。利用可能なエクスポート形式は、EPS、JPEG、TIFF、PNG、および BMP です。Windows オペレーティング・システムでは、EMF (拡張メタファイル) 形式も使用できません。

「上位のフォルダを開く」。エクスポートによって作成されたファイルを含むフォルダーを開きます。

HTML のオプション

HTML のエクスポートには、HTML 5 と互換性のあるブラウザが必要です。HTML 形式での出力のエクスポートでは、次のオプションを使用できます。

ピボット・テーブルの層。デフォルトでは、ピボット・テーブル層の包含または除外は各ピボット・テーブルのテーブル・プロパティによりコントロールされます。この設定を上書きして、すべての層を含めることも、または現在見えている層以外のすべてを除くこともできます。詳しくは、トピック 97 ページの『テーブルのプロパティ: 印刷』を参照してください。

「階層化テーブルを対話型としてエクスポート」。階層化されたテーブルはビューアでの表示どおりに表示され、ユーザーはブラウザに表示される階層をインタラクティブに変更できます。このオプションを選択しない場合、各テーブル層は独立したテーブルとして表示されます。

「HTML 形式のテーブル」。エクスポートされたピボット・テーブルに含めるスタイル情報を制御します。

- 「スタイルおよび固定の列幅を指定してエクスポート」。ピボット・テーブルのすべてのスタイル情報 (フォントのスタイル、背景色など) および列幅を維持します。
- 「スタイルを指定せずにエクスポート」。ピボット・テーブルをデフォルトの HTML テーブルに変換します。スタイル属性は保存されません。列幅は自動的に決定されます。

「脚注とキャプションを含める」。ピボット・テーブルのすべての脚注とキャプションを含めるか除外するかを制御します。

モデルのビュー。デフォルトでは、モデル・ビューの包含または除外は、各モデルのモデル・プロパティによってコントロールされます。この設定を上書きして、すべてのビューを含むことも、または現在見えているビュー以外のすべてを除くこともできます。詳しくは、トピックモデル・プロパティを参照してください。(注: テーブルを含むすべてのモデル・ビューがグラフィックスとしてエクスポートされます。)

注: HTML の場合、エクスポートされたグラフのイメージ・ファイル形式も制御できます。詳しくは、トピック 84 ページの『グラフィックス形式オプション』を参照してください。

HTML エクスポート・オプションを設定するには

1. エクスポート形式として「HTML」を選択します。
2. 「オプション変更」をクリックします。

Web レポートのオプション

Web レポートは、ほとんどのブラウザと互換性のあるインタラクティブ文書です。ビューアで使用可能なピボット・テーブルのインタラクティブ機能は、その多くが Web レポートでも使用可能です。

「レポートのタイトル」。レポートのヘッダーに表示されるタイトルです。デフォルトではファイル名が使用されます。ファイル名の代わりに使用するカスタム・タイトルを指定できます。

「形式」。レポート形式には以下の 2 つのオプションがあります。

- 「**SPSS Web レポート (HTML 5)**」。この形式を使用するには、HTML 5 と互換性のあるブラウザが必要です。
- 「**Cognos Active Report (mht)**」。この形式を使用するには、MHT 形式のファイルまたは Cognos Active Report アプリケーションをサポートするブラウザが必要です。

「オブジェクトの除外」。選択したオブジェクト・タイプをレポートから除外できます。

- 「テキスト」。ログではないテキスト・オブジェクト。このオプションでは、アクティブなデータ・セットについての情報を含むテキスト・オブジェクトが含まれます。
- 「ログ」。実行されたコマンド シンタックスのリストを含むテキスト オブジェクト。ログ項目には、ビューア出力を生成しないコマンドで発生した警告およびエラー メッセージも含まれます。
- 「メモ テーブル」。統計およびグラフ作成プロシージャの出力には、「メモ テーブル」が含まれます。このテーブルには、使用されたデータ セット、欠損値、およびプロシージャを実行するために使用されたコマンド シンタックスに関する情報が含まれています。
- 「警告とエラー メッセージ」。統計およびグラフ作成プロシージャの警告およびエラー・メッセージ。

「Web レポートに合うようにテーブルやグラフのスタイルを変更します」。このオプションは、標準 Web レポートのスタイルをすべてのテーブルとグラフに適用します。この設定は、ビューアに表示されている出力内のフォント、色などのスタイルをすべて上書きします。標準 Web レポートのスタイルは変更できません。

「**Web サーバー接続 (Web Server Connection)**」。IBM SPSS Statistics Web レポート アプリケーション サーバーを実行している 1 つ以上のアプリケーション サーバーの URL ロケーションを含めることができます。Web アプリケーション サーバーは、テーブルのピボット、グラフの編集、および変更済み Web レポートの保存を行うための機能を提供します。

- Web レポートに含める各アプリケーション サーバーで「使用」を選択します。
- Web レポートに URL の指定が含まれる場合、その Web レポートはそのアプリケーション サーバーに接続して追加の編集機能を提供します。
- 複数の URL を指定すると、Web レポートはその指定順で各サーバーへの接続を試みます。

IBM SPSS Statistics Web レポート アプリケーション サーバーは、<http://www.ibm.com/developerworks/spssdevcentral> からダウンロードできます。

Word/RTF のオプション

Word 形式での出力のエクスポートでは、次のオプションを使用できます。

ピボット・テーブルの層。デフォルトでは、ピボット・テーブル層の包含または除外は各ピボット・テーブルのテーブル・プロパティによりコントロールされます。この設定を上書きして、すべての層を含めることも、または現在見えている層以外のすべてを除くこともできます。詳しくは、トピック 97 ページの『テーブルのプロパティ: 印刷』を参照してください。

幅広のピボット・テーブル。定義された文書の幅に対して広すぎるテーブルの処理をコントロールします。デフォルトでは、テーブルはちょうど収まるように折り返されます。テーブルはセクションに分割され、行ラベルはテーブルのセクションごとに繰り返されます。あるいは、幅の広いテーブルを縮小するか、変更せずに定義された文書の幅を超えて拡張することができます。

「ブレイクポイントを保持」。ブレイクポイントが定義されている場合、それらの設定が Word テーブルで保持されます。

「脚注とキャプションを含める」。ピボット・テーブルのすべての脚注とキャプションを含めるか除外するかを制御します。

モデルのビュー。デフォルトでは、モデル・ビューの包含または除外は、各モデルのモデル・プロパティによってコントロールされます。この設定を上書きして、すべてのビューを含むことも、または現在見えているビュー以外のすべてを除くこともできます。詳しくは、トピックモデル・プロパティを参照してください。(注: テーブルを含むすべてのモデル・ビューがグラフィックスとしてエクスポートされます。)

「エクスポートのページ設定」。エクスポートする文書のページ・サイズおよび余白を定義できるダイアログが開きます。折り返しおよび縮小の動作を決定するために使用される文書の幅は、ページの幅から左右の余白を引いた長さです。

Word エクスポート・オプションを設定するには

1. エクスポート形式として「**Word/RTF**」を選択します。
2. 「オプション変更」をクリックします。

Excel のオプション

Excel 形式での出力のエクスポートでは、次のオプションを使用できます。

「ワークシートの作成」、「ワークブックの作成」、または「既存のワークシートを変更」。デフォルトでは、新しいワークブックが作成されます。指定した名前のファイルが既に存在する場合は、そのファイルが上書きされます。ワークシートを作成するオプションを選択した場合、指定した名前のワークシートが指定したファイルに既に存在する場合は、そのワークシートが上書きされます。既存のワークシートを変更するオプションを選択した場合、ワークシート名も指定する必要があります (これはワークシート作成のオプションです)。ワークシート名は 31 文字以下にする必要があり、スラッシュ、バックスラッシュ、大括弧、疑問符、アスタリスクを含めることはできません。

Excel 97-2004 にエクスポートする場合、既存のワークシートを変更すると、グラフ、モデル・ビュー、およびツリー図はエクスポートされた出力に含まれません。

「ワークシート内の位置」。エクスポートされた出力のワークシート内の場所を制御します。デフォルトでは、エクスポートされた出力は、内容が入力されている最後の列の後の最初の行から、既存の内容を変更せずに追加されます。これは、既存のワークシートに新しい列を追加する場合に適しています。既存のワークシートに新しい行を追加する場合は、エクスポートされた出力を最後の行の後に追加するのが適しています。エクスポートした出力を特定のセルの場所に追加すると、出力が追加された領域の既存の内容が上書きされます。

ピボット・テーブルの層。デフォルトでは、ピボット・テーブル層の包含または除外は各ピボット・テーブルのテーブル・プロパティによりコントロールされます。この設定を上書きして、すべての層を含めることも、または現在見えている層以外のすべてを除くこともできます。詳しくは、トピック 97 ページの『テーブルのプロパティ: 印刷』を参照してください。

「脚注とキャプションを含める」。ピボット・テーブルのすべての脚注とキャプションを含めるか除外するかを制御します。

モデルのビュー。デフォルトでは、モデル・ビューの包含または除外は、各モデルのモデル・プロパティによってコントロールされます。この設定を上書きして、すべてのビューを含むことも、または現在見えているビュー以外のすべてを除くこともできます。詳しくは、トピックモデル・プロパティを参照してください。(注: テーブルを含むすべてのモデル・ビューがグラフィックスとしてエクスポートされます。)

Excel エクスポート・オプションを設定するには

1. エクスポート形式として「**Excel**」を選択します。
2. 「オプション変更」をクリックします。

PowerPoint のオプション

PowerPoint では次のオプションを使用できます。

ピボット・テーブルの層。デフォルトでは、ピボット・テーブル層の包含または除外は各ピボット・テーブルのテーブル・プロパティによりコントロールされます。この設定を上書きして、すべての層を含めることも、または現在見えている層以外のすべてを除くこともできます。詳しくは、トピック 97 ページの『テーブルのプロパティ: 印刷』を参照してください。

幅広のピボット・テーブル。定義された文書の幅に対して広すぎるテーブルの処理をコントロールします。デフォルトでは、テーブルはちょうど収まるように折り返されます。テーブルはセクションに分割され、行ラベルはテーブルのセクションごとに繰り返されます。あるいは、幅の広いテーブルを縮小するか、変更せずに定義された文書の幅を超えて拡張することができます。

「脚注とキャプションを含める」。ピボット・テーブルのすべての脚注とキャプションを含めるか除外するかを制御します。

「スライド・タイトルとしてビューアのアウトライン・エントリを使用」。エクスポートによって作成される各スライドにタイトルを含めます。各スライドには、ビューアからエクスポートされた項目が 1 つ含まれます。タイトルは、ビューアのアウトライン・ペインにある項目のアウトライン・エントリから作成されます。

モデルのビュー。デフォルトでは、モデル・ビューの包含または除外は、各モデルのモデル・プロパティによってコントロールされます。この設定を上書きして、すべてのビューを含むことも、または現在見えているビュー以外のすべてを除くこともできます。詳しくは、トピックモデル・プロパティを参照してください。(注: テーブルを含むすべてのモデル・ビューがグラフィックスとしてエクスポートされます。)

「エクスポートのページ設定」。エクスポートする文書のページ・サイズおよび余白を定義できるダイアログが開きます。折り返しおよび縮小の動作を決定するために使用される文書の幅は、ページの幅から左右の余白を引いた長さです。

PowerPoint エクスポート・オプションを設定するには

1. エクスポート形式として「**PowerPoint**」を選択します。
2. 「オプション変更」をクリックします。

注: PowerPoint へのエクスポートを利用できるのは Windows オペレーティング・システムだけです。

PDF のオプション

PDF では次のオプションを使用できます。

ブックマークを埋め込む。このオプションを選択すると、ビューアのアウトラインのエントリーに対応するブックマークが PDF 文書に挿入されます。ブックマークを埋め込むことにより、ビューアのアウトライン・ペインと同様に、多数の出力オブジェクトを含む文書内の移動が簡単になります。

フォントを埋め込む。フォントを埋め込むと、PDF 文書をどのコンピューターでも同じように表示させることができます。フォントが埋め込まれていない場合、PDF 文書を表示 (または印刷) するとき、コンピューターで利用できない一部のフォントに対して代用フォントが使用されます。ただし、この場合は、最適な結果が得られないことがあります。

ピボット・テーブルの層。デフォルトでは、ピボット・テーブル層の包含または除外は各ピボット・テーブルのテーブル・プロパティによりコントロールされます。この設定を上書きして、すべての層を含めることも、または現在見えている層以外のすべてを除くこともできます。詳しくは、トピック 97 ページの『テーブルのプロパティ: 印刷』を参照してください。

モデルのビュー。デフォルトでは、モデル・ビューの包含または除外は、各モデルのモデル・プロパティによってコントロールされます。この設定を上書きして、すべてのビューを含むことも、または現在見えているビュー以外のすべてを除くこともできます。詳しくは、トピックモデル・プロパティを参照してください。(注: テーブルを含むすべてのモデル・ビューがグラフィックスとしてエクスポートされます。)

PDF エクスポート・オプションを設定するには

1. エクスポート形式として「PDF 形式」を選択します。
2. 「オプション変更」をクリックします。

PDF 出力に影響を与えるその他の設定

ページ設定/ページ属性。PDF 文書のページ・サイズ、方向、余白、ページのヘッダーやフッターの内容と表示、およびグラフの印刷サイズは各種のページ設定オプションおよびページ属性オプションで制御します。

テーブル・プロパティ/テーブル・ルック。幅の広いテーブルや縦に長いテーブルの拡大縮小、およびテーブル層の印刷は、各テーブルのテーブル・プロパティで制御します。これらのプロパティはテーブル・ルックに保存することもできます。

デフォルト/現在のプリンター。PDF 文書の解像度 (DPI) は、デフォルトの解像度設定または現在選択されているプリンタに適用される現在の解像度設定となります (「ページ設定」で変更できます)。最大解像度は 1200 DPI です。プリンタ設定の方が解像度が高い場合、PDF 文書の解像度は 1200 DPI になります。

注: 高解像度の文書を低解像度プリンタで印刷すると、最適な印刷結果が得られないことがあります。

テキストのオプション

テキストのエクスポートでは次のオプションを使用できます。

「ピボット・テーブル形式」。ピボット・テーブルは、タブ区切り形式またはスペース区切り形式でエクスポートできます。スペース区切り形式では、次の制御も可能です。

- 列幅。「オートフィット」は表示内容を折り返さず、各列はその列の最大幅のラベルまたは値と同じ幅です。「カスタム」はテーブルのすべての列に適用される最大列幅と、その幅を超え、列の次の行に折り返す値を設定します。
- 行と列の罫線文字。行と列の罫線の作成に使用される文字をコントロールします。行と列の罫線を非表示にするには、値にスペースを入力します。

ピボット・テーブルの層。デフォルトでは、ピボット・テーブル層の包含または除外は各ピボット・テーブルのテーブル・プロパティによりコントロールされます。この設定を上書きして、すべての層を含めることも、または現在見えている層以外のすべてを除くこともできます。詳しくは、トピック 97 ページの『テーブルのプロパティ: 印刷』を参照してください。

「脚注とキャプションを含める」。ピボット・テーブルのすべての脚注とキャプションを含めるか除外するかを制御します。

モデルのビュー。デフォルトでは、モデル・ビューの包含または除外は、各モデルのモデル・プロパティによってコントロールされます。この設定を上書きして、すべてのビューを含むことも、または現在見えているビュー以外のすべてを除くこともできます。詳しくは、トピックモデル・プロパティを参照してください。(注: テーブルを含むすべてのモデル・ビューがグラフィックスとしてエクスポートされます。)

テキスト・エクスポート・オプションを設定するには

1. エクスポート形式として「テキスト」を選択します。
2. 「オプション変更」をクリックします。

グラフィックスのみのオプション

グラフィックスのみのエクスポートでは、次のオプションを使用できます。

モデルのビュー。デフォルトでは、モデル・ビューの包含または除外は、各モデルのモデル・プロパティによってコントロールされます。この設定を上書きして、すべてのビューを含むことも、または現在見えているビュー以外のすべてを除くこともできます。詳しくは、トピックモデル・プロパティを参照してください。(注: テーブルを含むすべてのモデル・ビューがグラフィックスとしてエクスポートされます。)

グラフィックス形式オプション

HTML およびテキスト文書およびグラフ限定のエクスポートでは、グラフィック形式を選択できます。各グラフィック形式でさまざまなオプションの設定を制御できます。

エクスポートするグラフのグラフィック形式とオプションを選択するには

1. 文書タイプとして「HTML」、「テキスト」、または「なし (グラフィックのみ)」を選択します。
2. ドロップダウン・リストからグラフィック・ファイル形式を選択します。
3. 選択したグラフィック・ファイル形式のオプションを変更するには、「オプション変更」をクリックします。

JPEG グラフのエクスポート・オプション

- 「イメージ・サイズ」。元のグラフのサイズのパーセントで、最大 200 パーセントです。
- 「グレースケールに変換」。カラーから灰色の陰影に変換します。

BMP グラフのエクスポート・オプション

- 「イメージ・サイズ」。元のグラフのサイズのパーセントで、最大 200 パーセントです。

- ・ 「イメージを圧縮してファイル・サイズを縮小」。可逆的圧縮技術でイメージ品質を損なわずにより小さなファイルを作成します。

PNG グラフのエクスポート・オプション

「イメージ・サイズ」。元のグラフのサイズのパーセントで、最大 200 パーセントです。

「色の深さ」。エクスポートされるグラフ内の色の数を決定します。ある深さで保存されたグラフには、実際に使用される色の最小数と、深さで許可される色の最大数があります。例えば、3 つの色 (赤、白、および黒) を含むグラフを 16 色で保存する場合、グラフは 3 色のままで保持されます。

- ・ グラフの色数が深さで指定された色数を越えている場合、グラフの色を再現するために色がディザリングされます。
- ・ 「現在の画面の濃さ」は、コンピューターのモニターに現在表示されている色の数です。

EMF および TIFF グラフのエクスポート・オプション

「イメージ・サイズ」。元のグラフのサイズのパーセントで、最大 200 パーセントです。

注: EMF (拡張メタファイル) 形式は、Windows オペレーティング・システムでのみ使用できます。

EPS グラフのエクスポート・オプション

「イメージ・サイズ」。元のイメージ・サイズのパーセント (200 パーセントまで) でサイズを指定するか、またはピクセルでイメージの幅を指定できます (高さは幅の値と縦横比により決定されます)。エクスポートされたイメージは元のイメージに常に比例します。

「**TIFF** プレビュー・イメージを含む」。EPS イメージを含むプレビューを TIFF 形式で保存して、画面上に EPS イメージを表示できないアプリケーションでも表示できるようにします。

「フォント」。EPS イメージでのフォント処理を制御します。

- ・ 「フォント参照を使用」。出力デバイスがグラフで使用されているフォントに対応している場合、そのフォントを使用します。対応していないフォントについては、代替フォントが使用されます。
- ・ 「フォントを曲線に置換」。フォントを PostScript 曲線データに変換します。テキスト自体は、EPS グラフィックスを編集できるアプリケーションで、テキストとして編集できなくなります。出力デバイスがグラフに使用されているフォントに対応していない場合は、このオプションが便利です。

ビューアの印刷

「ビューア」ウィンドウの内容を印刷するには、次の 2 つのオプションがあります。

「表示されているすべての出力」。コンテンツ・ペインに現在表示されている項目だけを印刷します。非表示の項目 (アウトライン・ペインに閉じたブック・アイコンのある項目、または省略されたアウトライン層に隠れた項目) は印刷されません。

「選択」。アウトライン・ペインかコンテンツ・ペイン、またはその両方で現在選択されている項目だけを印刷します。

出力とグラフを印刷するには

1. ビューアをアクティブ・ウィンドウにします (ウィンドウ内の任意の場所をクリックします)。
2. メニューから次の項目を選択します。

「ファイル」 > 「印刷」

3. 目的の印刷設定を選択します。
4. 「OK」をクリックして印刷します。

印刷プレビュー

「印刷プレビュー」は、ビューア文書の各ページに印刷される内容を表示します。次のような項目はビューアのコンテンツ・ペインには表示されない場合がありますが、「印刷プレビュー」では表示されるので、ビューア文書を実際に印刷する前に「印刷プレビュー」で確認することをお勧めします。

- 改ページ
- ピボット・テーブルの隠れ層
- 幅広のテーブルでのブレイク
- 各ページに印刷されるヘッダーとフッター

ビューアで出力が現在選択されている場合、プレビューには選択した出力だけが表示されます。すべての出力のプレビューを表示する場合は、ビューアで何も選択されていないことを確認してください。

ページ属性: ヘッダーとフッター

ヘッダーとフッターは、各ページの上部と下部に印刷される情報です。ヘッダーやフッターとして使用したい任意のテキストを入力することができます。また、ダイアログ・ボックスの中央にあるツールバーを使用して、次の項目を挿入することもできます。

- 日付と時刻
- ページ番号
- ビューアのファイル名
- アウトラインの見出しラベル
- ページ・タイトルおよびサブタイトル
- 「デフォルトの作成」は、ここでデフォルト設定として指定された設定を新しいビューア文書に使用します。(注: これによって、「ヘッダー/フッター」タブおよび「オプション」タブ両方の現在の設定がデフォルト設定になります。)
- アウトラインの見出しラベルは、各ページの最初の項目が第 1、第 2、第 3、第 4 のどのレベルのアウトライン見出しであるかを示します。
- ページ・タイトルおよびサブタイトルには、現在のページのタイトルとサブタイトルが印刷されます。これらを作成するには、ビューアの「挿入」メニューの「新規ページ・タイトル」を使用するか、TITLE および SUBTITLE コマンドを使用します。ページ・タイトルまたはサブタイトルを指定していない場合、この設定は無視されます。

注: 新規ページ・タイトルおよびサブタイトルのフォント特性は、「オプション」ダイアログ・ボックス (「編集」メニューで「オプション」を選択) の「ビューア」タブ上で設定します。既存のページ・タイトルおよびサブタイトルのフォント特性は、ビューアでタイトルを編集して変更できます。

ヘッダーとフッターがページ上にどのように印刷されるかを確認するには、「ファイル」メニューの「印刷プレビュー」を選択します。

ページのヘッダーとフッターを挿入するには

1. ビューアをアクティブ・ウィンドウにします (ウィンドウ内の任意の場所をクリックします)。
2. メニューから次の項目を選択します。

「ファイル」 > 「ヘッダーとフッター...」

3. 各ページに表示するヘッダーやフッターを入力します。

ページ属性: オプション

このダイアログ・ボックスでは、印刷するグラフのサイズ、印刷される出力項目間のスペース、およびページの番号付けを制御します。

- 「印刷するグラフ・サイズ」。印刷するグラフのサイズを、定義されたページ・サイズに対して相対的に制御します。グラフの縦横比 (幅と高さの比率) は、印刷するグラフ・サイズにかかわらず一定です。グラフの全体的な印刷サイズは、その高さの両方で制限されます。グラフの外枠がページの左右の境界線に達すると、ページの高さもそれ以上大きくできなくなります。
- 「項目間のスペース」。印刷される項目間のスペースを制御します。各ピボット・テーブル、グラフ、およびテキスト・オブジェクトは個別の項目です。この設定は、ビューアでの項目の表示には影響しません。
- 「開始ページ数」。指定した番号から始めて、順番にページ番号を付けます。
- 「デフォルトの作成」。このオプションは、ここでデフォルト設定として指定された設定を新しいビューアの文書に使用します。(このオプションを使用すると、現在の「ヘッダー/フッター」設定と「オプション」設定がデフォルトになるので注意してください。)

グラフの印刷サイズ、ページの番号付け、および印刷項目間のスペースを変更するには

1. ビューアをアクティブ・ウィンドウにします (ウィンドウ内の任意の場所をクリックします)。
2. メニューから次の項目を選択します。

「ファイル」 > 「ページ属性...」

3. 「オプション」タブをクリックします。
4. 設定を変更し、「OK」をクリックします。

出力の保存

ビューアの内容は保存が可能です。

- 出力オブジェクト (*.cou)。この形式では、グラフ、タブ、注釈などを含め、出力コンテナ全体を保存します。この形式は、IBM SPSS Modeler で開いて表示したり、プロジェクトに追加したり、IBM SPSS Collaboration and Deployment Services Repository を使用して公開および追跡したりすることができます。この形式は IBM SPSS Statistics と互換性がありません。
- ビューア ファイル (*.spv)。「ビューア」ウィンドウにファイルを表示するために使用される形式。IBM SPSS Modeler のモデル ナゲットからこの形式に保存すると、「モデル」タブのビューアの内容のみが保存されます。

Web レポートの保存に関するオプションを制御する場合、および結果を他の形式 (テキスト、Word、Excel など) で保存する場合は、「ファイル」メニューの「エクスポート」を使用します。

ビューア文書を保存するには

1. 「ビューア」ウィンドウのメニューから次の項目を選択します。

「ファイル」 > 「保存」

2. 文書の名前を入力し、「保存」をクリックします。

オプションで、以下のことが可能です。

IBM SPSS Smartreader で編集されないようにファイルをロック。

ビューア文書がロックされている場合、ピボット・テーブルを操作することはできます (行と列をスワップする、表示される層を変更するなど) が、IBM SPSS Smartreader (ビューア文書を処理する別の製品) で出力を編集したり、ビューア文書への変更を保存したりすることはできません。この設定は、IBM SPSS Statistics または IBM SPSS Modeler で開いたビューア文書では無効になります。

パスワードでファイルを暗号化

文書をパスワードで暗号化することによって、ビューア文書に保存されている機密情報を保護することができます。暗号化すると、文書はパスワードを入力しなければ開くことができません。IBM SPSS Smartreader ユーザーも、ファイルを開くにはパスワードを入力する必要があります。

ビューア文書を暗号化するには、次の操作を行います。

- a. 「出力に名前を付けて保存」ダイアログ・ボックスで、「パスワードでファイルを暗号化」を選択します。
- b. 「保存」をクリックします。
- c. 「ファイルの暗号化」ダイアログ・ボックスで、パスワードを入力し、「パスワードの確認」テキスト・ボックスにパスワードを再入力します。パスワードは 10 文字以内で、大文字と小文字が区別されます。

警告: パスワードがわからなくなった場合、復旧することはできません。パスワードをなくした場合、ファイルを開くことはできません。

強固なパスワードの作成

- 使う文字を 8 文字以上にします。
- パスワードに数字、記号、句読点を含めます。
- 「123」や「abc」のような連続する数字や文字、「111aaa」のような繰り返しを避けます。
- 誕生日やニックネームなどの個人情報を含むパスワードは作成しないようにします。
- 定期的にパスワードを変更します。

注: IBM SPSS Collaboration and Deployment Services Repositoryへの暗号化ファイルの保存はサポートされていません。

暗号化されたファイルの変更

- 暗号化されたファイルを開いて変更し、「ファイル」>「保存」を選択すると、変更したファイルは同じパスワードで保存されます。
- 暗号化されたファイルのパスワードを変更するには、ファイルを開き、暗号化の手順を繰り返して、「ファイルの暗号化」ダイアログ・ボックスで異なるパスワードを指定します。
- 暗号化ファイルを暗号化されていない状態で保存するには、ファイルを開き、「ファイル」>「名前を付けて保存」を選択して、「出力に名前を付けて保存」ダイアログ・ボックスで「パスワードでファイルを暗号化」を選択解除します。

注: 暗号化されたデータ ファイルと出力ドキュメントは、バージョン 21 より前のバージョンの IBM SPSS Statistics で開くことはできません。暗号化されたシンタックス ファイルは、バージョン 22 より前のバージョンで開くことはできません。

必要なモデル情報を出力ドキュメントとともに保管 (Store required model information with the output document)

このオプションが適用されるのは、一部のインタラクティブ機能を有効にするための補助情報を必要とするモデル ビューア項目が出力ドキュメント内にある場合のみです。これらのモデル ビューア項目のリストおよび補助情報が必要なインタラクティブ機能を表示するには、「詳細 (More Info)」をクリックします。この情報を出力ドキュメントとともに保管すると、ドキュメントのサイズが大幅に増加する可能性があります。この情報を保管しないことを選択した場合でも、これらの出力項目を開くことはできますが、指定されたインタラクティブ機能は使用できません。

ピボット・テーブル

ピボット・テーブル

多くの結果は、インタラクティブにピボットできるテーブルで表現されます。つまり、行、列、および層を再配置することができます。

ピボット・テーブルの操作

ピボット・テーブルを操作するには、次のようなオプションがあります。

- 行と列の入れ替え
- 行と列の移動
- 多次元レイヤーの作成
- 行と列のグループ化とグループ化の解除
- 行、列、および他の情報の表示と非表示
- 行ラベルと列ラベルの回転
- 項目の定義の検索

ピボット・テーブルのアクティブ化

ピボット・テーブルを操作または変更するには、まずテーブルをアクティブ化する必要があります。テーブルをアクティブ化するには、次の手順に従います。

1. テーブルをダブルクリックします。

または

2. テーブルを右クリックし、ポップアップ・メニューから「内容編集」を選択します。
3. サブメニューから、「ビューア内」または「別ウィンドウ」を選択します。

テーブルのピボット

テーブルには、行、列、および層の 3 つのディメンションがあります。ディメンションには、複数の要素を含める (または、何も含めない) ことができます。ディメンション間またはディメンション内の要素を移動して、テーブル構造を変更できます。要素を移動するには、目的の位置にドラッグ・アンド・ドロップします。

ディメンション内の要素の表示順序の変更

テーブル・ディメンション内の要素 (行、列、または層) の表示順序を変更するには

1. ピボット・トレイがまだオンになっていない場合は、「ピボット テーブル」メニューから次の項目を選択します。

「ピボット」 > 「ピボット トレイ」

2. ピボット・トレイで、ディメンション内の要素をドラッグ・アンド・ドロップします。

ディメンション要素内の行と列の移動

1. ピボット・トレイではなくテーブル自体で、移動する行または列のラベルをクリックします。
2. ラベルを新しい位置にドラッグします。

行と列の入れ替え

行と列を入れ替えるには、ピボット・トレイの使用に代わる簡単な代替方法があります。

1. メニューから次の項目を選択します。

「ピボット」 > 「行と列の入れ替え」

この操作には、すべての行要素を列ディメンションにドラッグし、すべての列要素を行ディメンションにドラッグするのと同じ効果があります。

行または列をグループ化する

1. グループ化する行または列のラベルを選択します (複数のラベルを選択するには、クリックしてドラッグするか、Shift を押しながらクリックします)。
2. メニューから次の項目を選択します。

「編集」 > 「グループ」

グループ・ラベルは自動的に挿入されます。グループ・ラベルをダブルクリックしてラベル・テキストを編集します。

注: 既存のグループに行または列を追加するには、現在グループ内にある項目を一度グループ化解除する必要があります。それから追加の項目を含む新しいグループを作成できます。

行または列のグループ化の解除

グループ化を解除すると、グループ・ラベルが自動的に削除されます。

行ラベルまたは列ラベルの回転

テーブルの最も内側の列ラベルと最も外側の行ラベルについて、垂直表示と水平表示の間でラベルを回転できます。

1. メニューから次の項目を選択します。

「形式」 > 「列内側のラベルを回転」

または

「形式」 > 「行外側のラベルを回転」

回転できるのは、最も内側の列ラベルと最も外側の行ラベルだけです。

行のソート

ピボット・テーブルの行をソートするには、次の手順を実行します。

1. テーブルをアクティブにします。

2. ソートに使用する列のセルを選択します。選択した行のグループだけをソートするには、ソートに使用する列の連続する複数のセルを選択します。
3. メニューから次の項目を選択します。

「編集」 > 「行をソートする」

4. サブメニューから「昇順」または「降順」を選択します。
 - 行ディメンションにグループが含まれている場合、ソートは選択内容を含むグループにのみ影響を与えます。
 - グループの境界を超えてソートすることはできません。
 - 行ディメンションに複数の項目が含まれているテーブルをソートすることはできません。

行と列の挿入

ピボット・テーブルの行または列を挿入するには、次の手順を実行します。

1. テーブルをアクティブにします。
2. テーブルの任意のセルを選択します。
3. メニューから次の項目を選択します。

「前に挿入」

or

「後に挿入」

サブメニューから次の項目を選択します。

「行」

or

「列」

- 新規の行または列の各セルには正符号 (+) が挿入され、新規の行または列が空であるために自動的に非表示になることを防ぎます。
- ディメンションがネストまたは階層化されたテーブルでは、対応する各ディメンション・レベルで列または行が挿入されます。

変数ラベルと値ラベルの表示の制御

変数に記述的な変数ラベルまたは値ラベルが含まれている場合は、ピボット・テーブルの変数名と変数ラベルおよびデータ値と値ラベルの表示を制御できます。

1. ピボット・テーブルをアクティブにします。
2. メニューから次の項目を選択します。

「表示」 > 「変数ラベル」

or

「表示」 > 「値ラベル」

3. サブメニューから次のいずれかのオプションを選択します。
 - 「名前」または「値」。変数名 (または値) だけが表示されます。記述ラベルは表示されません。

- 「ラベル」。記述ラベルだけが表示されます。変数名 (または値) は表示されません。
- 「両方」。名前 (または値) と記述ラベルの両方が表示されます。

出力言語の変更

ピボット・テーブルの出力言語を変更するには、次の手順を実行します。

1. テーブルをアクティブにします。
2. メニューから次の項目を選択します。

「表示」 > 「言語」

3. 使用可能ないずれかの言語を選択します。

言語の変更は、テーブル・タイトル、行ラベルと列ラベル、脚注テキストなど、アプリケーションによって生成されたテキストにだけ影響を与えます。変数名および記述的な変数ラベルと値ラベルは影響を受けません。

大きなテーブルのナビゲート

ナビゲーション・ウィンドウを使用して大きなテーブルをナビゲートするには、次の手順を実行します。

1. テーブルをアクティブにします。
2. メニューから次の項目を選択します。

「表示」 > 「ナビゲーション」

変更を元に戻す

アクティブ化されたピボット・テーブルへの最近の変更またはすべての変更を元に戻すことができます。どちらの操作も、テーブルの最新のアクティブ化以降に行われた変更にのみ適用されます。

最新の変更を元に戻すには、次のように操作します。

1. メニューから次の項目を選択します。

「編集」 > 「元に戻す」

すべての変更を元に戻すには、次のように操作します。

2. メニューから次の項目を選択します。

「編集」 > 「復元」

層の処理

各カテゴリまたはカテゴリの組み合わせに対し、個別の 2 次元テーブルを表示できます。テーブルは、層内で積み上げられ、一番上の層だけが見えていると考えることができます。

層の作成と表示

層を作成するには

1. ピボット・トレイがまだオンになっていない場合は、「ピボット テーブル」メニューから次の項目を選択します。

「ピボット」 > 「ピボット トレイ」

2. 行ディメンションまたは列ディメンションからレイヤー・ディメンションに要素をドラッグします。

レイヤー・ディメンションに要素を移動すると、多次元テーブルが作成されますが、単一の 2 ディメンションの「スライス」だけが表示されます。表示されているテーブルは、最上層のテーブルです。例えば、レイヤー・ディメンションに「はい」 / 「いいえ」 カテゴリー変数がある場合、多次元テーブルには「はい」 カテゴリーと「いいえ」 カテゴリーの 2 つの層があります。

表示される層の変更

1. (ピボット・トレイではなく、ピボット・テーブル自体で) 層のドロップダウン・リストからカテゴリーを選択します。

層カテゴリーに移動

「層カテゴリーに移動」では、ピボット・テーブル内の層を変更することができます。このダイアログ・ボックスは、層が多数ある場合、または選択した層が多数のカテゴリーを持つ場合に、特に便利です。

項目の表示と非表示

次のようなさまざまなタイプのセルを非表示にできます。

- ディメンション・ラベル
- 行内または列内のラベル・セルとデータ・セルを含むカテゴリー
- カテゴリー・ラベル (データ・セルは表示)
- 脚注、タイトル、およびキャプション

テーブルの行と列の非表示

テーブルの非表示の行と列の表示

1. メニューから次の項目を選択します。

「表示」 > 「すべてのカテゴリを表示」

これにより、テーブルの非表示の行と列が表示されます (このテーブルに対して「テーブル プロパティ」で「空白の行と列を隠す」が選択されている場合、完全に空白の行または列は非表示のままになります)。

ディメンション・ラベルの表示と非表示

1. ディメンション・ラベル、またはディメンション内のカテゴリー・ラベルを選択します。
2. 「表示」メニューまたはポップアップ・メニューから、「次元ラベル非表示」または「次元ラベルを表示」を選択します。

テーブル・タイトルの表示と非表示

タイトルを非表示にするには、次のように操作します。

1. ピボット・テーブルをアクティブにします。
2. タイトルを選択します。
3. 「表示」メニューから、「非表示」を選択します。

非表示のタイトルを表示するには、次のように操作します。

4. 「表示」メニューから、「すべて表示」を選択します。

テーブル・ルック

テーブル・ルックは、テーブルの表示を定義する一連のプロパティです。すでに定義されているテーブル・ルックを選択することもできますが、独自のテーブル・ルックを作成することもできます。

- テーブル・ルックを適用する前後に、セル・プロパティを使用して個別のセルまたはセルのグループのセル形式を変更することができます。編集したセルの形式は、新しいテーブル・ルックを適用するときにもそのまま使用されます。
- オプションで、すべてのセルを現在のテーブル・ルックで定義されたセルの形式にリセットすることができます。これにより、編集を加えられたすべてのセルがリセットされます。「テーブル・ルック・ファイル」リストで「表示されているとおり」が選択されている場合、編集されたすべてのセルがテーブルの現在のプロパティにリセットされます。
- テーブル・ルックには、「テーブル プロパティ」ダイアログで定義したテーブル プロパティだけが保存されます。テーブル・ルックには、個別のセルの変更は含まれません。

テーブル・ルックを適用するには

1. ピボット・テーブルをアクティブ化します。
2. メニューから次の項目を選択します。

「形式」 > 「テーブルルック...」

3. ファイルのリストから、テーブル・ルックを選択します。別のディレクトリーからファイルを選択するには、「参照」をクリックします。
4. 「**OK**」をクリックして、選択したピボット・テーブルにテーブル・ルックを適用します。

テーブル・ルックを編集または作成するには

1. 「テーブルルック」ダイアログ・ボックスで、ファイルのリストからテーブル・ルックを選択します。
 2. 「テーブル・ルックの編集」をクリックします。
 3. 対象の属性のテーブル・プロパティを調整し、「**OK**」をクリックします。
 4. 「ルックの保存」をクリックして、編集したテーブル・ルックを保存するか、または「名前を付けて保存」をクリックして、新規のテーブル・ルックとして保存します。
- テーブル・ルックを編集した場合、選択したピボット・テーブルだけが影響を受けます。編集したテーブル・ルックは、テーブルを選択してテーブル・ルックを再適用しない限り、そのテーブル・ルックを使用している他のどのテーブルにも適用されません。
 - テーブル・ルックには、「テーブル プロパティ」ダイアログで定義したテーブル プロパティだけが保存されます。テーブル・ルックには、個別のセルの変更は含まれません。

テーブル・プロパティ

「テーブル プロパティ」を使用すると、テーブルの全般的なプロパティの設定、およびテーブルの各部に対するセル・スタイルの設定を行うことができます。以下を行うことができます。

- 空の行または列の非表示、印刷プロパティの調整などの全般的なプロパティの制御。
- 脚注マーカーの形式と位置の制御。
- データ領域内のセル、行ラベルと列ラベル、およびテーブルの他の領域に対して特定の形式を決定する。
- テーブルの各領域の境界線の幅と色の制御。

ピボット・テーブル・プロパティを変更するには

1. メニューから次の項目を選択します。

「形式」 > 「テーブル プロパティ...」

2. タブを選択します（「一般」、「脚注」、「セル書式」、「罫線」、または「印刷」）。
3. 必要なオプションを選択します。
4. 「OK」または「適用」をクリックします。

新しいプロパティが、選択したピボット・テーブルに適用されます。

テーブルのプロパティ：全般

プロパティの中には、テーブル全体に適用されるものがあります。以下を行うことができます。

- 空の行および列の表示と非表示を切り替える（空の行または列には、データ・セルに何も無い行と列です）。
- 行ラベルの配置を制御します。行ラベルは、左上のコーナーに配置するか、入れ子にすることができます。
- 最大および最小の列幅を制御します（ポイント数で表されます）。

テーブルの全般的なプロパティを変更するには

1. 「全般」タブをクリックします。
2. 必要なオプションを選択します。
3. 「OK」または「適用」をクリックします。

表示する行を設定:

注: この機能は、レガシー・テーブルのみ適用されます。

デフォルトでは、行数が多いテーブルは、100 行ごとに表示されます。テーブルに表示される行数を制御するには、次を実行します。

1. 「行ごとにテーブルを表示」を選択します。
2. 「表示する行を設定」をクリックします。

or

3. アクティブ化されたピボット・テーブルの「表示」メニューから、「行ごとにテーブルを表示」および「表示する行を設定」を選択します。

「表示する行」。一度に表示する最大行数を制御します。ナビゲーション・コントロールを使用して、テーブルのさまざまなセクションに移動できます。最小値は 10 です。デフォルトは 100 です。

「ウィドウ/オフアン許容度」。テーブルの表示ビューを分割するために、テーブルの最も内側の行ディメンションの最大行数を制御します。例えば、最も内側の行ディメンションの各グループに 6 つのカテゴリがある場合、値に 6 を指定すると、どのグループも表示ビューで分割されなくなります。この設定を使用すると、表示されたビューの合計行数が、指定された最大表示行数を超える場合があります。

テーブル・プロパティ：メモ

「テーブル プロパティ」ダイアログの「メモ」タブでは、脚注の書式とテーブルのコメント・テキストを制御します。

「脚注」。脚注マーカーのプロパティには、テキストに関連するスタイルと位置があります。

- 脚注マーカーのスタイルは、数字 (1、2、3、...) または文字 (a、b、c、...) のいずれかです。
- 脚注マーカーは、上付き文字または下付き文字としてテキストに付加することができます。

「コメント テキスト」。それぞれのテーブルにコメント・テキストを追加できます。

- コメント・テキストは、ビューアでテーブルにカーソルを移動したときに、ツールチップに表示されます。
- スクリーン・リーダーは、テーブルにフォーカスがあるときにコメント・テキストを読み上げます。
- ビューアのツールチップには、コメントの先頭から 200 文字だけが表示されますが、スクリーン・リーダーはテキスト全体を読み上げます。
- 出力を HTML または Web レポートにエクスポートすると、コメント・テキストは alt (代替) テキストとして使用されます。

テーブルのプロパティ: セルの書式

書式設定を行う場合、テーブルはタイトル、層、コーナー・ラベル、行ラベル、列ラベル、データ、キャプション、脚注の各領域に分類されます。テーブルの各領域に関連付けられているセル書式を変更できます。セル書式には、テキストの特性 (フォント、大きさ、色、スタイルなど)、水平方向と垂直方向の位置合わせ、背景の色、セル内の余白などがあります。

セル書式は、領域 (情報のカテゴリー) に適用されます。領域は個々のセルの特性を表すものではありません。この区別は、テーブルを変更するときに重要になります。

例えば、

- 列ラベルのセル書式として太字フォントを指定すると、現在列ディメンションにどのような情報が表示されているかに関係なく、列ラベルは必ず太字で表示されます。項目を列ディメンションから他のディメンションに移動した場合、列ラベルの太字特性は保持されません。
- アクティブなピボット・テーブルでセルを強調表示にしてからツールバーの「太字」ボタンをクリックすることにより列ラベルを太字にした場合は、セルの内容をどのディメンションに移動したとしてもセルの内容は太字のまま保持されますが、列ディメンションに移動された他の項目については、列ラベルに太字特性が保持されることはありません。

セル書式を変更するには

1. 「セル書式」タブをクリックします。
2. ドロップダウン・リストから領域を選択するか、またはサンプル内の領域をクリックします。
3. その領域の特性を選択します。選択した内容がサンプルに反映されます。
4. 「OK」または「適用」をクリックします。

行の色の変更

別の背景色やテキスト色を適用して、テーブルのデータ領域内の行を変更するには

1. 「領域」ドロップダウン・リストから「データ」を選択します。
2. 「背景色」グループで、「行の色を変更」を選択します (チェック・マークを付けます)。
3. 行の背景とテキストの変更に使用する色を選択します。

変更した行の色は、テーブルのデータ領域にのみ有効です。行や列のラベル領域には影響しません。

テーブル・プロパティ：罫線

テーブル内の罫線の各位置について、線種と色を選択できます。スタイルに「なし」を選択した場合、選択した場所には線が表示されません。

テーブルの罫線を変更するには

1. 「罫線」タブをクリックします。
2. リストで罫線の名前をクリックするか、または「サンプル」領域で線をクリックして、罫線の位置を選択します。
3. 線種を選択するか、または「なし」を選択します。
4. 色を選択します。
5. 「OK」または「適用」をクリックします。

テーブルのプロパティ：印刷

ピボット・テーブルの印刷に関する、次のプロパティを制御できます。

- すべての層またはテーブルの最上層だけを印刷する。また、別々のページに各層を印刷することもできます。
- 印刷するページに合うようにテーブルを水平または垂直方向に縮小できます。
- 定義されたページ・サイズに対してテーブルが広すぎるか長すぎる場合、テーブルの印刷部分に取り入れられる行と列の最小数である改ページ後の行や列数をコントロールできます。

注: その他の出力がテーブルの上にあるため、テーブルが長過ぎて現在の行に収まりきれないが、定義されたページの長さには収まる場合、ウィドウ/オーファン設定に関わらず、テーブルは自動的に新規ページに印刷されます。

- 1 ページに収まりきれないテーブルにも継続テキストを付けることができます。各ページの下部と各ページの上部に継続テキストを表示することができます。どちらも選択されなかった場合には、継続テキストは表示されません。

ピボット・テーブルの印刷プロパティを制御するには

1. 「印刷」タブをクリックします。
2. 必要な印刷オプションを選択します。
3. 「OK」または「適用」をクリックします。

セル・プロパティ

セル・プロパティは、選択したセルに適用されます。フォント、値形式、位置合わせ、余白、および色を変更することができます。セル・プロパティはテーブル・プロパティより優先されるので、テーブル・プロパティを変更しても、個別に適用したセル・プロパティは変更されません。

セル・プロパティを変更するには

1. テーブルのセルを選択します。
2. 「形式」メニューまたはポップアップ・メニューから「セル プロパティ」を選択します。

フォントと背景

「フォントと背景色」タブでは、テーブル内の選択したセルのフォントのスタイルと色、および背景色を制御します。

形式値

「形式値」タブでは、選択したセルの値の形式を制御します。数値、日付、時刻、通貨の形式を選択したり、表示される小数の桁数を調整したりできます。

位置合わせと余白

「位置合わせおよび余白」タブでは、選択したセルの水平および垂直方向の位置合わせと、上下左右の余白を制御します。「混合」の水平位置合わせでは、セルの型に応じて内容が位置合わせされます。例えば、日付は右揃えにされ、テキスト値は左揃えにされます。

脚注とキャプション

テーブルには、脚注およびキャプションを追加できます。脚注またはキャプションの非表示、脚注マーカーの変更、および脚注の再番号付けを行うこともできます。

脚注とキャプションの追加

テーブルにキャプションを追加するには、次の手順を実行します。

1. 「挿入」メニューから、「解説」を選択します。

脚注は、テーブル内のどの項目にも付加することができます。脚注を追加するには、次の手順を実行します。

1. アクティブ化されたピボット・テーブル内で、タイトル、セル、またはキャプションをクリックします。
2. 「挿入」メニューから、「脚注」を選択します。
3. 指定した領域に脚注テキストを挿入します。

キャプションを表示するか、または非表示にするには

キャプションを非表示にするには、次のように操作します。

1. キャプションを選択します。
2. 「表示」メニューから、「非表示」を選択します。

非表示のキャプションを表示するには、次のように操作します。

1. 「表示」メニューから、「すべて表示」を選択します。

テーブルの脚注を表示するか、または非表示にするには

脚注を非表示にするには、次のように操作します。

1. 脚注参照を含むセルを右クリックし、ポップアップ・メニューから「脚注を隠す」を選択します。

または

2. テーブルの脚注領域で脚注を選択し、ポップアップ・メニューから「非表示」を選択します。

注: レガシー テーブルの場合、テーブルの脚注領域を選択し、ポップアップ・メニューから「脚注を編集」を選択して、非表示にする脚注の「表示」プロパティを選択解除 (クリア) します。

セルに複数の脚注が含まれている場合は、後者の方法を使用して脚注を選択的に非表示にします。

テーブルのすべての脚注を非表示にするには、次のように操作します。

1. テーブルの脚注領域ですべての脚注を選択 (クリック・アンド・ドラッグを使用するか、または Shift キーを押しながらクリックして脚注を選択) し、「表示」メニューから「非表示」を選択します。

非表示の脚注を表示するには、次のように操作します。

1. 「表示」メニューから、「すべての脚注を表示」を選択します。

脚注マーカー

「脚注マーカー」では、脚注をマークするために使用する文字を変更します。デフォルトでは、標準の脚注マーカーは、テーブル・プロパティの設定に応じた、連続する文字または数字です。特殊なマーカーを割り当てることもできます。脚注を再番号付けしたり、標準マーカーの数値と文字を切り替えたりしても、特殊マーカーは影響を受けません。標準マーカーの数値または文字の表示および脚注マーカーの下付き文字または上付き文字は、「テーブル プロパティ」ダイアログの「脚注」タブによって制御されます。

脚注マーカーを変更するには

1. 脚注を選択します。
2. 「形式」メニューから、「脚注マーカー」を選択します。

特殊マーカーは 2 文字までに制限されています。特殊マーカーを持つ脚注は、テーブルの脚注エリア内で、連続文字や通し番号の数字での脚注より上位になるため、特殊マーカーに変更すると脚注リストが並べ替えられます。

脚注の再番号付け

行、列、および層を切り替えてテーブルをピボットすると、脚注の順序が変化する場合があります。脚注の再番号付けを行うには、次の手順に従います。

1. 「形式」メニューから、「脚注番号のつけ直し」を選択します。

レガシー・テーブルの脚注の編集

レガシー・テーブルの場合、「脚注を編集」ダイアログを使用して、脚注のテキストおよびフォント設定の入力や変更、脚注マーカーの変更、および脚注を非表示にしたり削除したりすることができます。

レガシー・テーブルに新しい脚注を挿入すると、「脚注を編集」ダイアログが自動的に開きます。「脚注を編集」ダイアログを使用して、新しい脚注を作成せずに既存の脚注を編集するには、次の手順を実行します。

「マーカー」。デフォルトでは、標準の脚注マーカーは、テーブル・プロパティの設定に応じた、連続する文字または数字です。特殊なマーカーを割り当てるには、「マーカー」列に新しいマーカー値を入力します。脚注を再番号付けしたり、標準マーカーの数値と文字を切り替えたりしても、特殊マーカーは影響を受けません。標準マーカーの数値または文字の表示および脚注マーカーの下付き文字または上付き文字は、「テーブル プロパティ」ダイアログの「脚注」タブによって制御されます。詳しくは、トピック 95 ページの『テーブル・プロパティ: メモ』を参照してください。

特殊マーカーを標準マーカーに戻すには、「脚注を編集」ダイアログでマーカーを右クリックし、ポップアップ・メニューで「脚注マーカー」を選択し、「脚注マーカー」ダイアログ・ボックスで「標準マーカー」を選択します。

「脚注」。脚注の内容。現在のフォントおよび背景の設定が表示に反映されます。フォントの設定は、「形式」サブダイアログを使用して脚注ごとに変更できます。詳しくは、トピック 100 ページの『脚注のフォントと色の設定』を参照してください。すべての脚注には 1 つの背景色が適用されます。この背景色は、「セル プロパティ」ダイアログの「フォントと背景」タブで変更できます。詳しくは、トピック 97 ページの『フォントと背景』を参照してください。

「表示」。デフォルトでは、すべての脚注が表示されます。脚注を非表示にするには、「表示」チェック・ボックスを選択解除します (チェック・マークを外します)。

脚注のフォントと色の設定: レガシー・テーブルの場合、「形式」ダイアログを使用して、1 つ以上の選択した脚注のフォント・ファミリー、スタイル、サイズ、および色を変更できます。

1. 「脚注を編集」ダイアログで、「脚注」グリッドの脚注を 1 つ以上選択 (クリック) します。
2. 「形式」ボタンをクリックします。

選択したフォント・ファミリー、スタイル、サイズ、および色が選択したすべての脚注に適用されます。

背景色、配置、余白は、「セル プロパティ」ダイアログで設定でき、すべての脚注に適用されます。これらの設定は脚注ごとに変更できません。詳しくは、トピック 97 ページの『フォントと背景』を参照してください。

データ・セルの幅

「データ・セルの幅の設定」を使用して、すべてのデータ・セルを同じ幅に設定します。

すべてのデータ・セルの幅を設定するには

1. メニューから次の項目を選択します。
「形式」 > 「データ・セルの幅の設定...」
2. セル幅の値を入力します。

列の幅の変更

1. 列罫線をクリックしてドラッグします。

ピボット・テーブルの隠れた罫線を表示する

見えている罫線が少ないテーブルでは、隠れた罫線を表示することができます。それにより、列幅の変更のような作業が簡単になります。

1. 「ビュー」メニューから「グリッド線」を選択します。

ピボット・テーブルの行、列、およびセルの選択

行または列全体、またはデータ・セルおよびラベル・セルの指定したグループを選択できます。

複数のセルを選択するには、次のように操作します。

「選択」 > 「データとラベルのセル」

ピボット・テーブルの印刷

印刷したピボット・テーブルの外観に影響を与える可能性がある要素がいくつかあります。これらの要素は、ピボット・テーブルの属性を変更することで制御できます。

- 多次元ピボット・テーブル (複数の層のあるテーブル) では、すべての層を印刷するか、または最上層 (表示されている層) だけを印刷するかを選択できます。詳しくは、トピック 97 ページの『テーブルのプロパティ: 印刷』を参照してください。
- 縦長または幅広のピボット・テーブルでは、ページに合わせてテーブルを自動的にサイズ変更したり、テーブル・ブレイクの位置および改ページの位置を制御したりできます。詳しくは、トピック 97 ページの『テーブルのプロパティ: 印刷』を参照してください。

印刷後のピボット・テーブルの外観を確認するには、「ファイル」メニューの「印刷プレビュー」を使用します。

幅広または縦長のテーブルのテーブル・ブレイクの制御

広すぎるか、または長すぎるために、定義したページ・サイズ内に印刷できないピボット・テーブルは、自動的に分割されて複数のセクションに印刷されます。以下を行うことができます。

- 大きなテーブルを分割する行と列の位置を制御します。
- テーブルを分割するときにまとめる必要がある行と列を指定します。
- 定義したページ・サイズに合わせて、大きなテーブルをスケール変更します。

印刷するピボット テーブルの行ブレイクおよび列ブレイクを指定するには、以下のようになります。

1. ピボット・テーブルをアクティブにします。
2. ブレイクを挿入する場所の左にある列のセルをクリックするか、またはブレイクを挿入する行の前の行のセルをクリックします。
3. メニューから次の項目を選択します。

「形式」 > 「分割点」 > 「垂直方向の分割点」

または

「形式」 > 「分割点」 > 「水平方向の分割点」

1. ピボット・テーブルをアクティブにします。
2. ブレイクを挿入する場所の左にある列のセルをクリックするか、またはブレイクを挿入する行の前の行のセルをクリックします。
3. メニューから次の項目を選択します。

「形式」 > 「分割点」 > 「垂直方向の分割点」

または

「形式」 > 「分割点」 > 「水平方向の分割点」

まとめる行または列を指定するには、以下のようになります。

1. まとめる行または列のラベルを選択します。クリックしてドラッグするか、または Shift キーを押しながらクリックして、複数の行ラベルまたは列ラベルを選択します。
2. メニューから次の項目を選択します。

「形式」 > 「分割点」 > 「まとめる」

ブレイクポイントおよびまとめられたグループを表示するには、以下のようになります。

1. メニューから次の項目を選択します。

「形式」 > 「分割点」 > 「分割点の表示」

ブレイクポイントは、垂直方向または水平方向の線として表示されます。まとめられたグループは、濃い罫線で囲まれた長方形の領域としてグレー表示されます。

注: レガシー・テーブルでは、ブレイクポイントおよびまとめられたグループの表示はサポートされていません。

ブレイクポイントおよびまとめられたグループをクリアするには

ブレイクポイントをクリアするには、次のように操作します。

1. 垂直方向のブレイクポイントの左にある列のセルをクリックするか、または水平方向のブレイクポイントの上にある行のセルをクリックします。
2. メニューから次の項目を選択します。

「形式」 > 「分割点」 > 「分割点またはグループをクリア」

まとめられたグループをクリアするには、次のように操作します。

3. グループを指定する列ラベルまたは行ラベルを選択します。
4. メニューから次の項目を選択します。

「形式」 > 「分割点」 > 「分割点またはグループをクリア」

行または列をピボットするか、または並べ替えると、すべてのブレイクポイントおよびまとめられたグループが自動的にクリアされます。この動作は、レガシー・テーブルには適用されません。

ピボット・テーブルからのグラフの作成

1. ピボット・テーブルをダブルクリックして、アクティブにします。
2. グラフに表示する行、列、またはセルを選択します。
3. 選択した領域の任意の場所を右クリックします。
4. ポップアップ・メニューから「グラフの作成」を選択し、グラフ・タイプを選択します。

レガシー・テーブル

テーブルをレガシー・テーブル (リリース 19 ではフル機能のテーブルと見なされる) として表示するように選択できます。これは、リリース 20 より前の IBM SPSS Statistics リリースと完全に互換性があります。レガシー・テーブルは、表示に時間がかかる場合があるため、リリース 20 より前のリリースとの互換性が必要な場合のみ推奨されます。レガシー・テーブルの作成方法については、103 ページの『ピボット・テーブル・オプション』を参照してください。

オプション

オプション

オプションはさまざまな設定を制御します。

オプション設定を変更するには

1. メニューから次の項目を選択します。

「編集」 > 「オプション...」

2. 変更する設定のタブをクリックします。
3. 設定を変更します。
4. 「OK」または「適用」をクリックします。

全般オプション

最大スレッド数

マルチスレッド プロシージャで結果を計算する際に使用するスレッドの数。「自動」設定は、使用可能なプロセッシング コアの数に基づきます。マルチスレッド プロシージャの実行中に他のアプリケーションで追加処理リソースを使用できるようにするには、より小さい値を指定してください。このオプションは、分散分析モードでは無効です。

出力

「10 進値の先行ゼロを表示する」。小数部分のみで構成される数値の先行ゼロを表示します。例えば、先行ゼロを表示する場合、値 .123 は 0.123 として表示されます。この設定は、通貨およびパーセント形式の数値には適用されません。固定 ASCII ファイル (*.dat) を除き、データが外部ファイルに保存されるときに先行ゼロは含まれません。

「測定システム」。印刷する際のピボット・テーブルのセルの余白、セル幅、テーブル間のスペースなどの属性を指定するために使用する測定システム (ポイント、インチ、またはセンチメートル)。

ビューアーのオプション

ビューアーの出力表示オプションは、設定を変更した後で作成した新しい出力にのみ影響します。ビューアーに既に表示されている出力は、これらの設定を変更しても影響されません。

「初期出力状態」。プロシージャを実行するたびにどの項目を自動的に表示または非表示にするか、また項目の初期の位置合わせの方法を制御します。ログ、警告、メモ、タイトル、ピボット・テーブル、グラフ、ツリー図、およびテキスト出力の表示を制御できます。また、ログ内のコマンドの表示のオンとオフを切り替えることもできます。ログからコマンド・シンタックスをコピーして、シンタックス・ファイルに保存できます。

注: ビューアーには、すべての出力項目が左揃えで表示されます。印刷出力の配置だけが、位置調整設定の影響を受けます。中央揃えと右揃えの項目は、小さい記号で識別されます。

「タイトル」。新しい出力タイトルのフォント・スタイル、サイズ、および色を制御します。

「ページ・タイトル」。新しいページ・タイトル、TITLE および SUBTITLE コマンド・シンタックスによって生成されたページ・タイトル、または「挿入」メニューの「新しいページ表題」で作成されたページ・タイトルのフォント・スタイル、サイズ、および色を制御します。

「テキスト出力」。テキスト出力に使用されるフォント。テキスト出力は、モノスペース (固定ピッチ) フォントでを使用することを意図しています。プロポーショナル・フォントを選択すると、表形式出力は正しく位置合わせされません。

「デフォルトのページ設定」。印刷の方向と余白に関するデフォルトのオプションを制御します。

ピボット・テーブル・オプション

「ピボット テーブル」オプションは、ピボット・テーブルの表示に関するさまざまなオプションを設定します。

テーブル・ルック(T)

ファイルのリストからテーブル・ルックを選択し、「OK」または「適用」をクリックします。IBM SPSS Statistics で提供されるテーブル・ルックの 1 つを使用することも、ピボット・テーブル・エディター (「形式」メニューで「テーブルルック」を選択) で独自のテーブル・ルックを作成することもできます。

- 「参照」。別のディレクトリーからテーブル・ルックを選択できます。
- 「デフォルトのディレクトリーに設定」。デフォルトのテーブル・ルック・ディレクトリーを変更できます。「参照」を使用して使用するディレクトリーに移動し、そのディレクトリーでテーブル・ルックを選択して、「デフォルトのディレクトリーに設定」を選択します。

注: 以前のバージョンの IBM SPSS Statistics で作成されたテーブル・ルックは、バージョン 16.0 以降では使用できません。

列幅

これらのオプションは、ピボット・テーブルでの列幅の自動調整を制御します。

- 「ラベルのみに合わせて調整」。列幅を列ラベルの幅に調整します。これを選択するとよりコンパクトなテーブルが生成されますが、ラベルより幅の広いデータ値は切り捨てられます。
- 「すべてのテーブルについてラベルとデータに合わせて調整」。列ラベルか最大データ値のいずれか大きい方に合わせて列幅を調整します。これを選択すると幅の広いテーブルが作成されますが、すべての値が確実に表示されるようになります。

デフォルトの編集モード

このオプションは、ピボット・テーブルを「ビューア」ウィンドウでアクティブ化するか、別のウィンドウでアクティブするかを制御します。デフォルトでは、ピボット・テーブルをダブルクリックすると、非常に大きなテーブルを除き、すべてのテーブルが「ビューア」ウィンドウで開かれます。ピボット・テーブルを別のウィンドウでアクティブ化することを選択したり、サイズ設定を選択して、それより小さなピボット・テーブルは「ビューア」ウィンドウで、大きなピボット・テーブルは別のウィンドウで開くように設定することもできます。

幅の広いテーブルをリッチ・テキスト形式でクリップボードにコピー

ピボット・テーブルを Word/RTF 形式で貼り付けると、文書の幅より広すぎるテーブルは、折り返されるか、文書の幅に合わせて縮小されるか、または変更されないままになります。

出力オプション

出力オプションは、さまざまな出力オプションのデフォルト設定を制御します。

「スクリーン・リーダーのアクセシビリティ」。ピボット・テーブルの行および列ラベルをスクリーン・リーダーで読み上げる方法を制御します。データ・セルごとに行ラベルおよび列ラベルの全体を読み上げるか、テーブル内のデータ・セルを移動したときに変更されたラベルのみを読み上げることができます。

第 7 章 欠損値の処理

欠損値の概要

データ・マイニングのデータの前処理フェーズでは、データの欠損値の置換が必要な場合がしばしばあります。欠損値とは、データ・セット中の不明な値、収集されていない値、または誤って入力された値です。通常、そのような値はフィールドに対して無効になります。例えば、「性別」フィールドの値は M または F でなくてはなりません。このフィールドに Y や Z などの値があった場合は、値は無効で空白として解釈しても問題ありません。同様に、「年齢」フィールドに負の値がある場合も、意味がないため空白として扱われます。このような明白に誤った値が故意に入力されたり、質問に対して回答したくないことを示すために、フィールドを空欄のまま放置することはよくあります。時には、年齢を入力していないなどの未回答が、特定の結果を予測するための因子であるかかどうかを判断するために、空白の意味を詳しく調べる必要があることもあります。

一部のモデル作成手法では、欠損値がより効果的に処理されます。例えば、C5.0、および Apriori は、データ型ノードで明示的に「欠損値」と宣言された値を効果的に処理しています。他のモデル作成技法では、欠損値を処理できないため、学習に時間がかかったり、モデルの精度が低下してしまいます。

IBM SPSS Modeler の欠損値には、次の 2 種類があります。

- **ヌル値またはシステム欠損値**：これらの値は、データベースまたはソース・ファイルに空白のまま残された文字列以外の値であり、入力ノードまたはデータ型ノードで特に「欠損値」として定義されていません。システム欠損値は `$null$` 値として表示されます。空の文字列が特定のデータベースでヌルとして処理される場合でも、IBM SPSS Modeler では空の文字列をヌルとは見なさないことに注意してください。
- **空文字列と空白文字**：空文字の値と空白文字（表示されない文字による文字列）をヌル値の重複レコードとして処理します。空白の文字列は、ほとんどの目的に対してホワイトスペースとして扱われます。例えば、オプションを選択してソースまたはデータ型ノードで空白文字を空白値として扱う場合、この設定は空白の文字列も同様に適用します。
- **空白値またはユーザー定義の欠損値**：これらは、入力ノードまたはデータ型ノード内で欠損値として明示的に定義されている `unknown`、`99`、`-1` などの値です。オプションでヌルと空白文字を「空白」として処理することもできます。そうすることによって、特別な処理のためにフラグを付けたり、ほとんどの計算から除外することができるようになります。例えば、`@BLANK` 関数を使用して、これらの値を他の欠損値と共に空白値として処理することができます。

混在データの読み込み：数値ストレージ（整数、実数、時間、タイムスタンプ、または日付のいずれか）をフィールドに読み込む場合は、数値以外のすべての値がヌル またはシステム欠損値 に設定されることに注意してください。これは、他のアプリケーションと異なり、フィールド内にストレージの混在を許さないためです。これを回避するためには、必要に応じて入力ノードまたは外部アプリケーション内の記憶域タイプを変更し、混在データを含むフィールドを文字列として読み込む必要があります。

Oracle からの空の文字列の読み取り：Oracle データベースとの間で読み書きするときには、Oracle が IBM SPSS Modeler やその他のほとんどのデータベースとは異なり、空の文字列値をヌル値と同様に処理および格納することに注意してください。つまり、Oracle データベースから抽出されたデータは、同じデータがファイルやその他のデータベースから抽出された場合とは異なって動作し、また異なる結果が返ることがあります。

欠損値の処理

業務やドメインの知識の観点から、欠損値をどのように取り扱うかを決定する必要があります。多くの場合、学習時間を短縮し、精度を向上させるには、データセットから空白を削除する必要があります。その一方で、空白値から、新しいビジネス・チャンスを見出したり、新規の洞察を加えられることもあります。最適テクニックを選択するときに、次に示すデータの側面を検討する必要があります。

- データ・セットのサイズ
- 空白を含むフィールドの数
- 失われた情報の量

一般的に、アプローチは次の 2 通りあります。

- 欠損値を含むフィールドまたはレコードを除外する。
- 様々な方法で欠損値を代入、置換または強制する。

これら 2 つのアプローチは、データ検査ノードを使用し、大半を自動化することができます。例えば、フィルター・ノードを生成し、モデル作成で有用な欠損値をあまりに多く含むフィールドを除外することができます。また、スーパー・ノードを作成して維持するいずれかまたはすべてのフィールドに対して欠損値を代入することができます。このため検査によって実際にデータの現在の状態を評価するだけでなく、評価に基づいて行動を起こすことができます。

欠損値を含むレコードの処理

欠損値の大部分が少数のレコードに集中している場合、それらのレコードだけを除外することができます。例えば、銀行にはローン顧客に関する詳細で完全なレコードがあるのが普通です。しかし、行員用のローンの承認では空白に関する制限が緩和されており、行員ローン用に収集されたデータにいくつかの空白のフィールドが含まれていることがあります。このような場合のこれらの欠損値の処理には、2 つのオプションがあります。

- 条件抽出ノードを使用して、スタッフレコードを削除できます。
- データ・セットが大きい場合、空白を含むレコードを破棄することができます。

欠損値を含むフィールドの処理

欠損値の大部分が少数のフィールドに集中している場合は、レコードレベルではなくフィールドレベルで欠損値に対処することができます。この方法では、欠損値の処理方法を決める前に、特定のフィールドの相対重要度を調べることがもできます。フィールドがモデル作成において重要ではない場合、欠損値の過多に関わらず、そのフィールドを保持する価値はないと思われます。

例えば、マーケットリサーチ会社がデータの収集に 50 の質問から成る一般的なアンケートを使用する場合を考えてみましょう。そのうちの 2 つの質問は年齢と政治信条に関するもので、これは多くの人が提供をためらう情報です。この場合、年齢と政治信条には多くの欠損値が存在することになります。

フィールドの尺度

どの方法を使用するかを決める際には、欠損値のあるフィールドの尺度も考慮する必要があります。

数値型フィールド: 連続型などの数値型フィールドの場合、多くのモデルは数値型フィールドに空白が含まれていると機能しないため、常に数値以外のすべての値を削除してからモデルを作成する必要があります。

カテゴリー・フィールド: 名義型 やフラグ型 などのカテゴリー・フィールドの場合は、欠損値の変更は必須ではありませんが、変更するとモデルの精度が向上します。例えば、フィールド「性別」を使用するモデルは、Y や Z のような意味のない値があっても機能しますが、M や F 以外の値をすべて削除すればモデルの精度が向上します。

スクリーニングまたは削除フィールド

あまりに多くの欠損値を含むフィールドのスクリーニングには、次のようなオプションがあります。

- データ検査ノードを使用し、品質に基づいてフィールドをフィルタリングすることができます。
- フィールド選択ノードを使用して、指定した欠損値のパーセンテージ以上のフィールドをスクリーニングし、指定した目標値を基準にした重要度に基づいてフィールドをランク付けします。
- フィールドを削除する代わりに、データ型ノードを使用してフィールドの役割をなしに設定することもできます。これにより、フィールドはデータ・セット内に維持されますが、モデル作成プロセスからは除外されます。

システム欠損値を含むレコードの処理

システム欠損値とは

システム欠損値は、不明または適用できないデータ値を表します。このような値はデータベースではしばしばヌル 値と呼ばれます。

システム欠損値は空白値とは異なります。空白値は通常、データ型ノード内で特定の値または値範囲として定義され、ユーザー定義の欠損と見なすことができます。空白値はモデル作成のコンテキストに応じてさまざまに処理されます。

システム欠損値の構築

システム欠損値はデータ・ソースから読み取られたデータ内に存在する場合があります (例えば、データベース表にヌル 値が含まれる場合)。

システム欠損値は、式で値 **undef** を使用することによって構築できます。例えば、以下の CLEM 式は年齢が 30 以下の場合には年齢を返し、30 を超える場合は欠損値を返します。

```
if Age > 30 then undef else Age endif
```

欠損値は、外部結合が実行された場合、数値がゼロで除算された場合、負数の平方根が計算された場合などにも作成されます。

システム欠損値の表示

システム欠損値は、テーブルおよびその他の出力に **\$null\$** と表示されます。

システム欠損値のテスト

引数値がシステム欠損値である場合に **true** を返すには、下の例のように特殊関数 **@NULL** を使用します。

```
if @NULL(MyFieldName) then 'It is null' else 'It is not null' endif
```

関数に渡されるシステム欠損値

関数に渡されるシステム欠損値は、通常、欠損値を出力に伝達します。例えば、フィールド **f1** の値が特定の列のシステム欠損値である場合、式 **log(f1)** もその行のシステム欠損値に評価されます。**@NULL** 関数は例外です。

算術演算子を含む式でのシステム欠損値

システム欠損値を含む値に算術演算子を適用すると、その結果はシステム欠損値となります。例えば、フィールド `f1` の値が特定の列のシステム欠損値である場合、式 `f1 + 10` もその行のシステム欠損値に評価されます。

論理演算子を含む式でのシステム欠損値

論理演算子を含む式でシステム欠損値を処理する場合、3 値論理 (真、偽、および欠損値) のルールが適用され、真理値表に記述できます。一般的な論理演算子、`not`、`and`、および `or` の真理値表を以下の表に示します。

表 4. NOT の真理値表

オペランド	NOT オペランド
真	偽
偽	真
欠損値	欠損値

表 5. AND の真理値表

オペランド 1	オペランド 2	オペランド 1 AND オペランド 2
真	真	真
真	偽	偽
真	欠損値	欠損値
偽	真	偽
偽	偽	偽
偽	欠損値	偽
欠損値	真	欠損値
欠損値	偽	偽
欠損値	欠損値	欠損値

表 6. OR の真理値表

オペランド 1	オペランド 2	オペランド 1 OR オペランド 2
真	真	真
真	偽	真
真	欠損値	真
偽	真	真
偽	偽	偽
偽	欠損値	欠損値
欠損値	真	真
欠損値	偽	欠損値
欠損値	欠損値	欠損値

比較演算子を含む式でのシステム欠損値

システム欠損値と非システム欠損値とを比較する場合、結果は、真 (true) または偽 (false) の結果ではなく、システム欠損値と評価されます。システム欠損値同士は互いに比較できます。2 つのシステム欠損値は等しいと見なされます。

if/then/else/endif 式でのシステム欠損値

条件式を使用する際に、条件式がシステム欠損値を返す場合は、else 節からの値は条件式から返されません。

条件抽出ノードでのシステム欠損値

特定のレコードで、選択式が欠損値に評価される場合、レコードは条件抽出ノードから出力されません (この処理は包含 (Include) と破棄 (Discard) の両方のモードに適用されます)。

レコード結合ノードでのシステム欠損値

キーを使用して結合する場合、キー・フィールドにシステム欠損値を持つレコードは結合されません。

集計でのシステム欠損値

列のデータを集計する場合、欠損値は計算に含まれません。例えば、3 つの値 { 1、2、および undef } がある列では、列の値の和は 3 と計算され、平均値は 1.5 と計算されます。

欠損値の代入または置換

このようにわずかな欠損値が存在しているような場合は、値を挿入して空白と置換すると効果的な場合があります。この処理はデータ検査レポートから行うことができ、それにより必要に応じて特定のフィールドに対しオプションを指定し、様々な方法を使用して値を代入するスーパー・ノードを作成することができます。これは最もフレキシブルな方法で、単一ノード内にある多くのフィールドに対し処理を指定することができます。

次の方法で、欠損値の代入ができます。

固定: 固定値で置き換えます (指定のフィールド計測、範囲の中間または一定数)。

無作為: 正常または均一分布に基づいたランダム値で置き換えます。

式: ユーザー設定の式を指定することができます。例えば、値をグローバル値設定ノードで作成されたグローバル変数と置き換えることができます。

アルゴリズム: C&RT アルゴリズムの基づいたモデルによって予測された値で置き換えます。この方法で代入された各フィールドに対し、空白値やヌル値をモデルで予測された値と置き換える置換ノードとともに、個別の C&RT モデルが作成されます。フィルター・ノードを使用して、モデルが生成した予測値を削除します。

代わりに、特定のフィールドに対して強制するには、データ型ノードを使用して、フィールドのデータ型が正しい値だけを利用するために、空白値の置換が必要なフィールドの「検査」列を「強制」に設定します。

欠損値用 CLEM 関数

欠損値の処理には、さまざまな関数を利用できます。条件抽出ノードや置換ノードで欠損値を破棄したり、値を代入するためによく用いられる関数を次に示します。

- `count_nulls(LIST)`
- `@BLANK(FIELD)`
- `@NULL(FIELD)`
- `undef`

@FIELD 関数とともに @ 関数を使用することで、1 つまたは複数のフィールドに空白やヌル値が存在するかどうかを判断することができます。空白またはヌル値があるフィールドに単にフラグを設定することも、それを他の演算子を使用して適切な値に置換することもできます。

次に示すように、フィールドのリスト全般にわたってヌルをカウントできます。

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

フィールドのリストを入力として受け入れる関数を使用する場合は、特殊関数 @FIELDS_BETWEEN および @FIELDS_MATCHING を使用できます。以下に例を示します。

```
count_nulls(@FIELDS_MATCHING('card*'))
```

undef 関数を使用して、フィールドを **\$null\$** として表示されるシステム欠損値で埋めることができます。例えば、任意の数値を置換するには、次のような条件ステートメントを使用することができます。

```
if not(Age > 17) or not(Age < 66) then undef else Age endif
```

このステートメントでは、範囲内にはない値を、**\$null\$** として表示されるシステム欠損値で置き換えます。not() 関数を使うと、負の数値も含めた他のすべての数値を指定することもできます。詳しくは、トピック 159 ページの『空白値とヌル値処理関数』を参照してください。

レコードの破棄に関する注意

条件抽出ノードを使用してレコードを破棄する場合、シンタックスでは 3 値論理を用いて、選択したステートメントのヌル値を自動的に含めることに注意してください。選択した式からヌル値 (システム欠損値) を除外するには、式で `and not` を使用して、これを明示的に指定する必要があります。例えば、処方薬の種類が薬品 C であるすべてのレコードを選択し、含めるには、次のステートメントを使用します。

```
Drug = 'drugC' and not(@NULL(Drug))
```

前のバージョンでは、このような状況の場合ヌル値を除外していました。

第 8 章 CLEM 式の作成

CLEM について

Control Language for Expression Manipulation (CLEM) は、IBM SPSS Modeler ストリームで利用するデータの分析と処理を行うための強力な言語です。CLEM を使用すれば、経費と収入データから利益を算出するような簡単な操作から、Web ログ・データを有益な情報を含む一連のフィールドやレコードに変換するような複雑な操作まで、さまざまなストリーム操作を行うことができます。

CLEM は、IBM SPSS Modeler で次の目的のために使用されます。

- レコード・フィールドに対する条件の比較と評価。
- 新規フィールドの値の作成。
- 既存のフィールドの新しい値の作成。
- レコードのシーケンスの推測。
- レコードからレポートへのデータ挿入。

CLEM 式は IBM SPSS Modeler でのデータ準備に欠くことができず、レコードとフィールド設定（選択、バランス、置換）からプロットと出力（分析、レポート、テーブル）まで、広範囲にわたるノードに使用できます。例えば、CLEM をフィールド作成ノードに使用して、比率などの、式に準拠する新規フィールドを作成できます。

CLEM 式は、また、グローバル検索と置き換え操作にも使用できます。例えば、@NULL(@FIELD) 式を置換ノードに使用して、整数値 0 でシステム欠損値を置き換えることができます（空白とも呼ばれるユーザー欠損値を置き換える場合は、@BLANK 関数を使います）。

また、より複雑な CLEM 式を作成することも可能です。例えば、条件ルールのセットに基づいて、新規フィールドを派生させることができます。例えば、以下の式を使用して、新規の値カテゴリーを作成します。
If: CardID = @OFFSET(CardID,1), Then: @OFFSET(ValueCategory,1), Else: 'exclude'

この例では @OFFSET 関数が使用され、「あるレコードの CardID フィールドの値が前のレコードの値と同じ場合は、前のレコードの ValueCategory と呼ばれるフィールドの値を返す」ことを指示しています。それ以外の場合は、文字列「exclude」を割り当てます。つまり、隣接するレコードの CardID が同じ場合、それらには同じ値カテゴリーが割り当てられます（exclude 文字列があるレコードは、後に、条件抽出ノードを使用して抜き出すことができます）。

CLEM の例

正しい指定形式や、CLEM で利用できる式の種類を理解するために、次の式の例を参照してください。

簡単な式

フィールド *After* および *Before* の値に基づいて新しいフィールドを作成する簡単な式の例を次に示します。

$(After - Before) / Before * 100.0$

フィールドの値を参照する際には、フィールド名を引用符で囲まないことに注意してください。

同様に、次の式は単にフィールド *salary* の各値の対数を返します。

```
log(salary)
```

複雑な式

より複雑で長い式を使用することもできます。次の式は、2つのフィールド (*\$KX-Kohonen* および *\$KY-Kohonen*) の値が、指定した範囲内の場合に *true* (真) の値を返します。ここでは、フィールド名に特殊文字が含まれているため、フィールド名を単一引用符で囲んでいることに注意してください。

```
(' $KX-Kohonen' >= -0.2635771036148072 and ' $KX-Kohonen' <= 0.3146203637123107  
and ' $KY-Kohonen' >= -0.18975617885589602 and  
' $KY-Kohonen' <= 0.17674794197082522) -> T
```

文字列関数など多くの関数では、さまざまなパラメーターを正しい形式で指定する必要があります。次の例では、商品が有機栽培 (*organic*)、遺伝子組み換え食品 (*genetically modified*)、または従来農法 (*conventional*) かを示すフィールド *produce_ID* の先頭文字を取得するために、関数 *subscrs* を以下のように使用しています。式の結果は、「-> `結果`」で示されます。

```
subscrs(1,produce_ID) -> `c`
```

次の式も同様です。

```
stripchar(`3`,`123`) -> `12`
```

文字は常に単一逆引用符で囲むことに注意してください。

式中での関数の組み合わせ

CLEM 式は複数の関数の組み合わせで成り立っていることもあります。関数 *subscr* および *lowertoupper* を組み合わせて、*produce_ID* の先頭文字を取得し、それを大文字に変換する式の例を次に示します。

```
lowertoupper(subscr(1,produce_ID)) -> `C`
```

この式は、次のようにさらに短い形式で記述することもできます。

```
lowertoupper(produce_ID(1)) -> `C`
```

よく使われる関数の組み合わせの例を、もう一つ次に示します。

```
locchar_back(`n`, (length(web_page)), web_page)
```

この式は、フィールド *web_page* 中の値を、値の最後の文字から先頭方向に向かって、文字 *`n`* を検索していきます。この式で値の長さに 7 のような特定の数字を使用すると、その値 (7 文字) 未満の値に対して式が不正になってしまいます。そこで、この式では *length* 関数と組み合わせることによって、動的に現在の値の長さを計算して使用します。

特殊関数

さまざまな特殊関数 (先頭に @ 記号がある) を利用することができます。よく使われる特殊関数を次に示します。

```
@BLANK('referrer ID') -> T
```

特殊関数は、組み合わせて使用され、同時に複数のフィールドに対して、フィールドが空の場合にフラグを設定するためによく用いられる方法です。

```
@BLANK(@FIELD)-> T
```

CLEM マニュアルには、ほかにもさまざまな例が記載されています。詳しくは、トピック 127 ページの『CLEM リファレンス概要』を参照してください。

値とデータ型

CLEM 式は、値、フィールド名、演算子、および関数から構成される式と同じようなものです。有効で最も単純な CLEM 式は、1 個の値またはフィールド名になります。次に、有効な値の例を示します。

```
3
1.79
'banana'
```

次に、有効なフィールド名の例を示します。

```
Product_ID
'$P-NextField'
```

「*Product_ID*」は Market Basket データ・セットからのフィールド名で、「*\$P-NextField'*」はパラメータ名になります。式の値は、指定されたフィールドの値になります。通常、フィールド名は文字から始まります。数字や下線 () を記述することもできます。フィールド名を引用符で囲めば、この規則に従わないフィールド名も使用できます。CLEM の値は次のいずれかになります。

- 文字列: "c1"、"Type 2"、"a piece of free text" など
- 整数: 12、0、-189 など
- 実数: 12.34、0.0、-0.0045 など
- 日付/時刻フィールド: 05/12/2002、12/05/2002、12/05/02 など

次の要素も使用できます。

- 文字コード: `a` や 3 など
- 項目のリスト: [1 2 3]、['Type 1' 'Type 2'] など

通常は文字コードやリストがフィールドの値になることはありません。これらの要素は、CLEM 関数の引数として使われます。

引用規則

本ソフトウェアでは、CLEM 式で使われるフィールド、値、パラメーター、および文字列などを柔軟に指定することができますが、次の規則にしたがって式を作成することをお勧めします。

- 文字列: 文字列を指定する場合は、常に二重引用符を使用します ("Type 2" や "value")。単一引用符を使用することもできますが、引用符で囲まれたフィールドと誤解される危険性があります。
- 文字: 常に ` ` のような単一逆引用符を使用します。例えば、関数 `stripchar(`d`,`drugA`) 内の文字 `d` を参照してください。これに関する唯一の例外は、文字列中の特定の文字を参照するために整数を使用する場合です。例えば、関数 `lowertoupper("druga"(5))` → "A" 内の文字 5 を参照してください。
注: 英国および米国の標準キーボードでは、逆引用符文字 (アクサングラフ、Unicode 0060) のキーは、Esc キーのすぐ下にあります。
- フィールド: CLEM 式内で使用される場合、通常、フィールドは引用符で囲まれません (`subscr(2,arrayID)`) → CHAR)。スペースや他の特殊文字を囲む必要がある場合には、単一引用符を使用します ('Order Number')。単一引用符で囲まれているのにデータセットで未定義のフィールドがあると、それは文字列として読み込まれてしまいます。
- パラメーター: 常に単一引用符を使用します ('\$P-threshold')。

式と条件

CLEM 式が返す結果の例を次に示します (新しい値を作成する場合に使用)。

```
Weight * 2.2
Age + 1
sqrt(Signal-Echo)
```

または、次のように、真 (*true*) あるいは偽 (*false*) を評価することもできます (条件に基づいて抽出する場合に使用)。

```
Drug = "drugA"
Age < 16
not(PowerFlux) and Power > 2000
```

CLEM 式の中では、次のように演算子と関数を自由に組み合わせることができます。

```
sqrt(abs(Signal)) * max(T1, T2) + Baseline
```

括弧と演算子の優先順位によって、式が計算される順序が決まります。この例では、計算は次の順序で行われます。

- `abs(Signal)` が計算され、その結果に対して `sqrt` が適用されます。
- `max(T1, T2)` が計算されます。
- 2 つの結果を乗算します。x は + より優先されます。
- 最後に、上の結果に `Baseline` が加算されます。

優先順位を降順 (最初に計算されるものから最後に計算されるもの順) に記載すると、次のようになります。

- 関数の引数
- 関数の呼び出し
- **xx**
- **x / mod div rem**
- + -
- > < >= <= /== == = /=

優先順位を変更する場合、または計算の順序がよくわからない場合は、次のように括弧を使用して、計算の順序を明確に指定することができます。

```
sqrt(abs(Signal)) * (max(T1, T2) + Baseline)
```

ストリーム、セッション、およびスーパーノード・パラメーター

パラメーターは、CLEM 式とスクリプトで使用するために定義できます。実際のところ、パラメーターはユーザー定義の変数であり、保存されて、現在のストリーム、セッション、またはスーパーノードで存続します。さらに、スクリプトを使用する場合と同様に、ユーザー・インターフェースからもアクセスできます。例えば、ストリームを保存すると、そのストリームに設定されているパラメーターも保存されます (これは、ローカル・スクリプト変数と異なる点です。ローカル・スクリプト変数は、宣言されたスクリプト内でのみ使用できます)。多くの場合、パラメーターはスクリプトで使用され、スクリプト内でハードコーディングする必要のないフィールドや値の情報を提供することにより、スクリプトの動作を制御します。

パラメーターの有効範囲は、それがどこで設定されたかによって異なります。

- ストリーム・パラメーターは、ストリーム・スクリプト内またはストリーム・プロパティのダイアログ・ボックス内で設定でき、ストリーム内のすべてのノードで使用できます。Clem 式ビルダーの「パラメーター」リストに表示されます。
- セッション・パラメーターは、スタンドアロン・スクリプト内または「セッション・パラメーター」ダイアログ・ボックス内で設定できます。セッション・パラメーターは、現在のセッションのすべてのストリーム（「マネージャ」ウィンドウの「ストリーム」タブに表示されているすべてのストリーム）で利用できます。

パラメーターは、スーパーノード用にも設定できます。この場合、スーパーノード内にカプセル化されたノードでだけ表示できます。

CLEM 式でのパラメーターの使用

CLEM 式で、パラメーターは \$P-pname の形式で表されます。ここで、pname はパラメーターの名前です。CLEM 式でパラメーターを使用する場合は、'\$P-scale' のように単一引用符で囲む必要があります。

利用できるパラメーターは、Clem 式ビルダーを使用して簡単に参照することができます。現在のパラメーターを表示するには

1. CLEM 式を利用できる任意のダイアログ・ボックスで、Clem 式ビルダーボタンをクリックします。
2. 「フィールド」リストで、「パラメーター」を選択します。

CLEM 式に挿入するパラメーターをリストから選択することができます。詳しくは、トピック 122 ページの『フィールド、パラメーター、およびグローバル変数の選択』を参照してください。

文字列の処理

文字列に対しては、次のような操作を行うことができます。

- 大文字または小文字への文字列の変換 — `uppertolower(CHAR)`。
- 文字列変数からの `ID_` や `\$` などの指定された文字の削除 — `stripchar(CHAR,STRING)`。
- 文字列変数の長さ (文字数) の判別 — `length(STRING)`。
- 文字列値のアルファベット順のチェック — `alphabefore(STRING1, STRING2)`。
- 値からの先頭または末尾の空白文字の削除 — `trim(STRING)`、`trim_start(STRING)`、または `trimend(STRING)`。
- 文字列からの最初または最後の n 文字の抽出 — `startstring(LENGTH, STRING)` または `endstring(LENGTH, STRING)`。例えば、製品名と 4 桁の ID コードを結合する (ACME CAMERA-D109) フィールド名 *item* があるとします。4 桁のコードのみを含む新規フィールドを作成するには、フィールド作成コードで次の式を指定します。
`endstring(4, item)`
- 特定のパターンの一致 — `STRING matches PATTERN`。例えば、職名に「market」が付いている人を選択するには、条件抽出ノードで次のように指定します。
`job_title matches "*market*"`
- 文字列内のサブ文字列のすべてのインスタンスの置換 — `replace(SUBSTRING, NEWSUBSTRING, STRING)`。垂直線 (|) など、サポートされていない文字のすべてのインスタンスを、テキスト・マイニングを行う前にセミコロンと置き換えるには、置換ノードの `replace` 関数を使用します。「対象フィールド」で、サポートされていない文字のあるすべてのフィールドを選択します。**Replace:** 条件の場合、「常時」を選択し、「置換値」の下で以下の条件を指定します。`replace('|',';',',@FIELD)`

- 特定のサブ文字列の有無に基づいた、フラグ型フィールドの作成。例えば、それぞれの回答に次のような式を使用して個別のフラグ型フィールドを生成するために、フィールド生成ノードの文字列関数を使用できます。

```
hassubstring(museums,"museum_of_design")
```

詳しくは、トピック 143 ページの『文字列関数』を参照してください。

空白および欠損値の処理

空白または欠損値を置き換えることは、データ・マイナーに一般的なデータ準備タスクです。CLEM では、空白処理を自動化するさまざまなツールを用意しています。空白の処理には、置換ノードがよく使用されますが、次の関数は CLEM 式を利用できる任意のノードで使用することができます。

- @BLANK(FIELD) は、Age などの特定のフィールドの値が空白であるレコードを判別するために使用できます。
- @NULL(FIELD) は、指定されたフィールドの値がシステム欠損値であるレコードを判別するために使用できます。IBM SPSS Modeler では、システム欠損値は \$null\$ 値として表示されます。

詳しくは、トピック 159 ページの『空白値とヌル値処理関数』を参照してください。

数値の処理

IBM SPSS Modeler では、次のような数値に対する標準の操作を利用することができます。

- 指定した角度のサインの計算 — $\sin(\text{NUM})$
- 数値フィールドの自然対数の計算 — $\log(\text{NUM})$
- 2 つの数値の合計の計算 — $\text{NUM1} + \text{NUM2}$

詳しくは、トピック 138 ページの『数値関数』を参照してください。

時間と日付の処理

時間および日付の形式は、データ・ソースやロケールによってさまざまです。使用する日付と時間の書式はストリームごとに異なり、「ストリームのプロパティ」ダイアログ・ボックスで設定されます。日付/時間フィールドの処理に一般的に使われる関数の例を次に示します。

経過時間の算出

基準日からの経過時間は、次のような関数ファミリーを使用すれば簡単に算出することができます。この関数は、基準日から日付文字列 DATE で指定された日付までの月数を、実数で返します。これは、1 か月を 30.0 日と仮定した近似値になります。

```
date_in_months(Date)
```

日付/時間値の比較

日付/時間フィールドの値は、次のような関数を使用してレコード間で比較することができます。この関数は、日付文字列 DATE1 で指定された日付が、日付文字列 DATE2 で指定された日付よりも前の場合に、true (真) の値を返します。それ以外の場合、この関数は 0 の値を返します。

```
date_before(Date1, Date2)
```

差分の計算

次のような関数を使用すれば、2 つの時間や 2 つの日付の差分を算出することもできます。

```
date_weeks_difference(Date1, Date2)
```

この関数は、日付文字列 DATE1 が表す日付から日付文字列 DATE2 が表す日付までの時間を、実数の週数で返します。ここでは、1 週間を 7.0 日と仮定しています。DATE2 が DATE1 よりも前の場合、この関数は負の数値を返します。

今日の日付

関数 @TODAY を使用すれば、データ・セットに現在の日付を追加することができます。現在の日付は、「ストリームのプロパティ」ダイアログ・ボックスで指定されている日付の形式を使用して、指定されたフィールドまたは新しいフィールドに文字列として追加されます。詳しくは、トピック 148 ページの『日付および時刻の関数』を参照してください。

複数フィールドの要約

CLEM 言語には、複数フィールドにわたって要約統計量を返すさまざまな関数が含まれます。これらの関数は調査データの分析に特に役立つことがあり、その場合、1 つの質問に対する複数の回答を複数フィールドに保存することもできます。詳しくは、トピック 118 ページの『複数回答データの処理』を参照してください。

比較関数

複数フィールドにわたって、min_n 関数と max_n 関数を使用して値を比較できます。以下に例を示します。

```
max_n(['card1fee' 'card2fee' 'card3fee' 'card4fee'])
```

また、さまざまな個数関数を使用して特定基準を満たす値のカウンタを取得することができ、たとえば、値が複数フィールドに保存されている場合でもそれは可能です。例えば、5 年を越えて保持されてきたカードの数をカウントするには次のようにします。

```
count_greater_than(5, ['cardtenure' 'card2tenure' 'card3tenure'])
```

同じフィールド・セットにわたってヌル値をカウントするには次のようにします。

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

この例でカウントされているのは、カードを保持する人たちの数ではなく、保持されているカードの数であることに注意してください。詳しくは、トピック 136 ページの『比較関数』を参照してください。

指定された値が複数のフィールドで出現する回数をカウントするには、count_equal 関数を使用できます。以下の例では、値 Y を含むリスト内のフィールド数をカウントします。

```
count_equal("Y", [Answer1, Answer2, Answer3])
```

リスト内のフィールドが以下の値の場合、示されたように関数は値 Y の結果を返します。

表 7. 関数値：

Answer1	Answer2	Answer3	カウント
Y	N	Y	2
Y	N	N	1

数値関数

複数フィールドにわたって、sum_n 関数、mean_n 関数、および sdev_n 関数を使用して統計を取得できます。以下に例を示します。

```
sum_n(['card1ba1' 'card2ba1' 'card3ba1'])
mean_n(['card1ba1' 'card2ba1' 'card3ba1'])
```

詳しくは、トピック 138 ページの『数値関数』を参照してください。

フィールドのリストを生成

フィールドのリストを入力として受け入れる関数を使用する場合は、特殊関数 @FIELDS_BETWEEN(start, end) および @FIELDS_MATCHING(pattern) を入力として使用できます。例えば、フィールドの順序が上記の sum_n の例に示すようなものと仮定すると、以下は同等になります。

```
sum_n(@FIELDS_BETWEEN(card1ba1, card3ba1))
```

また、「card」で始まるすべてのフィールドにわたってヌル値の数をカウントするには、次のようにします。

```
count_nulls(@FIELDS_MATCHING('card*'))
```

詳しくは、トピック 160 ページの『特殊フィールド』を参照してください。

複数回答データの処理

さまざまな比較関数を使用して、次のような複数回答データを分析することができます。

- value_at
- first_index / last_index
- first_non_null / last_non_null
- first_non_null_index / last_non_null_index
- min_index / max_index

例えば、特定の購買を決定した 1 番目、2 番目、3 番目に重要な理由 (例えば、価格、人に勧められた、レビュー、地元の提供業者など) を尋ねる複数回答の質問があるとします。この場合、次のように最初に表示されたフィールドのインデックスを作成して価格の重要度を決定します。

```
first_index("price", [Reason1 Reason2 Reason3])
```

同様に、顧客に質問して購買の尤度の順に 3 つの車をランク付けし、次のように 3 つの個別のフィールドにコード化します。

表 8. 車のランク付けの例：

customer id	car1	car2	car3
101	1	3	2
102	3	2	1
103	2	3	1

この場合、min_index 関数を使用して、最も好きな車 (ランク #1、または最も低いランク) のフィールドのインデックスを決定します。

```
min_index(['car1' 'car2' 'car3'])
```

詳しくは、トピック 136 ページの『比較関数』を参照してください。

複数回答セットの参照

特別な @MULTI_RESPONSE_SET 関数を使用して、複数回答セットのすべてのフィールドを参照することができます。例えば、上記の 3 つの *car* フィールドが *car_rankings* という名前の複数回答セットにある場合、次の関数が同じ結果を返します。

```
max_index(@MULTI_RESPONSE_SET("car_rankings"))
```

Clem 式ビルダー

CLEM 式を手動で入力したり式ビルダーを使用したりすることができ、それによって、現在ストリームのデータ・フィールドに加えて、CLEM の関数と演算子のリストがすべて表示されます。このため、フィールド名と関数名を正確に覚えていなくても、すばやく CLEM 式を作成できます。加えて、式ビルダーのコントロールによってフィールドや値に適切な引用符が自動的に追加され、正しい式をシンタックス的に作成することが簡単になります。

注: スクリプトまたはパラメーターの設定では、Clem 式ビルダーを利用できません。

注: データ ソースを変更する場合は、ソースの変更前に、Clem 式ビルダーが選択した関数を引き続きサポートできることを確認する必要があります。すべてのデータベースがすべての関数をサポートするわけではないため、新しいデータ ソースに対して実行した場合にエラーが発生することがあります。

Clem 式ビルダーへのアクセス

Clem 式ビルダーは、条件抽出、バランス、フィールド作成、置換、精度分析、レポート、およびテーブル・ノードの式などの CLEM 式を使うノードのすべてで使用できます。式フィールドのすぐ右にある計算機ボタンをクリックして、式ビルダーを開くことができます。

CLEM 式の作成

Clem 式ビルダーには、フィールド、関数、および演算子の完全なリストがあるだけでなく、データがインスタンス化されている場合はデータ値へのアクセス手段も提供されています。

Clem 式ビルダーを使用して CLEM 式を作成する手順

1. Clem 式フィールドとフィールドのリストを参照しながら、ウィンドウに入力します。

または

2. スクロール・リストから適切なフィールドと関数を選択します。
3. 黄色の矢印ボタンをダブルクリックまたはクリックして、CLEM 式フィールドにフィールドや関数を追加します。
4. オペランドを式に挿入するには、ダイアログ・ボックスの中央にあるオペランド ボタンを使用します。

関数の選択

関数リストには、使用可能なすべての CLEM 関数と演算子が表示されています。リストをスクロールして関数を選択します。または、すばやく目的のものを探すには、ドロップダウン・リストを使用して関数または演算子のサブセットを表示します。利用できる関数は、探しやすいようにカテゴリーにグループ化されています。

これらのカテゴリーの多くは、CLEM の言語記述の参照セクションに説明されています。詳しくは、132 ページの『関数のリファレンス』を参照してください。

他のカテゴリーは次の通りです。

- 一般的な関数: 最も使用される関数のうちいくつかの選択が表示されます。
- 最近使用した演算子: 現在のセッション内で使われた CLEM 関数が表示されます。
- @ 関数: 名前の前に「@」記号が付く、すべての特別な関数のリストが表示されます。
- データベース関数: ストリームにデータベース接続が含まれている場合 (データベース入力ノードによって)、この選択にはユーザー定義関数 (UDF) など、そのデータベースから使用できる関数が表示されます。詳しくは、『データベース関数』を参照してください。
- データベース集計: ストリームに (データベース・ソース・ノードによって) データベース接続が含まれている場合、これを選択すると、そのデータベースで使用できる集計オプションがリストされます。これらのオプションは、レコード集計ノードの Clem 式ビルダーで使用できます。
- データベース・ウィンドウ集計: ストリームに (データベース・ソース・ノードによって) データベース接続が含まれている場合、これを選択すると、そのデータベースで使用できるウィンドウ集計オプションがリストされます。これらのオプションは、「フィールド設定」パレットのノード内の Clem 式ビルダーでのみ使用できます。

注: SPSS Modeler はデータベース システム ビューからウィンドウ集計関数を取得するため、使用可能なオプションはデータベースの動作によって異なります。

これらのオプションは、「集計」という名前が付いていますが、レコード集計ノードで使用するものではなく、フィールド作成または選択などのノードで使用するのに適しています。これは、これらのオプションの出力が実際のレコード集計ではなくスカラーであるためです。したがって、出力に示されるデータの量をレコード集計ノードと同じ方法で削減することはありません。例えば、この種の集計を使用してデータの行の移動平均 (「現在行とそれより前のすべての行の平均」など) を求めることができます。

- 組み込み集計: 使用できる集計モードのリストを表示します。
- 演算子: 式を作成するときに使用できるすべての演算子が表示されます。演算子は、ダイアログ・ボックスの中央にあるボタンからも使用できます。
- すべての関数: 利用できるすべての CLEM 関数が表示されます。

関数のグループを選択して、ダブルクリックすると CLEM 式フィールド中のカーソルの位置に関数が挿入されます。

データベース関数

データベース関数はさまざまな場所にリストできます。SPSS Modeler が関数の詳細を探す際に検索する場所を以下の表に示します。データベース管理者はこの表を使用して、ユーザーがさまざまな関数を使用するために必要な領域へのアクセス権を持っていることを確認できます。

さらにこの表には、データベースおよび関数の種類に基づいて、関数が使用可能な場合をフィルタリングするために使用する条件もリストします。

注: Amazon Redshift のデータベース関数を使用する場合、データベース管理者は、次の 6 つのデータベース オブジェクトへのアクセス許可をユーザーに付与することが必要になる場合があります。最初の 4 つはシステム カタログ テーブル、後の 2 つはスキーマです。

- pg_type

- pg_proc
- pg_namespace
- pg_aggregate
- information_schema
- pg_catalog

表 9. 式ビルダーでのデータベース関数

データベース	関数の種類	関数を検索する場所	関数のフィルタリングに使用する条件
DB2 LUW	UDF	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE が F、および FUNCTIONTYPE が S
DB2 LUW	UDA	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE が F、および FUNCTIONTYPE が C
DB2 iSeries	UDF	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE が F、および FUNCTION_TYPE が S
DB2 iSeries	UDA	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE が F、および FUNCTION_TYPE が C
DB2 z/OS	UDF	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE が F、および FUNCTIONTYPE が S
DB2 z/OS	UDA	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE が F、および FUNCTIONTYPE が C
SQL Server	UDF	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE が FN または FS のいずれか
SQL Server	UDA	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE が AF
Oracle	UDF	ALL_ARGUMENTS ALL_PROCEDURES	以下の条件がすべて満たされる場合: <ul style="list-style-type: none"> • OBJECT_TYPE が FUNCTION • AGGREGATE が NO • PLS_TYPE が NULL ではない
Oracle	UDA	ALL_ARGUMENTS ALL_PROCEDURES	以下の条件がすべて満たされる場合: <ul style="list-style-type: none"> • ARGUMENT_NAME が NULL • AGGREGATE が YES • PLS_TYPE が NULL ではない
Teradata	UDF	DBC.FUNCTIONS DBC.ALLRIGHTS	以下の条件がすべて満たされる場合: <ul style="list-style-type: none"> • FUNCTIONTYPE が F • COLUMNNAME が RETURN0 • SPPARAMETERTYPE が 0 • ACCESSRIGHT が EF
Teradata	UDA	DBC.FUNCTIONS DBC.ALLRIGHTS	以下の条件がすべて満たされる場合: <ul style="list-style-type: none"> • FUNCTIONTYPE が A • COLUMNNAME が RETURN0 • SPPARAMETERTYPE が 0 • ACCESSRIGHT が EF

表 9. 式ビルダーでのデータベース関数 (続き)

データベース	関数の種類	関数を検索する場所	関数のフィルタリングに使用する条件
Netezza	UDF	####.._V_FUNCTION NZA.._V_FUNCTION INZA.._V_FUNCTION	####.._V_FUNCTION に対しては、以下の条件が適用されます。 <ul style="list-style-type: none"> RESULT が TABLE% のような値を持つ文字列を含んでいない FUNCTION が '/_%' escape '/' のような値を持つ文字列を含んでいない VARARGS が FALSE NZA.._V_FUNCTION と INZA.._V_FUNCTION の両方に対しては、以下の条件が適用されます。 <ul style="list-style-type: none"> RESULT が TABLE% のような値を持つ文字列を含んでいない FUNCTION が '/_%' escape '/' のような値を持つ文字列を含んでいない BUILTIN が f VARARGS が FALSE
Netezza	UDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	以下の条件が両方とも満たされる場合: <ul style="list-style-type: none"> AGGTYPE が ANY または GROUPED VARARGS が FALSE
Netezza	WUDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	####.._V_AGGREGATE に対しては、以下の条件が適用されます。 <ul style="list-style-type: none"> AGGTYPE が ANY または ANALYTIC AGGREGATE が MAX_LABEL でない VARARGS が FALSE NZA.._V_FUNCTION と INZA.._V_FUNCTION の両方に対しては、以下の条件が適用されます。 <ul style="list-style-type: none"> AGGTYPE が ANY または ANALYTIC BUILTIN が f VARARGS が FALSE

表で使用されている用語の説明

- UDF ユーザー定義関数
- UDA ユーザー定義の集計関数
- WUDA ユーザー定義のウィンドウ集計関数
- ##### 現在接続しているデータベース。

フィールド、パラメーター、およびグローバル変数の選択

「フィールド」リストには、現在データ・ストリームで利用できるすべてのフィールドが表示されます。リストをスクロールして、目的のフィールドを選択してください。フィールドを上の式に追加するには、ダブルクリックするか、または黄色の矢印キーを使用します。

詳しくは、トピック 114 ページの『ストリーム、セッション、およびスーパーノード・パラメーター』を参照してください。

フィールドの他に、次の項目から選択することもできます。

複数回答セット。詳しくは、「*IBM SPSS Modeler* ソース、プロセス、および出力ノード」を参照してください。

最近使用した演算子: 現在のセッション内で使われたフィールド、マルチアンサー・セット、パラメーター、グローバル値が表示されます。

パラメーター。詳しくは、トピック 114 ページの『ストリーム、セッション、およびスーパーノード・パラメーター』を参照してください。

グローバル値。詳しくは、「*IBM SPSS Modeler* ソース、プロセス、および出力ノード」を参照してください。

値の表示または選択

フィールドの値は、式ビルダー、データ検査レポートを含めたシステム内の多くの場所からと、時間区分ノード内で将来の値を編集しているときに、表示できます。この機能を使用するには、データがソースまたはデータ型ノード内で完全にインスタンス化されていることが必要であり、そのようにすると、ストレージ、データ型、および値が明らかになります。

Clem 式ビルダーまたは時間区分ノードからフィールドの値を表示するには、表示するフィールドを選択して値ピッカー・ボタンをクリックすると、選択したフィールドの値を表示するダイアログ・ボックスが表示されます。次に値を選択して、その値を現在の式またはリストに貼り付けるために、「挿入」をクリックします。



図 14. 値ピッカーボタン

フラグ型と名義型フィールドの場合、すべての定義済みの値が表示されます。連続型 (数値範囲型) フィールドの場合、最小値と最大値が表示されます。

CLEM 式の検査

Clem 式ビルダーで (右下隅にある) 「検査」 をクリックして、CLEM 式が正しいかどうかを検証します。検査されていない式は赤で表示されます。エラーが検出されると、原因を示すメッセージが表示されます。

検査する項目を次に示します。

- 値やフィールド名の引用符が正しいかどうか
- パラメーターやグローバル変数の使用法が正しいかどうか
- 演算子が適切に使用されているかどうか
- 参照されているフィールドが存在しているかどうか
- 参照されているグローバル値が存在しているか、また定義が正しいかどうか

シンタックスにエラーを発見した場合は、手動で式を入力するのではなく、リストや演算子ボタンを使用して式を作成してください。この方法では、フィールドや値に適切な引用符が自動的に追加されます。

IBM Analytical Decision Management で式をビルドするときは、以下の制限に注意してください。式には、以下のいずれの項目も含めることはできません。

- IBM SPSS Modeler ストリーム パラメータへの参照
- IBM SPSS Modeler ストリーム グローバルへの参照
- データベース関数への参照
- 以下のいずれかの特特殊フィールドまたはフィールド値 @ 関数への参照:
 - @TARGET
 - @PREDICTED
 - @FIELD
 - @PARTITION_FIELD
 - @TRAINING_PARTITION
 - @TESTING_PARTITION
 - @VALIDATION_PARTITION

検索と置換

「検索/置換」ダイアログ・ボックスは、スクリプト・エディター、CLEM 式ビルダーなど、スクリプトまたは式のテキストを編集する場合、またはレポート・ノードでテンプレートを定義する場合に使用できます。これらの領域のいずれかでテキストを編集する場合、Ctrl + F キー を押してダイアログ・ボックスにアクセスし、カーソルがテキスト領域にフォーカスしていることを確認します。「フィルター」ノードを使用している場合、例えば、「設定」タブのテキスト領域から、または CLEM 式ビルダーのテキスト・フィールドからダイアログ・ボックスにアクセスできます。

1. テキスト領域内にカーソルを置いて、Ctrl + F キーを押して「検索/置換」ダイアログ・ボックスにアクセスします。
2. 検索するテキストを入力するか、最近検索した項目のドロップダウン・リストから選択します。
3. 置換テキストがある場合は、入力します。
4. 「次を検索」 をクリックして、検索を開始します。
5. 「置換」 をクリックして現在の選択内容を置換するか、「すべてを置換」 をクリックしてすべてまたは選択したインスタンスを更新します。
6. 各操作が終了すると、ダイアログ・ボックスが閉じます。テキスト領域で F3 を押すと最後の検索操作が繰り返され、または Ctrl + F キーを押すとダイアログに再度アクセスします。

検索オプション

大文字と小文字を区別：検索操作で、例えば *myvar* が *myVar* と位置するかどうかなど、大文字と小文字を区別するかどうかを指定します。この設定に関係なく、置換テキストは常に入力したとおりに挿入されます。

語全体のみ：検索操作が語内に埋め込まれたテキストに一致するかどうかを指定します。このオプションを選択すると、*spider* に関する検索は、*spiderman* または *spider-man* に一致しません。

正規表現：正規表現のシンタックスを使用するかどうかを指定します (次項参照)。このオプションを選択すると、「語全体のみ」 オプションは無効化され、その値は無視されます。

選択されたテキストのみ：「すべてを置換」 オプションを使用する場合、検索の範囲を制御します。

正規表現シンタックス

正規表現を使用すると、タブまたは改行文字などの特殊文字、*a* から *d* までなど文字のクラスまたは範囲、行の開始または終了などの境界について検索することができます。次の種類の表現がサポートされています。

表 10. 文字の一致：

文字	一致
x	文字 x
\\	円記号
¥0n	8 進法の値を持つ文字 0n (0 ≤ n ≤ 7)
¥0nn	8 進法の値を持つ文字 0nn (0 ≤ n ≤ 7)
¥0mnn	8 進法の値を持つ文字 0mnn (0 ≤ m ≤ 3, 0 ≤ n ≤ 7)
¥xhh	16 進法の値を持つ文字 0xhh
¥uhhhh	16 進法の値を持つ文字 0xhhhh
¥t	タブ文字 ('¥u0009')
¥n	改行文字 ('¥u000A')
¥r	復帰文字 ('¥u000D')
¥f	改ページ文字 ('¥u000C')
¥a	アラート (ベル) 文字 ('¥u0007')
¥e	エスケープ文字 ('¥u001B')
¥cx	xに対応する制御文字

表 11. 文字クラスの一致：

文字クラス	一致
[abc]	a、b、または c (単純クラス)
[^abc]	a、b、または c 以外の文字 (減法)
[a-zA-Z]	a から z または A から Z の各文字 (範囲)
[a-d[m-p]]	a から d、または m から p (和集合)。または、「a-dm-p」と指定することもできます
[a-z&&[def]]	a から z、および d、e、または f (交差)
[a-z&&[^bc]]	a から z のうち、b と c を除いたもの (差集合)。または、「ad-z」と指定することもできます
[a-z&&[^m-p]]	a から z のうち、m から p までを除いたもの (差集合)。または、「a-lq-z」と指定することもできます

表 12. 事前設定された文字クラス：

事前設定された文字クラス	一致
.	任意の文字 (行末に一致する場合または一致しない場合があります)
¥d	任意の数字: [0-9]
¥D	数字以外: [^0-9]
¥s	空白文字: [¥t¥n¥x0B¥f¥r]
¥S	空白文字以外: [^¥s]
¥w	ワード文字: [a-zA-Z_0-9]

表 12. 事前設定された文字クラス (続き):

事前設定された文字クラス	一致
¥W	非ワード文字: [^¥w]

表 13. 境界の一致:

境界の一致	一致
^	行頭
\$	行末
¥b	語の境界
¥B	語以外の境界
¥A	入力の開始
¥Z	最後の行末以外の入力の終了
¥z	入力の終了

第 9 章 CLEM 言語リファレンス

CLEM リファレンス概要

この項では、Control Language for Expression Manipulation (CLEM) について説明していきます。CLEM は、IBM SPSS Modeler ストリーム内で使われるデータの分析と操作の非常に役に立つツールです。ノード内で CLEM を使用して、条件の評価や値の新規作成からレポートへのデータ挿入まで、作業を実行できます。

CLEM 式は、値、フィールド名、演算子、および関数で構成されます。正しい構文を使用して、さまざまなデータ操作を作成することができます。

CLEMデータ型

CLEM データ型は次のいずれかで構成できます。

- 整数値
- 実数
- 文字
- 文字列
- リスト
- フィールド
- 日付/時刻

引用符の使用規則

IBM SPSS Modeler では、CLEM 式で使われるフィールド、値、パラメーター、および文字列などを柔軟に指定することができます。次の規則に従って式を作成することをお勧めします。

- 文字列: 文字列を指定する場合は、常に二重引用符を使用します (例: "Type 2")。単一引用符を使用することもできますが、引用符で囲まれたフィールドと誤解される危険性があります。
- フィールド: スペースや他の特殊文字を囲む必要がある場合にのみ、単一引用符を使用します (例: 'Order Number')。単一引用符で囲まれているのにデータセットで未定義のフィールドがあると、それは文字列として読み込まれてしまいます。
- パラメーター: パラメーターを使用する場合は、常に単一引用符を使用します (例: '\$P-threshold')。
- 文字: 常に単一逆引用符 (`) を使用します (例: stripchar(`d`, "drugA"))。

これらの規則は、以降の項目で詳細に説明しています。

整数値

整数は、10 進数のシーケンスとして表されます。必要に応じて、整数の前にマイナス符号 (-) を付けて負の数を表すことができます。例えば、1234、999、-77 のように記述します。

CLEM 言語は任意の精度の整数を処理します。整数の最大サイズは使用するプラットフォームによって異なります。値が大きすぎて整数フィールドに表示できない場合は、通常、フィールドのデータ型を Real に変更すると値を正確に表示できます。

実数

実数 は浮動小数点数を意味しています。実数は、1 つ以上の数値と、その後続く小数点、その後続く 1 つ以上の数値で表されます。CLEM 実数は倍精度で保持されます。

必要に応じて、実数の前にマイナス符号 (-) を付けて負の数を表すことができます。例えば、1.234、0.999、-77.001 のように記述します。指数表記で実数を表すには、<数値> e <指数> の形式を使用します。例えば、1234.0e5、1.7e-2 のように記述します。IBM SPSS Modeler アプリケーションがファイルから数値文字列を読み込んで自動的に数値に変換する場合、小数点の前に数字がない数値や小数点の後に数字がない数値も受け入れます。例えば、999. や .11 などです。ただし、これらの形式は CLEM 式では不正です。

注：CLEM 式の実数を参照する場合、現在のストリームまたはロケールの設定に関わらず、小数点区切り文字としてピリオドを使用する必要があります。例えば、次のとおりです。

```
Na > 0.6
```

次のようには、指定できません。

```
Na > 0,6
```

これは、「ストリームのプロパティ」ダイアログ・ボックスでコンマが小数点として選択された場合でも適用されます。また、コード構文が特定のロケールまたは表記方法から独立する必要があるという一般的なガイドラインを検討します。

文字

一般的に文字 (通常 CHAR と表記) は、CLEM 式内で文字列のテストを実行するために用いられます。例えば、isuppercode 関数を使用して、文字列の先頭文字が大文字かどうかを判断することができます。文字列の先頭文字に対してテストを行う必要があることを示すために、文字を使用する CLEM 式を次に示します。

```
isuppercode(subscrs(1, "MyString"))
```

CLEM 式中の特定文字のコード (場所ではなく) を表すには、単一逆引用符を `<文字>` の形式で使用します。例えば、`A`、`Z` のように記述します。

注：フィールドに対する CHAR ストレージ・タイプはありません。そのため、結果が CHAR となる式でフィールドが作成または置換された場合、その結果は文字列に変換されます。

文字列

基本的に、文字列は二重引用符で囲んでください。文字列の例は "c35product2" や "referrerID" などです。文字列内で特殊文字を指定するには、"¥\$65443" のように円記号を使用します。(円記号文字を指定するには、¥ のように円記号を 2 つ使用します)。文字列を単一引用符で囲むこともできますが、その場合引用符で囲まれたフィールド ('referrerID') と区別できない可能性があります。詳しくは、トピック 143 ページの『文字列関数』を参照してください。

リスト

リストは、順序付けられた要素のシーケンスであり、データ型が混在していることもあります。リストは、大カッコ ([]) で囲みます。リストの例は [1 2 4 16] や ["abc" "def"] などです。リストは IBM SPSS Modeler フィールドの値としては使用されません。リストは、member や oneof などの関数に引数を渡すために使用します。

注: リストの構成要素として使用できるのは、静的オブジェクト (例えば、文字列、数値、フィールド名) のみであり、関数の呼び出しではありません。

フィールド

CLEM 式内で、関数名以外の名前はフィールド名とみなされます。これらは単純に `Power`、`val27`、`state_flag` のように記述できますが、名前が数字から始まる場合またはスペースなどアルファベット以外の文字 (アンダースコアを除く) を含む場合は、`'Power Increase'`、`'2nd answer'`、`'#101'`、`'$P-NextField'` のように名前を単一引用符で囲みます。

注: 単一引用符で囲まれているのにデータセットで未定義のフィールドがあると、それは文字列として読み込まれてしまいます。

日付

日付の計算は、基準日に基づいて行われます。基準日は、「ストリームのプロパティ」ダイアログ・ボックスで指定します。デフォルトの基準日は、1900 年 1 月 1 日です。

CLEM 言語は、次の日付の形式をサポートします:

表 14. CLEM 言語の日付の形式

形式	例
DDMMYY	150163
MMDDYY	011563
YYMMDD	630115
YYYYMMDD	19630115
YYYYDDD	4 桁の年、およびそれに続く年単位の日を表す 3 桁の数字。例えば、2000032 は、2000 年の 32 番目の日付、つまり 2000 年 2 月 1 日を表します。
DAY	現在のロケールの曜日。例えば、英語の場合、Monday、Tuesday などです。
MONTH	現在のロケールの月。例えば、January、February などです。
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

表 14. CLEM 言語の日付の形式 (続き)

形式	例
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	四半期を表す 1 桁の数字 (1 から 4)、およびそれに続く文字 Q と 4 桁の年。例えば、2004 年 12 月 25 日は、4 Q 2004 として表されます。
ww WK YYYY	1 年の内での週を表す 2 桁の数で、次に文字 WK と 4 桁の年が続きます。年内の週は、週の最初の日が月曜日で、また、少なくとも最初の週に 1 日以上あるという仮定の下に計算されます。

時間

CLEM 言語は、次の時間の形式をサポートします:

表 15. CLEM 言語の時間の形式

形式	例
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 演算子

次の演算子が利用できます。

表 16. CLEM 言語の演算子:

演算	コメント	優先順位 (次項参照)
or	2 つの CLEM 式間で使用されます。どちらかが真 (true) の場合、または両方が真 (true) の場合に、値を返します。	10

表 16. CLEM 言語の演算子 (続き):

演算	コメント	優先順位 (次項参照)
and	2 つのCLEM 式間で使用されます。両方が真 (true) の場合に、値を返します。	9
=	任意の比較可能な 2 つの項目間で使用されます。ITEM1 が ITEM2 と等しい場合に true を返します。	7
==	= と同等	7
/=	任意の比較可能な 2 つの項目間で使用されます。ITEM1 が ITEM2 と等しくない場合に真が返されます。	7
/==	/= と同等	7
>	任意の比較可能な 2 つの項目間で使用されます。ITEM1 が ITEM2 より大きい場合に true を返します。	6
>=	任意の比較可能な 2 つの項目間で使用されます。ITEM1 が ITEM2 以上の場合に true を返します。	6
<	任意の比較可能な 2 つの項目間で使用されます。ITEM1 が厳密に ITEM2 より小さい場合に真を返します。	6
<=	任意の比較可能な 2 つの項目間で使用されます。ITEM1 が ITEM2 以下の場合に true を返します。	6
&&=_0	2 つの整数間に用いられます。ブール式の $INT1 \ \&\& \ INT2 = 0$ と同じになります。	6
&&/=_0	2 つの整数間に用いられます。ブール式の $INT1 \ \&\& \ INT2 = 0$ と同じになります。	6
+	2 つの数値で加算します。NUM1 + NUM2。	5
><	2 つの文字列を次のように連結します。 STRING1 >< STRING2.	5
-	1 つの数値をもう 1 つの数値から減算します。 NUM1 - NUM2。1 つの数値の前にも使用できます。- NUM。	5
*	2 つの数値を乗算するのに使用されます。 NUM1 * NUM2。	4
&&	2 つの整数間に用いられます。結果は、整数 INT1 と INT2 のビット単位の「論理積」になります。	4
&&^^	2 つの整数間に用いられます。結果は、INT1 と、INT2 のビット単位の補数との、ビット単位の「論理積」になります。	4
	2 つの整数間に用いられます。結果は、INT1 と INT2 のビット単位の「包含論理和」になります。	4
^^	整数の前に用いられます。INT のビット単位の補数を生成します。	4

表 16. CLEM 言語の演算子 (続き):

演算	コメント	優先順位 (次項参照)
/&	2 つの整数間に用いられます。結果は、INT1 と INT2 のビット単位の「排他的論理和」になります。	4
INT1 << N	2 つの整数間に用いられます。N の数だけ位置を左にシフトした INT のビット・パターンを生成します。	4
INT1 >> N	2 つの整数間に用いられます。N の数だけ位置を右にシフトした INT のビット・パターンを生成します。	4
/	1 つの数値をもう 1 つの数値で除算するのに使われます: NUM1 / NUM2。NUM1 / NUM2。	4
**	2 つの数字の間に用いられます: BASE ** POWER。BASE の POWER 乗を返します。	3
rem	2 つの整数間に用いられます。剰余 INT1 - (INT1 div INT2) * INT2 を返します。	2
div	2 つの整数間に用いられます。整数の除算を実行します。	2

演算子の優先順位

優先順位は、複数の 2 項演算子を使ったカッコで囲まれていない式などの、複雑な式の解析方法を定めるものです。例を次に示します。

3 + 4 * 5

相対的な優先順位に従って * は + より前に解析されるため、(3 + 4) * 5 ではなく 3 + (4 * 5) として解析されます。CLEM 言語中のすべての演算子には、それに対応した優先順位があります。この値が小さいほど、その演算子は処理リスト上で重要な意味を持ち、他の演算子よりも先に処理されます。

関数のリファレンス

IBM SPSS Modeler でデータを処理するために、次の CLEM 関数を利用できます。これらの関数は、フィールド作成ノードやフラグ設定ノードなど、さまざまなダイアログ ボックスにコードとして入力できます。または、Clem 式ビルダーを利用して、有効な CLEM 式を作成することができます。関数やフィールド名を覚えておく必要はありません。

表 17. IBM SPSS Modeler データで使用される CLEM 関数

関数の種類	説明
情報	フィールド値を詳しく調べる場合に用いられます。例えば、関数 is_string は、データ型が文字列型のすべてのレコードに対して真を返します。
変換	新しいフィールドの作成や、ストレージ・タイプの変換に用いられます。例えば、関数 to_timestamp は選択されているフィールドをタイムスタンプに変換します。
比較	フィールドの値を互いに比較したり、指定した文字列と比較する場合に用いられます。例えば、<= は、あるフィールドの値がもう 1 つのフィールドの値以下かどうかを比較します。

表 17. IBM SPSS Modeler データで使用される CLEM 関数 (続き)

関数の種類	説明
論理	if、then、else などの論理演算を行うために用いられます。
数値	フィールド値の自然ログ数の算出など、数値計算に用いられます。
三角関数	指定された角度のアークコサインの算出など、三角関数の計算に用いられます。
確率	学生からの <i>t</i> 分布値が特定値に満たなくなる確率など、さまざまな分布を基準にして確率を返します。
空間	地理空間データで空間の計算を実行するために用いられます。
ビット単位	整数をビット・パターンとして操作する場合に用いられます。
ランダム	無作為に項目を選択したり、無作為な数字を生成するために用いられます。
ストリング	指定した文字を削除する stripchar など、文字列に関するさまざまな操作を行うために用いられます。
SoundEx	正しいスペルが分からない場合に、特定文字の発音方法についての音声的な仮定を基準にして、文字列を検索するために用いられます。
日時	日付、時間、タイムスタンプ・フィールドに対してさまざまな操作を行うために用いられます。
シーケンス	データ・セットのレコード・シーケンスの詳細を調べたり、そのシーケンスに基づいた操作を行うために用いられます。
グローバル	グローバル値の設定ノードが作成したグローバル値にアクセスするために用いられます。例えば、@MEAN は、データ・セット全体のフィールドのすべての値の平均を参照するために用いられます。
空白とヌル	アクセス、フラグ設定、およびユーザーが指定した空白やシステム欠損値を埋めるために用いられます。例えば、@BLANK(FIELD) は、空白があるレコードに真のフラグを設定するために用いられます。
特殊フィールド	調査対象の特定のフィールドを表すために用いられます。例えば、@FIELD は複数のフィールドを作成する場合に用いられます。

関数の表記方法について

このガイドでは、次の規約を関数中のアイテムを参照するために使用します。

表 18. 関数の表記方法について :

表記方法	説明
BOOL	真 (true) または偽 (false) を示すブールまたはフラグ。
NUM、NUM1、NUM2	任意の数値。
REAL、REAL1、REAL2	1.234 または -77.01 のような任意の実数
INT、INT1、INT2	1 または -77 のような任意の実数
CHAR	`A` のような文字コード。
STRING	"referrerID" のような文字列。
LIST	["abc" "def"] のような、アイテムのリスト。
ITEM	Customer または extract_concept のような任意の実数
DATE	start_date のような日付フィールド。ここで、値の形式は DD-MON-YYYY のようになります。

表 18. 関数の表記方法について (続き) :

表記方法	説明
TIME	power_flux のような時刻フィールド。ここで、値の形式は HHMMSS のようになります。

このガイドにある関数の一覧では、関数を最初の列に、結果のタイプ (整数、文字列等) を 2 番目の列に、説明 (存在する場合) を 3 番目の列に示しています。例えば、次に rem 関数の説明を示します。

表 19. rem 関数の説明 :

関数	結果	説明
INT1 rem INT2	Number	INT1 を INT2 で除算した剰余を返します。つまり、 $INT1 - (INT1 \text{ div } INT2) * INT2$ です。

項目をリストにする方法や、関数内で文字を指定する方法などの使用方法の詳細は、別の場所で説明されています。詳しくは、トピック 127 ページの『CLEMデータ型』を参照してください。

情報関数

情報関数は、特定のフィールドの値に対する洞察を行うために用いられます。通常これらは、フラグ型フィールドの作成に用いられます。例えば、@BLANK 関数を使用して、選択したフィールドに対する値が空白のレコードを示すフラグ型フィールドを作成することができます。同様に、is_string などのストレージ・タイプ関数を使用して、フィールドのストレージ・タイプを確認することもできます。

表 20. CLEM 情報関数 :

関数	結果	説明
@BLANK(FIELD)	Boolean	上流のデータ型ノードまたは入力ノードで設定された空白処理規則 (「データ型」タブ) にしたがって、値が空白のレコードに対して真を返します。
@NULL(ITEM)	Boolean	値が未定義のすべてのレコードに対して真を返します。未定義の値はシステムのヌル値で、IBM SPSS Modeler では \$null\$ として表されます。
is_date(ITEM)	Boolean	データ型が日付のすべてのレコードに対して真 (true) を返します。
is_datetime(ITEM)	Boolean	データ型が日付、時間またはタイムスタンプのすべてのレコードに対して真 (true) を返します。
is_integer(ITEM)	Boolean	データ型が整数のすべてのレコードに対して真 (true) を返します。
is_number(ITEM)	Boolean	データ型が数値のすべてのレコードに対して真 (true) を返します。
is_real(ITEM)	Boolean	データ型が実数のすべてのレコードに対して真 (true) を返します。
is_string(ITEM)	Boolean	データ型が文字列のすべてのレコードに対して真 (true) を返します。
is_time(ITEM)	Boolean	データ型が時間のすべてのレコードに対して真 (true) を返します。
is_timestamp(ITEM)	Boolean	データ型がタイムスタンプのすべてのレコードに対して真 (true) を返します。

変換関数

変換関数により、新規フィールドを作成し、既存のファイルのストレージ・タイプを変換することができます。例えば、文字列を結合したり、切り離したりして、新しい文字列を生成できます。文字列を結合するには、演算子 \gg を使用します。例えば、フィールド Site の値が "BRAMLEY" である場合、" \gg Site" は " \gg BRAMLEY" を返します。 \gg の結果は、引数が文字列でない場合でも、常に文字列となります。したがって、フィールド V1 が 3 であり、フィールド V2 が 5 である場合、 $V1 \gg V2$ は "35" (数値ではなく文字列) を返します。

変換の関数と日付や時刻の値のような、入力に特別な型が必要なその他の関数は、「ストリームのオプション」ダイアログ・ボックスに指定されている現在の形式に依存します。例えば、値が Jan 2003、Feb 2003 などの文字列フィールドを日付ストレージへ変換する場合、ストリームのデフォルトの日付形式として一致する「MON YYYY」を選択します。

表 21. CLEM 変換関数

関数	結果	説明
ITEM1 \gg ITEM2	String	2 つのフィールドの値を連結し、結果の文字列を ITEM1ITEM2 の形式で返します。
to_integer(ITEM)	Integer	指定されたフィールドのストレージを整数に変換します。
to_real(ITEM)	Real	指定されたフィールドのストレージを実数に変換します。
to_number(ITEM)	Number	指定されたフィールドのストレージを数値に変換します。
to_string(ITEM)	String	指定されたフィールドのストレージを文字列に変換します。この関数を使用して実数を文字列に変換すると、小数点以下 6 桁を持つ値が返されます。
to_time(ITEM)	Time	指定されたフィールドのストレージを時間に変換します。
to_date(ITEM)	Date	指定されたフィールドのストレージを日付に変換します。
to_timestamp(ITEM)	Timestamp	指定されたフィールドのストレージをタイムスタンプに変換します。
to_datetime(ITEM)	Datetime	指定されたフィールドのストレージを日付、時間またはタイムスタンプ値に変換します。
datetime_date(ITEM)	Date	数値、文字列またはタイムスタンプの日付値を返します。数値（秒単位）を日付へ変換しなおすことができるのは、この関数だけです。ITEM が文字列の場合は、現在のデータ形式で文字列を解析することにより日付を作成します。この関数が正常に機能するためには、「ストリームのプロパティ」ダイアログ・ボックスの「日付の形式」に、正しい値が指定されていなければなりません。ITEM が数値の場合は、基準日（または紀元）からの秒数として解釈します。日付の端数は切り捨てられます。ITEM がタイムスタンプの場合は、日付をタイムスタンプの一部として返します。ITEM が日付の場合は、変更せずに返します。
stb_centroid_latitude(ITEM)	Integer	geohash 引数の重心に対応する緯度の整数値を返します。
stb_centroid_longitude(ITEM)	Integer	geohash 引数の重心に対応する経度の整数値を返します。

表 21. CLEM 変換関数 (続き)

関数	結果	説明
to_geohash(ITEM)	String	<p>緯度と経度に対応する geohash スtring を、指定されたビット数の密度を使用して返します。</p> <p>geohash は、緯度と経度の詳細に基づいて地理的座標のセットを識別するために使用されるコードです。 to_geohash の 3 つのパラメーターは以下のとおりです。</p> <ul style="list-style-type: none"> 緯度: 範囲 (-180 から 180)。単位は WGS84 座標系の度数。 経度: 範囲 (-90 から 90)。単位は WGS84 座標系の度数。 ビット: ハッシュを保存するために使用するビット数。範囲 (1 から 75)。これは、返される String の長さ (5 ビットごとに 1 文字使用) とハッシュの正確性の両方に影響します。例えば、5 ビット (1 文字) はおよそ 2500 キロメートルを表し、45 ビット (9 文字) はおよそ 2.3 メートルを表します。

比較関数

比較関数は、フィールドの値を互いに比較したり、指定した文字列と比較する場合に用いられます。例えば、文字列が等しいかどうかは、= を使用して確認することができます。次のように、文字列が等しいかどうかを調べます。Class = "class 1"。

数値の比較を目的とする場合、greater (より大きい) は正の無限大に近いことを意味し、lesser (より小さい) は負の無限大に近いことを意味します。つまり、すべての負の数値は、すべての正の数値より小さいこととなります。

表 22. CLEM 比較関数

関数	結果	説明
count_equal(ITEM1, LIST)	Integer	フィールドの LIST から ITEM1 と等しい数値を返します。ITEM1 が NULL の場合は、NULL を返します。
count_greater_than(ITEM1, LIST)	Integer	フィールドの LIST から ITEM1 より大きい数値を返します。ITEM1 が NULL の場合は、NULL を返します。
count_less_than(ITEM1, LIST)	Integer	フィールドの LIST から ITEM1 より小さい数値を返します。ITEM1 が NULL の場合は、NULL を返します。
count_not_equal(ITEM1, LIST)	Integer	フィールドの LIST から ITEM1 と等しくない数値を返します。ITEM1 が NULL の場合は、NULL を返します。
count_nulls(LIST)	Integer	フィールドの LIST から NULL 値の数を返します。
count_non_nulls(LIST)	Integer	フィールドの LIST から NULL 値以外の数を返します。
date_before(DATE1, DATE2)	Boolean	日付値の順序の確認に用いられます。DATE1 が DATE2 より前の場合に真 (true) を返します。
first_index(ITEM, LIST)	Integer	フィールドの LIST から ITEM を含む最初のフィールドの索引、または値が見つからないなら 0 を返します。サポート対象は、文字列、整数、実数型のみです。
first_non_null(LIST)	Any	提供されるフィールド・リストの最初の非ヌル値を返します。ストレージ・タイプはすべてサポート対象です。

表 22. CLEM 比較関数 (続き)

関数	結果	説明
first_non_null_index(LIST)	Integer	非ヌル値を含む特定の LIST の最初のフィールドの索引、またはすべての値がヌル 0 を返します。ストレージ・タイプはすべてサポート対象です。
ITEM1 = ITEM2	Boolean	ITEM1 が ITEM2 と等しい場合に真 (true) を返します。
ITEM1 /= ITEM2	Boolean	2 つの文字列が異なるか、同じでも 0 の場合に真 (true) を返します。
ITEM1 < ITEM2	Boolean	ITEM1 が ITEM2 より小さい場合に真 (true) を返します。
ITEM1 <= ITEM2	Boolean	ITEM1 が ITEM2 以下の場合に真 (true) を返します。
ITEM1 > ITEM2	Boolean	ITEM1 が ITEM2 より大きい場合に真 (true) を返します。
ITEM1 >= ITEM2	Boolean	ITEM1 が ITEM2 以上の場合に真 (true) を返します。
last_index(ITEM, LIST)	Integer	フィールドの LIST から ITEMを含む最新のフィールドの索引、または値が見つからないなら 0 を返します。サポート対象は、文字列、整数、実数型のみです。
last_non_null(LIST)	Any	提供されるフィールド・リストの最後の非ヌル値を返します。ストレージ・タイプはすべてサポート対象です。
last_non_null_index(LIST)	Integer	非ヌル値を含む特定の LIST の最後のフィールドの索引、またはすべての値がヌル 0 を返します。ストレージ・タイプはすべてサポート対象です。
max(ITEM1, ITEM2)	Any	ITEM1 または ITEM2 のどちらか大きい方を返します。
max_index(LIST)	Integer	数値フィールドの LIST から最大値を含むフィールドの索引、またはすべての値がヌルなら 0 を返します。例えば、3番目にリストされたフィールドに最大値がある場合は、インデックス値 3 を返します。複数のフィールドに最大値がある場合は、最初にリストされたもの (左端) を返します。
max_n(LIST)	Number	数値フィールドの LIST から最大値を返します。フィールドのすべての値が NULL の場合は、NULL を返します。
member(ITEM, LIST)	Boolean	ITEM が、指定された LIST のメンバーの場合に真 (true) を返します。それ以外の場合は、偽 (false) の値が返されます。また、フィールド名のリストを定義することもできます。
min(ITEM1, ITEM2)	Any	ITEM1 または ITEM2 のどちらか小さい方を返します。
min_index(LIST)	Integer	数値フィールドの LIST から最小値を含むフィールドの索引、またはすべての値がヌルなら 0 を返します。例えば、3番目にリストされたフィールドに最小値がある場合は、インデックス値 3 を返します。複数のフィールドに最小値がある場合は、最初にリストされたもの (左端) を返します。
min_n(LIST)	Number	数値フィールドの LIST から最小値を返します。フィールドのすべての値が NULL の場合は、NULL を返します。
time_before(TIME1, TIME2)	Boolean	時間値の順序の確認に用いられます。TIME1 が TIME2 より前の場合に真 (true) が返されます。
value_at(INT, LIST)		オフセットが有効値の範囲外の場合 (つまり 1 以上またはリストされたフィールドの数値以下)、オフセット NT または NULL でそれぞれリストされたフィールドの値を返します。ストレージ・タイプはすべてサポート対象です。

論理関数

CLEM 式を使用して論理演算を行うことができます。

表 23. CLEM 論理関数

関数	結果	説明
COND1 and COND2	<i>Boolean</i>	この演算は論理積で、COND1 と COND2 の両方が真 (true) の場合に真 (true) の値を返します。COND1 が偽 (false) の場合、COND2 は評価されません。こうすることによって、COND2 の演算が正しいかどうかを COND1 で判断することができます。例えば、length(Label) >=6 and Label(6) = 'x' です。
COND1 or COND2	<i>Boolean</i>	この演算は、(包括的) 論理和で、COND1 と COND2 のどちらかが真 (true) の場合、または両方とも真 (true) の場合に、真 (true) の値を返します。COND1 が真 (true) の場合、COND2 は評価されません。
not(COND)	<i>Boolean</i>	この演算は論理否定で、COND が偽 (false) の場合に真 (true) の値を返します。それ以外の場合、この演算は 0 の値を返します。
if COND then EXPR1 else EXPR2 endif	<i>Any</i>	この演算は条件評価です。COND が真 (true) の場合、この演算は EXPR1 の結果を返します。それ以外の場合は、EXPR2 を評価した結果を返します。
if COND1 then EXPR1 elseif COND2 then EXPR2 else EXPR_N endif	<i>Any</i>	この演算は複数の分岐を持つ条件評価です。COND1 が真 (true) の場合、この演算は EXPR1 の結果を返します。それ以外の場合で、COND2 が真ならば、この演算は EXPR2 を評価した結果を返します。それ以外の場合は、EXPR_N を評価した結果を返します。

数値関数

CLEM には、一般的に使われるさまざまな数値関数が用意されています。

表 24. CLEM 数値関数 :

関数	結果	説明
-NUM	<i>Number</i>	NUM を否定する場合に用いられます。対応する数値の符号を逆にした値を返します。
NUM1 + NUM2	<i>Number</i>	NUM1 と NUM2 を合計した値を返します。
NUM1 - NUM2	<i>Number</i>	NUM1 から NUM2 を減算した値を返します。
NUM1 * NUM2	<i>Number</i>	NUM1 を NUM2 で乗算した値を返します。
NUM1 / NUM2	<i>Number</i>	NUM1 を NUM2 で除算した値を返します。
INT1 div INT2	<i>Number</i>	整数の除算を行うために用いられます。INT1 を INT2 で除算された値を返します。
INT1 rem INT2	<i>Number</i>	INT1 を INT2 で除算した剰余を返します。つまり、INT1 - (INT1 div INT2) * INT2 です。
INT1 mod INT2	<i>Number</i>	この関数は廃止されました。代わりに rem 関数を使用します。

表 24. CLEM 数値関数 (続き):

関数	結果	説明
BASE ** POWER	Number	POWER 乗までべき乗した BASE を返します。BASE と POWER はどちらも任意の数値です (ただし、POWER が整数の 0 以外のいずれかのデータ型のゼロの場合、BASE はゼロ以外である必要があります)。POWER が整数の場合は、BASE のべき乗を順次掛けていくことによって計算されます。したがって、BASE が整数の場合、結果は整数になります。POWER が整数の 0 の場合、結果は常に BASE と同じデータ型の 1 になります。POWER が整数ではない場合、結果は $\exp(\text{POWER} * \log(\text{BASE}))$ のように計算されます。
abs(NUM)	Number	NUM の絶対値を返します。この値は常に、同じデータ型の数値になります。
exp(NUM)	Real	NUM 乗までべき乗した e を返します。この e は自然対数の底です。
fracof(NUM)	Real	NUM-intof(NUM) として定義される、NUM の小数部を返します。
intof(NUM)	Integer	引数を切り捨てて整数にします。NUM と同じ符号で、abs(INT) <= abs(NUM) となる最大の絶対値を持つ整数を返します。
log(NUM)	Real	NUM の自然対数 (底 e) を返します。NUM は、ゼロ以外でなければなりません。
log10(NUM)	Real	NUM の常用対数を返します。この関数は、 $\log(\text{NUM}) / \log(10)$ として定義されます。
negate(NUM)	Number	NUM を否定する場合に用いられます。対応する数値の符号を逆にした値を返します。
round(NUM)	Integer	NUM が正の数の場合は intof(NUM+0.5)、NUM が負の数の場合は intof(NUM-0.5) を使用して、NUM を整数に丸めるために用いられます。
sign(NUM)	Number	NUM の符号を判断するために用いられます。NUM が整数の場合、この演算は -1, 0, または 1 を返します。NUM が実数の場合、NUM が負、0、または正の値かによって、-1.0, 0.0、または 1.0 を返します。
sqrt(NUM)	Real	NUM の平方根を返します。NUM は正でなければなりません。
sum_n(LIST)	Number	数値フィールドの LIST から合計値を返します。フィールドのすべての値が NULL の場合は、NULL を返します。
mean_n(LIST)	Number	数値フィールドの LIST から平均値を返します。フィールドのすべての値が NULL の場合は、NULL を返します。
sdev_n(LIST)	Number	数値フィールドの LIST から標準偏差を返します。フィールドのすべての値が NULL の場合は、NULL を返します。

三角関数

この項の関数は、すべて引数として角度を取るかまたは、結果として角度を返します。どちらの場合も、角度の単位 (ラジアンまたは度) は関連するストリーム・オプションの設定によって制御されます。

表 25. CLEM 三角関数

関数	結果	説明
arccos(NUM)	Real	指定した角度のアーコサインを計算します。
arccosh(NUM)	Real	指定した角度の双曲線アーコサインを計算します。

表 25. CLEM 三角関数 (続き)

関数	結果	説明
arcsin(NUM)	Real	指定した角度のアークサインを計算します。
arcsinh(NUM)	Real	指定した角度の双曲線アークサインを計算します。
arctan(NUM)	Real	指定した角度のアークタンジェントを計算します。
arctan2(NUM_Y, NUM_X)	Real	NUM_Y / NUM_X のアークタンジェントを計算し、2 つの数値の符号を使用して、象限情報を派生させます。結果は、 $-\pi < \text{ANGLE} \leq \pi$ (radians) - $180 < \text{ANGLE} \leq 180$ (degrees) の範囲内の実数です。
arctanh(NUM)	Real	指定した角度の双曲線アークタンジェントを計算します。
cos(NUM)	Real	指定した角度のコサインを計算します。
cosh(NUM)	Real	指定した角度の双曲線コサインを計算します。
pi	Real	この定数は、パイに最も近い値の実数です。
sin(NUM)	Real	指定した角度のサインを計算します。
sinh(NUM)	Real	指定した角度の双曲線サインを計算します。
tan(NUM)	Real	指定した角度のタンジェントを計算します。
tanh(NUM)	Real	指定した角度の双曲線タンジェントを計算します。

確率関数

確率分布で、学生からの t 分布値が特定値に満たなくなる確率など、さまざまな分布を基準にして確立が返されます。

表 26. CLEM 確率関数

関数	結果	説明
cdf_chisq(NUM, DF)	Real	指定した自由度のカイ 2 乗分布からの値が特定の数字より小さくなる確率を返します。
cdf_f(NUM, DF1, DF2)	Real	DF1 と DF2 の自由度の F 分布からの値が指定した数字より小さくなる確率を返します。
cdf_normal(NUM, MEAN, STDDEV)	Real	指定した平均と標準偏差の正規分布からの値が指定した数字より小さくなる確率を返します。
cdf_t(NUM, DF)	Real	指定した自由度の t 分布からの値が特定の数字より小さくなる確率を返します。

空間処理関数

空間処理関数は地理空間データで使用できます。例えば、これらの関数を使用して 2 地点間の距離やポリゴンの面積などを計算できます。空間述部 (within、close to など) に基づく複数の地理空間データ セットの結合が必要になる場合もあります。この結合は結合条件を使用して実行できます。

これらの空間処理関数は、「ツール」 > 「ストリームのプロパティ」 > 「オプション」 > 「地理空間」で指定した座標系とともに機能します。

注: これらの空間処理関数は、3 次元データには適用されません。3 次元データがストリームにインポートされた場合、これらの関数は最初の 2 次元のみを使用します。z 軸の値は無視されます。

表 27. CLEM 空間処理関数

関数	結果	説明
close_to(SHAPE,SHAPE,NUM)	Boolean	2 つの形状が互いに一定の距離内にあるかどうかを検証します。投影座標系が使用される場合、距離の単位はメートルです。座標系が使用されない場合は任意の単位になります。
crosses(SHAPE,SHAPE)	Boolean	2 つの形状が互いに交差しているかどうかを検証します。この関数は、2 つの行ストリング形状、または 1 つの行ストリングと 1 つのポリゴンの場合に適しています。
overlap(SHAPE,SHAPE)	Boolean	2 つのポリゴン間に交差部があるかどうか、およびその交差部が両方の形状の内部にあることを検証します。
within(SHAPE,SHAPE)	Boolean	形状 1 の全体がポリゴンに含まれているかどうかを検証します。
area(SHAPE)	Real	指定されたポリゴンの面積を返します。投影システムが使用されている場合、この関数によって平方メートル数が返されます。座標系が使用されない場合は任意の単位になります。形状はポリゴンまたは複数ポリゴンでなければなりません。
num_points(SHAPE,LIST)	Integer	ポリゴンの境界内に含まれている、1 つのポイント フィールドのポイント数 (マルチポイント) を返します。形状 1 はポリゴンまたは複数ポリゴンでなければなりません。
distance(SHAPE,SHAPE)	Real	形状 1 と形状 2 の間の距離を返します。投影座標系が使用される場合、この関数によってメートル数が返されます。座標系が使用されない場合は任意の単位になります。形状 1 と形状 2 はどの地理測定タイプでもかまいません。

ビット単位の整数演算

これらの関数を使用すると、2 の補数値を表すビット・パターンとして整数を操作できます。この場合、ビット位置 N は 2^{*N} の重みを持ちます。ビットは 0 から上方向に番号が付けられます。これらの演算は、整数の符号ビットが左方向に無限に拡張されているかのように処理します。つまり、最上位ビットを超えたすべての位置で、正の整数は 0 のビットを持ち、負の整数は 1 のビットを持ちます。

表 28. CLEM のビット単位の整数演算：

関数	結果	説明
~~ INT1	Integer	INT1 のビット単位の補数を生成します。つまり、INT1 で 0 である各ビットが 1 になります。~~ INT = -(INT + 1) は常に真です。
INT1 INT2	Integer	この演算の結果は、INT1 と INT2 のビット単位の「包括的論理和」または です。つまり、INT1 と INT2 のどちらかまたは両方に 1 がある各ビットが 1 になります。
INT1 /& INT2	Integer	この演算の結果は、INT1 と INT2 のビット単位の「排他的論理和」または です。つまり、INT1 と INT2 のどちらかにだけ (両方ではない) 1 がある各ビットが 1 になります。
INT1 && INT2	Integer	INT1 と INT2 のビット単位の「論理積」を生成します。つまり、INT1 と INT2 の両方で 1 である各ビットが 1 になります。

表 28. CLEM のビット単位の整数演算 (続き) :

関数	結果	説明
<code>INT1 &&~ INT2</code>	<i>Integer</i>	この演算の結果は、 <i>INT1</i> と、 <i>INT2</i> のビット単位の補数との、ビット単位の「論理積」です。つまり、 <i>INT1</i> のあるビット位置の値が 1 で、 <i>INT2</i> の同じビット位置の値が 0 の場合に、結果は 1 になります。これは <code>INT1 && (~INT2)</code> と同じであり、 <i>INT2</i> 内で設定された <i>INT1</i> のビットを消去する場合に役に立ちます。
<code>INT << N</code>	<i>Integer</i>	<i>N</i> の数だけ位置を左にシフトした <i>INT1</i> のビット・パターンを生成します。 <i>N</i> の値が負の場合は、右にシフトします。
<code>INT >> N</code>	<i>Integer</i>	<i>N</i> の数だけ位置を右にシフトした <i>INT1</i> のビット・パターンを生成します。 <i>N</i> の値が負の場合は、左にシフトします。
<code>INT1 &&=_0 INT2</code>	<i>Boolean</i>	ブール式 <code>INT1 && INT2 != 0</code> と同じですが、より効率的です。
<code>INT1 &&/=_0 INT2</code>	<i>Boolean</i>	ブール式 <code>INT1 && INT2 == 0</code> と同じですが、より効率的です。
<code>integer_bitcount(INT)</code>	<i>Integer</i>	<i>INT</i> の 2 の補数表現の、1 または 0 のビットの数をカウントします。 <i>INT</i> が負ではない場合、 <i>N</i> は 1 のビット数になります。 <i>INT</i> が負の場合、 <i>N</i> は 0 のビット数になります。符号の拡張のため、負ではない整数には無限大の数の 0 のビットがあります。また、負の整数には無限大の数の 1 のビットがあります。 <code>integer_bitcount(INT) = integer_bitcount(-(INT+1))</code> は常に真です。
<code>integer_leastbit(INT)</code>	<i>Integer</i>	整数 <i>INT</i> の最下位ビット・セットのビット位置 <i>N</i> を返します。 <i>N</i> は、最も大きな 2 のべき乗です。 <i>N</i> によって <i>INT</i> が正確に割られます。
<code>integer_length(INT)</code>	<i>Integer</i>	<i>INT</i> の長さのビット数を、2 の補数の整数として返します。つまり、 <i>N</i> は、 <code>INT < (1 << N) if INT >= 0 INT >= (-1 << N) if INT < 0</code> であるような最も小さい整数です。 <i>INT</i> が負ではない場合、符号なしの整数として <i>INT</i> を表すには、少なくとも <i>N</i> ビットのフィールドが必要です。または、 <i>INT</i> の符号にかかわらず、 <i>INT</i> を符号付きの整数として表すには、少なくとも <i>N+1</i> ビットが必要です。
<code>testbit(INT, N)</code>	<i>Boolean</i>	整数 <i>INT</i> の <i>N</i> の位置にあるビットを検定し、ビット <i>N</i> の状態をブール値として返します。このブール値は、1 の場合は真 (true)、0 の場合は偽 (false) になります。

乱数関数

次の関数は、無作為に項目を選択したり、乱数を生成する場合に用いられます。

表 29. CLEM 乱数関数 :

関数	結果	説明
<code>oneof(LIST)</code>	<i>Any</i>	無作為 (ランダム) に選択された <i>LIST</i> の要素を返します。 <i>LIST</i> 項目は <code>[ITEM1,ITEM2,...,ITEM_N]</code> のように入力する必要があります。フィールド名のリストを指定することもできます。

表 29. CLEM 乱数関数 (続き):

関数	結果	説明
random(NUM)	Number	1~NUM の範囲の、同じデータ型 (INT または REAL) の一様に分布した乱数を返します。整数を使用する場合、整数だけが返されます。実数 (10 進数) を使用する場合は、実数値が返されます (精度はストリーム・オプションによって決まります)。この関数で返される可能性がある最大の乱数は、NUM になります。
random0(NUM)	Number	random(NUM) と同じ性質を持ちますが、値の範囲が 0 から始まります。この関数で返される最大の乱数が NUM になることはありません。

文字列関数

CLEM では、文字列に対して次の操作を行うことができます。

- 文字列の比較
- 文字列の生成
- 文字へのアクセス

CLEM では、文字列とは、一組の二重引用符で囲まれた一連の任意の文字のことです (例: "string quotes")。任意の単一の英数字が、文字 (CHAR) になります。これらは、`<文字>` の形式で、単一逆引用符を使用して CLEM 式内で宣言されます。例えば、`z`、`A`、または `2` です。範囲外の文字、または文字列に対する逆索引文字は、未定義の動作が生じます。

注: SQL プッシュバックを使用する文字列と使用しない文字列と比較すると、接尾空白を含むさまざまな結果を生成する場合があります。

表 30. CLEM 文字列関数:

関数	結果	説明
allbutfirst(N, STRING)	String	STRING の先頭 N 文字を削除した文字列を返します。
allbutlast(N, STRING)	String	STRING の最後の文字を削除した文字列を返します。
alphabefore(STRING1, STRING2)	Boolean	文字列のアルファベット順を確認するために用いられます。STRING1 が STRING2 より前にある場合に真を返します。
endstring(LENGTH, STRING)	String	指定した文字列から最後の N 文字を抽出します。文字列の長さが指定した長さよりも短い場合、またはそれに等しい場合は、変更されません。
hasendstring(STRING, SUBSTRING)	Integer	この関数は、isendstring(SUBSTRING, STRING) と同じです。
hasmidstring(STRING, SUBSTRING)	Integer	この関数は、ismidstring(SUBSTRING, STRING) (埋め込みサブ文字列) と同じです。
hasstartstring(STRING, SUBSTRING)	Integer	この関数は、isstartstring(SUBSTRING, STRING) と同じです。
hassubstring(STRING, N, SUBSTRING)	Integer	この関数は、issubstring(SUBSTRING, N, STRING) と同じです。N のデフォルトは 1 です。

表 30. CLEM 文字列関数 (続き) :

関数	結果	説明
count_substring(String, Substring)	Integer	指定したサブ文字列が文字列内に発生する回数を返します。例を次に示します。 count_substring("foooo.txt", "oo") は 3 を返します。
hassubstring(String, Substring)	Integer	この関数は、issubstring(Substring, 1, String) と同じです。N のデフォルトは 1 です。
isalphacode(Char)	Boolean	CHAR が、文字コードが文字である指定された文字列 (通常フィールド名) 中の文字の場合に真を返します。それ以外の場合、この関数は 0 の値を返します。例えば、isalphacode(produce_num(1)) です。
isendstring(Substring, String)	Integer	文字列 String がサブ文字列 Substring で終わる場合、この関数は、String 内の Substring の整数の添字を返します。それ以外の場合、この関数は 0 の値を返します。
islowercode(Char)	Boolean	CHAR が指定された文字列 (通常フィールド名) の小文字の場合に、真 (true) の値を返します。それ以外の場合、この関数は 0 の値を返します。例えば、islowercode(``) と islowercode(country_name(2)) は両方とも有効な式です。
ismidstring(Substring, String)	Integer	Substring が String の部分文字列で、String の初めの文字から始まっていないか、または最後の文字で終わっていない場合、この関数は部分文字列が始まる位置の添字を返します。それ以外の場合、この関数は 0 の値を返します。
isnumbercode(Char)	Boolean	指定された文字列 (通常フィールド名) の CHAR が、文字コードが数字である文字の場合に真を返します。それ以外の場合、この関数は 0 の値を返します。例えば、isnumbercode(product_id(2)) です。
isstartstring(Substring, String)	Integer	文字列 String がサブ文字列 Substring から始まる場合、この関数は添字 1 を返します。そうでない場合は、この関数は 0 の値を返します。
issubstring(Substring, N, String)	Integer	この関数は、文字列 String の N 番目の文字から始めて、文字列 Substring と等しいサブ文字列を検索します。文字列が見つかった場合、一致する部分文字列が始まる位置の添字 M (整数) を返します。それ以外の場合、この関数は 0 の値を返します。N が与えられていない場合、この関数はデフォルトで 1 になります。

表 30. CLEM 文字列関数 (続き) :

関数	結果	説明
<code>issubstring(SUBSTRING, STRING)</code>	<i>Integer</i>	この関数は、文字列 <i>STRING</i> の <i>N</i> 番目の文字から始めて、文字列 <i>SUBSTRING</i> と等しいサブ文字列を検索します。文字列が見つかった場合、一致する部分文字列が始まる位置の添字 <i>M</i> (整数) を返します。それ以外の場合、この関数は 0 の値を返します。 <i>N</i> が与えられていない場合、この関数はデフォルトで 1 になります。
<code>issubstring_count(SUBSTRING, N, STRING):</code>	<i>Integer</i>	指定した <i>STRING</i> 内で <i>N</i> 番目に発生した <i>SUBSTRING</i> のインデックスを返します。 <i>N</i> 番目に発生する <i>SUBSTRING</i> よりも少ない場合、0 を返します。
<code>issubstring_lim(SUBSTRING, N, STARTLIM, ENDLIM, STRING)</code>	<i>Integer</i>	この関数は <code>issubstring</code> と同じですが、添字 <i>STARTLIM</i> から、またはその前から始まり、添字 <i>ENDLIM</i> で、またはその前で終わるように、マッチングが制限されます。 <i>STARTLIM</i> 制約または <i>ENDLIM</i> 制約は、どちらかの引数に偽 (<code>false</code>) の値を指定することによって無効にできます。例えば、 <code>issubstring_lim(SUBSTRING, N, false, false, STRING)</code> は <code>issubstring</code> と同じです。
<code>isuppercode(CHAR)</code>	<i>Boolean</i>	この関数は <i>CHAR</i> が大文字の場合に、真 (<code>true</code>) の値を返します。それ以外の場合、この関数は 0 の値を返します。例えば、 <code>isuppercode(`)`</code> と <code>isuppercode(country_name(2))</code> は両方とも有効な式です。
<code>last(CHAR)</code>	<i>String</i>	この関数は、 <i>STRING</i> の最後の文字 <i>CHAR</i> を返します(少なくとも 1 文字以上の長さがなければなりません)。
<code>length(STRING)</code>	<i>Integer</i>	文字列 <i>STRING</i> の長さ (つまり文字列内の半角文字数) を返します。
<code>locchar(CHAR, N, STRING)</code>	<i>Integer</i>	シンボル値フィールド中の文字の位置を識別するために用いられます。この関数は、文字列 <i>STRING</i> 内の文字 <i>CHAR</i> の検索を、 <i>STRING</i> の <i>N</i> 番目の文字から開始します。この関数は、文字が見つかった (<i>N</i> から始まる) 位置を示す値を返します。文字が見つからない場合は 0 を返します。関数のオフセット (<i>N</i>) が無効な場合 (例えばオフセットが文字列の長さを超えているなど)、この関数は <code>\$null\$</code> を返します。 例えば、 <code>locchar(`n`, 2, web_page)</code> は、フィールド <i>web_page</i> 内の <code>`n`</code> 文字の検索を、フィールド値の 2 番目の文字から開始します。 注：指定する文字を、忘れずに、単一逆引用符で囲むようにしてください。

表 30. CLEM 文字列関数 (続き) :

関数	結果	説明
locchar_back(CHAR, N, STRING)	Integer	locchar に似ていますが、N 番目の文字から前方向に検索される点が異なります。例えば、locchar_back(`n`, 9, web_page) と指定すると、フィールド web_page の 9 番目の文字から、文字列の先頭方向に向かって検索が開始されます。関数のオフセットが無効な場合 (例えばオフセットが文字列の長さを超えているなど)、この関数は \$null\$ を返します。できる限り、locchar_back を関数 length(<field>) とともに使用して、フィールドの現在の値の長さを動的に使用することをお勧めします。例えば、locchar_back(`n`, (length(web_page)), web_page) です。
lowertoupper(CHAR) lowertoupper (STRING)	CHAR または String	文字または文字列を入力にすることができ、同じデータ型の新しい項目を返すために用いられます。その際小文字はすべて同じ文字の大文字に変換されます。例えば、lowertoupper(`a`)、lowertoupper("My string")、および lowertoupper(field_name(2)) はすべて有効な式です。
matches	Boolean	文字列が指定したパターンに一致する場合、真を返します。パターンは文字列リテラルにする必要があります。パターンを含むフィールド名にはなりません。クエスチョン・マーク (?) をパターンに含めて正確に 1 つの文字に一致させることができます。アスタリスク (*) は 0 かそれ以上の文字数に一致します。リテラル・クエスチョン・マークまたはアスタリスクを (むしろ、ワイルドカードとして使用しないで) 一致させるために、バックスラッシュをエスケープ文字として使用することができます。
replace(SUBSTRING, NEWSUBSTRING, STRING)	String	指定した STRING 内で、SUBSTRING のすべてのインスタンスを NEWSUBSTRINGを使用して置き換えます。
replicate(COUNT, STRING)	String	指定した回数だけコピーされた元の文字列を含む文字列を返します。
stripchar(CHAR,STRING)	String	文字列またはフィールドから、指定した文字を削除します。この関数を利用すれば、データから通貨表記などの余分な記号を削除して、単純な数字または名前を取得できます。例えば、シンタックス stripchar(`\$`, 'Cost') を使用すると、すべての値からドル記号を削除した新しいフィールドが返されます。 注：指定する文字を、忘れずに、単一逆引用符で囲むようにしてください。

表 30. CLEM 文字列関数 (続き) :

関数	結果	説明
skipchar(CHAR, N, STRING)	Integer	文字列 <i>STRING</i> の <i>N</i> 文字目から、 <i>CHAR</i> 以外の文字を検索します。この関数は、見つかった文字の位置を示す整数サブ文字列を返します。 <i>N</i> 番目以降のすべての文字が <i>CHAR</i> の場合は、0 を返します。関数のオフセットが無効な場合 (例えばオフセットが文字列の長さを超えているなど)、この関数は \$null\$ を返します。 locchar は、よく関数 skipchar と一緒に、 <i>N</i> (文字列の検索開始点) の値を判断するために用いられます。例えば、skipchar(`s`, (locchar(`s`, 1, "MyString")), "MyString") です。
skipchar_back(CHAR, N, STRING)	Integer	skipchar に似ていますが、 <i>N</i> 番目の文字から前に戻る方向に検索される点が異なります。
startstring(LENGTH, STRING)	String	指定した文字列から最初の <i>N</i> 文字を抽出します。文字列の長さが指定した長さよりも短いか、またはそれに等しい場合は、変更されません。
strmember(CHAR, STRING)	Integer	locchar(CHAR, 1, STRING) と同じです。 <i>CHAR</i> が最初に出現する位置を示す整数サブ文字列、または 0 を返します。関数のオフセットが無効な場合 (オフセットが文字列の長さを超えているなど)、この関数は \$null\$ を返します。
subscrs(N, STRING)	CHAR	入力文字列 <i>STRING</i> の <i>N</i> 番目の文字 <i>CHAR</i> を返します。この関数は、STRING(<i>N</i>) という短い形式で記述することもできます。例えば、lowertoupper("name"(1)) は有効な式です。
substring(N, LEN, STRING)	String	文字列 <i>SUBSTRING</i> を返します。この文字列は、文字列 <i>STRING</i> の添字 <i>N</i> から <i>LEN</i> 文字分の文字列で構成されています。
substring_between(N1, N2, STRING)	String	添字 <i>N1</i> から始まり、添字 <i>N2</i> で終わる <i>STRING</i> のサブ文字列を返します。
trim(STRING)	String	指定した文字列から、文字列の前後の空白文字を削除します。
trim_start(STRING)	String	指定した文字列から、文字列の前の空白文字を削除します。
trimend(STRING)	String	指定した文字列から、文字列の後の空白文字を削除します。
unicode_char(NUM)	CHAR	16 進数値ではなく 10 進数を入力する必要があります。 <i>NUM</i> の Unicode 値を返します。
unicode_value(CHAR)	NUM	<i>CHAR</i> の Unicode 値を返します。

表 30. CLEM 文字列関数 (続き):

関数	結果	説明
uppertolower(CHAR) uppertolower (STRING)	CHAR または String	文字または文字列を入力にすることができ、同じデータ型の新しい項目を返すために用いられます。その際、大文字はすべて同じ文字の小文字に変換されます。 注：文字列は二重引用符で、文字は単一の逆引用符で忘れずに指定するようにしてください。単純なフィールド名の場合は、引用符は使用しません。

SoundEx 関数

SoundEx は、サウンドは分かっているにもかかわらず正しいスペルが分からない場合に、文字列を検索するために用いられる方法です。1918年に開発されたこの方法では、特定文字の発音方法についての音声的な仮定を基準にする、似通ったサウンドの単語が検索されます。この方法は、例えば、似通った名前前のスペルや発音がさまざまに異なる場合に、データベースで名前検索を行うために使用されます。基本的な SoundEx アルゴリズムはさまざまな文献で引用されており、また、(ph や f のように、文字列の前の文字の組み合わせが、同じサウンドを持つにもかかわらず一致しないことなど) 制約があることは知られていますが、ほとんどのデータベースで何らかの形でサポートされています。

表 31. CLEM soundex 関数

関数	結果	説明
soundex(STRING)	Integer	指定した STRING の 4 文字の SoundEx コードを返します。
soundex_difference(STRING1, STRING2)	Integer	2 つの文字列で SoundEx エンコードが同じ文字数を示す 0 から 4 の整数を返します。ここで、0 は類似性がないこと、また、4 は強い類似性があること、または同じ文字列であることを示します。

日付および時刻の関数

CLEM には、日付や時間を表す文字列変数の日付と時間ストレージのフィールドを操作する関数が用意されています。使用する日付と時間の書式は、ストリームごとに異なり、「ストリームのプロパティ」ダイアログ・ボックスで指定します。日付と時間の関数は、現在選択されている形式に従って、日付と時間の文字列を解析します。

日付に 2 桁だけを使用する (世紀を指定しない) 年を指定すると、IBM SPSS Modeler では、「ストリームのプロパティ」ダイアログ・ボックスで指定されているデフォルトの世紀が使用されます。

注: データ関数が SQL または IBM SPSS Analytic Server にプッシュバックされる場合、Analytic Server データ・ソースに続くブランチでは、そのデータ内の日付形式の文字列 (to_date) は、SPSS Modeler ストリームに指定されている日付形式に一致している必要があります。

表 32. CLEM 日付と時刻の関数 :

関数	結果	説明
@TODAY	<i>String</i>	「ストリームのプロパティ」ダイアログ・ボックスで「日/分をロールオーバー」を選択している場合、この関数は現在の日付形式を使用して、現在の日付を文字列として返します。2桁の日付形式を使用しており、「日/分をロールオーバー」を選択していない場合は、現在のサーバーの \$null\$ を返します。
to_time(ITEM)	<i>Time</i>	指定されたフィールドのストレージを時間に変換します。
to_date(ITEM)	<i>Date</i>	指定されたフィールドのストレージを日付に変換します。
to_timestamp(ITEM)	<i>Timestamp</i>	指定されたフィールドのストレージをタイムスタンプに変換します。
to_datetime(ITEM)	<i>Datetime</i>	指定されたフィールドのストレージを日付、時間またはタイムスタンプ値に変換します。
datetime_date(ITEM)	<i>Date</i>	数値、文字列またはタイムスタンプの日付値を返します数値（秒単位）を日付へ変換しなおすことができるのは、この関数だけです。ITEM が文字列の場合は、現在のデータ形式で文字列を解析することにより日付を作成します。この関数が正常に機能するためには、「ストリームのプロパティ」ダイアログ・ボックスの「日付の形式」に、正しい値が指定されていなければなりません。ITEM が数値の場合は、基準日（または紀元）からの秒数として解釈します。日付の端数は切り捨てられます。ITEM がタイムスタンプの場合は、日付をタイムスタンプの一部として返します。ITEM が日付の場合は、変更せずに返します。
date_before(DATE1, DATE2)	<i>Boolean</i>	DATE1 が DATE2 より前の日付またはタイムスタンプの場合に true を返します。それ以外の場合、この関数は 0 の値を返します。
date_days_difference(DATE1, DATE2)	<i>Integer</i>	日付またはタイムスタンプ DATE1 から日付またはタイムスタンプ DATE2 までの日数を整数で返します。DATE2 が DATE1 よりも前の場合、この関数は負の数値を返します。
date_in_days(DATE)	<i>Integer</i>	基準日から DATE で表される日付またはタイムスタンプまでの日数を整数で返します。DATE が基準日より前の場合、この関数は負の数値を返します。計算を適正に行うには、有効な日付を指定する必要があります。例えば、日付として 2001 年 2 月 29 日を指定することはできません。2001 年は閏年ではないので、この日付は存在しません。
date_in_months(DATE)	<i>Real</i>	基準日から DATE で表される日付またはタイムスタンプまでの月数を実数で返します。これは、1 か月を 30.4375 日と仮定した近似値になります。DATE が基準日より前の場合、この関数は負の数値を返します。計算を適正に行うには、有効な日付を指定する必要があります。例えば、日付として 2001 年 2 月 29 日を指定することはできません。2001 年は閏年ではないので、この日付は存在しません。

表 32. CLEM 日付と時刻の関数 (続き):

関数	結果	説明
date_in_weeks (DATE)	Real	基準日から DATE で表される日付またはタイムスタンプまでの週数を実数で返します。ここでは、1 週間を 7.0 日と仮定しています。DATE が基準日より前の場合、この関数は負の数値を返します。計算を適正に行うには、有効な日付を指定する必要があります。例えば、日付として 2001 年 2 月 29 日を指定することはできません。2001 年は閏年ではないので、この日付は存在しません。
date_in_years (DATE)	Real	基準日から DATE で表される日付またはタイムスタンプまでの年数を実数で返します。これは、1 年 365.25 日を基準とした近似値になります。DATE が基準日より前の場合、この関数は負の数値を返します。計算を適正に行うには、有効な日付を指定する必要があります。例えば、日付として 2001 年 2 月 29 日を指定することはできません。2001 年は閏年ではないので、この日付は存在しません。
date_months_difference (DATE1, DATE2)	Real	日付またはタイムスタンプ DATE1 から日付またはタイムスタンプ DATE2 までの月数を実数で返します。これは、1 か月を 30.4375 日と仮定した近似値になります。DATE2 が DATE1 よりも前の場合、この関数は負の数値を返します。
datetime_date (YEAR, MONTH, DAY)	Date	YEAR、MONTH、および DAY の日付値を作成します。引数は整数でなければなりません。
datetime_day (DATE)	Integer	指定された DATE またはタイムスタンプから、日付を返します。結果は 1 から 31 の範囲の整数になります。
datetime_day_name (DAY)	String	指定された DAY のフルネームを返します。引数は、1 (日曜) から 7 (土曜) の範囲の整数でなければなりません。
datetime_hour (TIME)	Integer	TIME またはタイムスタンプから時間を返します。結果は 0 から 23 の範囲の整数になります。
datetime_in_seconds (TIME)	Real	TIME に保存された秒の部分の部分を返します。
datetime_in_seconds (DATE)、 datetime_in_seconds (DATETIME)	Real	現在の DATE または DATETIME と基準日の間の差 (1900-01-01) から集計した数値を秒に変換して返します。
datetime_minute (TIME)	Integer	TIME またはタイムスタンプから分を返します。結果は 0 から 59 の範囲の整数になります。
datetime_month (DATE)	Integer	DATE またはタイムスタンプから月を返します。結果は 1 から 12 の範囲の整数になります。
datetime_month_name (MONTH)	String	指定された DAY のフルネームを返します。引数は、1 から 12 の範囲の整数でなければなりません。
datetime_now	Timestamp	現在の時刻をタイムスタンプとして返します。
datetime_second (TIME)	Integer	TIME またはタイムスタンプから秒を返します。結果は 0 から 59 の範囲の整数になります。
datetime_day_short_name (DAY)	String	DAY の名前を省略形で返します。引数は、1 (日曜) から 7 (土曜) の範囲の整数でなければなりません。
datetime_month_short_name (MONTH)	String	MONTH の名前を省略形で返します。引数は、1 から 12 の範囲の整数でなければなりません。
datetime_time (HOUR, MINUTE, SECOND)	Time	指定された HOUR、MINUTE、および SECOND の時間値を返します。引数は整数でなければなりません。
datetime_time (ITEM)	Time	ITEM の時間値を返します。

表 32. CLEM 日付と時刻の関数 (続き):

関数	結果	説明
<code>datetime_timestamp(YEAR, MONTH, DAY, HOUR, MINUTE, SECOND)</code>	<i>Timestamp</i>	与えられた YEAR、MONTH、DAY、HOUR、MINUTE、および SECOND のタイムスタンプ値を返します。
<code>datetime_timestamp(DATE, TIME)</code>	<i>Timestamp</i>	DATE および TIME のタイムスタンプ値を返します。
<code>datetime_timestamp (NUMBER)</code>	<i>Timestamp</i>	与えられた秒数のタイムスタンプ値を返します。
<code>datetime_weekday(DATE)</code>	<i>Integer</i>	指定されたDATE またはタイムスタンプから、曜日を返します。
<code>datetime_year(DATE)</code>	<i>Integer</i>	DATE またはタイムスタンプから年を返します。結果は 2002 のような整数になります。
<code>date_weeks_difference (DATE1, DATE2)</code>	<i>Real</i>	日付またはタイムスタンプ DATE1 から日付またはタイムスタンプ DATE2 までの週数を実数で返します。ここでは、1 週間を 7.0 日と仮定しています。DATE2 が DATE1 よりも前の場合、この関数は負の数値を返します。
<code>date_years_difference (DATE1, DATE2)</code>	<i>Real</i>	日付またはタイムスタンプ DATE1 から日付またはタイムスタンプ DATE2 までの年数を実数で返します。これは、1 年 365.25 日を基準とした近似値になります。DATE2 が DATE1 よりも前の場合、この関数は負の数値を返します。
<code>date_from_ywd(YEAR, WEEK, DAY)</code>	<i>Integer</i>	年、年の週、および曜日を、ISO 8601 標準を使用して日付に変換します。
<code>date_iso_day(DATE)</code>	<i>Integer</i>	ISO 8601 標準を使用して、日付の曜日を返します。
<code>date_iso_week(DATE)</code>	<i>Integer</i>	ISO 8601 標準を使用して、日付から年の週を返します。
<code>date_iso_year(DATE)</code>	<i>Integer</i>	ISO 8601 標準を使用して、日付から年を返します。
<code>time_before(TIME1, TIME2)</code>	<i>Boolean</i>	TIME1 が TIME2 より前の時間またはタイムスタンプの場合に真を返します。それ以外の場合、この関数は 0 の値を返します。
<code>time_hours_difference (TIME1, TIME2)</code>	<i>Real</i>	時間またはタイムスタンプ TIME1 と TIME2 間の時間差 (時間) を実数で返します。「ストリームのプロパティ」ダイアログ・ボックスで「日/分をロールオーバー」を選択している場合、TIME1 の値の方が大きいと、その値は前の日付を参照します。ロールオーバー・オプションをオンにしていなかった場合、TIME1 の値の方が大きいと、返される値は負になります。
<code>time_in_hours(TIME)</code>	<i>Real</i>	TIME で表される時間を実数で返します。例えば、時間の形式 HHMM では、式 <code>time_in_hours('0130')</code> は 1.5 として評価されます。TIME は時間またはタイムスタンプを示します。
<code>time_in_mins(TIME)</code>	<i>Real</i>	TIME で表される分を実数で返します。TIME は時間またはタイムスタンプを示します。
<code>time_in_secs(TIME)</code>	<i>Integer</i>	TIME で表される秒を整数で返します。TIME は時間またはタイムスタンプを示します。

表 32. CLEM 日付と時刻の関数 (続き):

関数	結果	説明
time_mins_difference(TIME1, TIME2)	Real	時間またはタイムスタンプ <i>TIME1</i> と <i>TIME2</i> 間の時間差 (分) を実数で返します。「ストリームのプロパティ」ダイアログ・ボックスで「日/分をロールオーバー」を選択している場合、 <i>TIME1</i> の値の方が大きいと、その値は前の日 (または、現在の形式で分と秒だけが指定されている場合は前の時間) を参照します。ロールオーバー・オプションをオンにしていない場合、 <i>TIME1</i> の値の方が大きいと、返される値は負になります。
time_secs_difference(TIME1, TIME2)	Integer	<i>TIME1</i> と <i>TIME2</i> の時間またはタイムスタンプの差異を秒で返します。「ストリームのプロパティ」ダイアログ・ボックスで「日/分をロールオーバー」を選択している場合、 <i>TIME1</i> の値の方が大きいと、その値は前の日 (または、現在の形式で分と秒だけが指定されている場合は前の時間) を参照します。ロールオーバー・オプションをオンにしていない場合、 <i>TIME1</i> の値の方が大きいと、返される値は負になります。

日付と時刻の値の変換:

変換関数および日付や時刻の値のような、入力に特別な型が必要なその他の関数は、「ストリームのオプション」ダイアログ・ボックスに指定されている現在の形式に依存します。例えば、*DATE* という名前のフィールド名がある場合は、*Jan 2003*、*Feb 2003* などの値で文字列として保存されており、次のように日付ストレージへ変換できます。

`to_date(DATE)`

この返還を行うには、ストリームのデフォルト日付形式として、一致する日付形式 **MON YYYY** を選択します。

フィルター・ノードを使用して文字列値を日付へ変換する例については、*streams* サブフォルダー内の *¥Demos* フォルダーにインストールされている、ストリーム *broadband_create_models.str* を参照してください。

数値として保存される日付: 上記の例の *DATE* がフィールド名であるのに対して、`to_date` は CLEM 関数であることに注意してください。数値として保存された日付がある場合は、数値が基準日 (または紀元) からの秒数として解釈される `datetime_date` 関数を用いることでそれらを変換できます。

`datetime_date(DATE)`

日付を秒数へ (および逆) 変換することで、次のように現在の日付に一定の日数をプラス、マイナスするといった計算を実行できます。

`datetime_date((date_in_days(DATE)-7)*60*60*24)`

シーケンス関数

一部の演算子では、イベントのシーケンス (順序) が重要になります。アプリケーションで使用できるレコード・シーケンスは、以下のとおりです。

- シーケンスと時系列
- シーケンス関数

- レコード・インデックスの作成
- 値の平均、合計、および比較
- 変化の把握 (差分)
- @SINCE
- オフセット値
- その他のシーケンス機能

多くのアプリケーションでは、ストリームを通過している各レコードは、それぞれ個別で、他のすべてのレコードから独立したものと見なされます。通常、このような場合は、レコードの順序は重要ではありません。

ただし、問題によっては、レコード・シーケンスが非常に重要になります。特に時系列の場合がそうで、レコードのシーケンスは、イベントまたは発生の順序、すなわちシーケンスを表します。各レコードは、特定の瞬間のスナップショットを示します。しかし、最も重要な情報は、瞬間的な値にあるのではなく、このような値が時間の経過に伴ってどのように変化し、動いていくのかということにあるのです。

もちろん、該当するパラメーターが時間以外のものであってもかまいません。例えば、レコードが、線からの距離について実行される分析を示している場合でも、同じ原則が適用されます。

シーケンスおよび特殊関数は、次の特徴によってすぐに判別できます。

- 関数名の最初に @ が付いている。
- 関数名が大文字である。

シーケンス関数は、ノードによって現在処理中のレコード、すでにノードを通過したレコード、あるいはまだノードに到達していないレコードを参照します。シーケンス関数は、CLEM 式の他の要素と自由に組み合わせ使用できますが、引数としての使用を制約されているものもあります。

例

ある事象が発生してから、またはある条件が真になってからの長さを知りたい場合があります。その場合は、次のように @SINCE 関数を使用します。

```
@SINCE(Income > Outgoings)
```

この関数は、指定した条件が真 (true) であった最後のレコードのオフセットを返します。つまり、指定した条件が真 (true) であった最後のレコード以前のレコード数を返します。指定した条件が一度も真 (true) にならなかった場合、@SINCE は @INDEX + 1 を返します。

@SINCE で使用される式の現在のレコードの値を参照したいこともあるでしょう。関数 @THIS を使用して、フィールド名が常に現在のレコードに適用されるように指定します。Concentration フィールドの値が、現在のレコードの 2 倍より大きい最後のレコードのオフセットを調べるには、次のように記述します。

```
@SINCE(Concentration > 2 * @THIS(Concentration))
```

定義により、現在のレコードに対して真 (true) である条件を @SINCE に指定する場合があります。次に例を示します。

```
@SINCE(ID == @THIS(ID))
```

この場合、@SINCE は現在のレコードに対して条件を評価しません。前のレコードと現在のレコードに対して条件を評価する場合は、同様な関数 @SINCE0 を使用します。現在のレコードで条件が真 (true) の場合は、@SINCE0 は 0 を返します。

表 33. CLEM シーケンス関数

関数	結果	説明
MEAN(FIELD)	Real	指定された <i>FIELD</i> または <i>FIELDS</i> に対して、値の平均値を返します。
@MEAN(FIELD, EXPR)	Real	現在のレコードを含めて、現在のノードが受け取った最後の <i>EXPR</i> レコードまでの、 <i>FIELD</i> の値の平均値を返します。 <i>FIELD</i> は数値型フィールドの名前でなければなりません。 <i>EXPR</i> には、0 より大きい整数として評価される任意の式を使用できます。 <i>EXPR</i> を省略した場合、または <i>EXPR</i> の値が現在までに受け取ったレコード数を超過している場合は、現在までに受け取ったすべてのレコードの平均が返されます。
@MEAN(FIELD, EXPR, INT)	Real	現在のレコードを含めて、現在のノードが受け取った最後の <i>EXPR</i> レコードまでの、 <i>FIELD</i> の値の平均値を返します。 <i>FIELD</i> は数値型フィールドの名前でなければなりません。 <i>EXPR</i> には、0 より大きい整数として評価される任意の式を使用できます。 <i>EXPR</i> を省略した場合、または <i>EXPR</i> の値が現在までに受け取ったレコード数を超過している場合は、現在までに受け取ったすべてのレコードの平均が返されます。 <i>INT</i> には、参照する値の最大数を指定します。この方法は、2 つの引数だけを使用するよりも効率的です。
@DIFF1(FIELD)	Real	<i>FIELD</i> の最初の差分を返します。したがって、1 つの引数を指定する形式では、単純にフィールドの現在値と前の値の差分を返します。前に関連するレコードが存在しない場合は、 \$null\$ を返します。
@DIFF1(FIELD1, FIELD2)	Real	2 つの引数を指定する形式では、 <i>FIELD2</i> に関する <i>FIELD1</i> の最初の差分を返します。前に関連するレコードが存在しない場合は、 \$null\$ を返します。それは、 $@DIFF1(FIELD1)/@DIFF1(FIELD2)$ で算出されます。
@DIFF2(FIELD)	Real	<i>FIELD</i> の 2 番目の差分を返します。したがって、1 つの引数を指定する形式では、単純にフィールドの現在値と前の値の差分を返します。前に関連するレコードが存在しない場合は、 \$null\$ を返します。 <i>@DIFF2</i> は $@DIFF(@DIFF(FIELD))$ で算出されます。
@DIFF2(FIELD1, FIELD2)	Real	2 つの引数を指定する形式では、 <i>FIELD2</i> に関する <i>FIELD1</i> の 2 番目の差分を返します。前に関連するレコードが存在しない場合は、 \$null\$ を返します。これは $@DIFF1(FIELD1)/@DIFF1(FIELD2) - @OFFSET(@DIFF1(FIELD1), 1)/@OFFSET(@DIFF1(FIELD2)) / @DIFF1(FIELD2)$ という複雑な計算です。
@INDEX	Integer	現在のレコードのインデックスを返します。インデックスは、レコードが現在のノードに到達したときにレコードに対して割り振られます。最初のレコードにはインデックス 1 が与えられます。インデックスは、その後の各レコードに対して 1 ずつ増やされます。

表 33. CLEM シーケンス関数 (続き)

関数	結果	説明
@LAST_NON_BLANK(FIELD)	Any	上流の入力ノードまたはデータ型ノードで定義されるように、空白でない <i>FIELD</i> の最後の値を返します。それまでに読み込んだレコードの <i>FIELD</i> の値がすべて空白である場合は、\$null\$ を返します。ユーザー欠損値とも呼ばれる空白値は、各フィールドに個別に定義することができることに注意してください。
@MAX(FIELD)	Number	指定された <i>FIELD</i> の最大値を返します。
@MAX(FIELD, EXPR)	Number	現在のレコードを含めて、現在までに受け取った過去 <i>EXPR</i> レコードの <i>FIELD</i> の最大値を返します。 <i>FIELD</i> は数値型フィールドの名前でなければなりません。 <i>EXPR</i> は、0 より大きい整数を評価する任意の式を使用できます。
@MAX(FIELD, EXPR, INT)	Number	現在のレコードを含めて、現在までに受け取った過去 <i>EXPR</i> レコードの <i>FIELD</i> の最大値を返します。 <i>FIELD</i> は数値型フィールドの名前でなければなりません。 <i>EXPR</i> は、0 より大きい整数を評価する任意の式を使用できます。 <i>EXPR</i> を省略した場合、または <i>EXPR</i> の値が、現在までに受け取ったレコード数を超過している場合は、現在までに受け取ったすべてのレコードの最大値が返されます。 <i>INT</i> には、参照する値の最大数を指定します。この方法は、2 つの引数だけを使用するよりも効率的です。
@MIN(FIELD)	Number	指定された <i>FIELD</i> の最小値を返します。
@MIN(FIELD, EXPR)	Number	現在のレコードを含めて、現在までに受け取った過去 <i>EXPR</i> レコードの <i>FIELD</i> の最小値を返します。 <i>FIELD</i> は数値型フィールドの名前でなければなりません。 <i>EXPR</i> は、0 より大きい整数を評価する任意の式を使用できます。
@MIN(FIELD, EXPR, INT)	Number	現在のレコードを含めて、現在までに受け取った過去 <i>EXPR</i> レコードの <i>FIELD</i> の最小値を返します。 <i>FIELD</i> は数値型フィールドの名前でなければなりません。 <i>EXPR</i> は、0 より大きい整数を評価する任意の式を使用できます。 <i>EXPR</i> を省略した場合、または <i>EXPR</i> の値が、現在までに受け取ったレコード数を超過している場合は、現在までに受け取ったすべてのレコードの最小値が返されます。 <i>INT</i> には、参照する値の最大数を指定します。この方法は、2 つの引数だけを使用するよりも効率的です。

表 33. CLEM シーケンス関数 (続き)

関数	結果	説明
@OFFSET(FIELD, EXPR)	Any	<p>現在のレコードから <i>EXPR</i> で指定された値のオフセットにあるレコードの <i>FIELD</i> の値を返します。正のオフセットがすでに通過したレコードを参照 (「後読み」) するのに対し、負のオフセットはまだ到着していないレコードに「先読み」を指定します。例えば、@OFFSET(Status, 1) は、前のレコードの Status フィールドの値を返します。一方、@OFFSET(Status, -4) は、値を取得するためにシーケンス内で 4 個先のレコードを (つまり、このノードをまだ通過していないレコードまで) 「先読み」します。負の (先読み) オフセットは、定数として指定する必要があります。負のオフセットに限っては、<i>EXPR</i> も任意の CLEM 式であり、現在のレコードに対してオフセットを与えるために評価されます。この場合、性能を改善するために、この関数の引数が 3 個のバージョンを使用する必要があります (次の関数を参照)。この式が、負ではない整数以外の値を返す場合、エラーになります。つまり、計算された先読みオフセットは正しく取得できません。</p> <p>注：自己参照の @OFFSET 関数では、リテラルの先読みは使用できません。例えば、置換ノードでは、field1 の値を @OFFSET(field1,-2) のような式を使用して置換できません。</p> <p>注: 置換ノードでは、フィールドを置換する際に、そのフィールドの値として事実上 2 つの異なる値 (具体的には、置換前の値と置換後の値) が存在します。@OFFSET は、それ自体を参照する場合は置換後の値を参照します。この置換後の値は過去の行にのみ存在するため、自己参照の @OFFSET は過去の行のみを参照できます。自己参照 @OFFSET は将来を参照できないため、オフセットについて以下の検査を実行します。</p> <ul style="list-style-type: none"> • オフセットがリテラルであり、将来を参照している場合、実行開始前にエラーが報告されます。 • オフセットが式であり、ランタイムに評価結果が将来になる場合、@OFFSET によって \$null\$ が返されます。 <p>注: 1 つのノード内で「先読み」と「後読み」の両方を使用することはサポートされていません。</p>

表 33. CLEM シーケンス関数 (続き)

関数	結果	説明
@OFFSET(FIELD, EXPR, INT)	Any	<p>@OFFSET 関数と同じ演算を行います。3 番目の引数 <i>INT</i> に、前方参照する値の最大数を指定することができます。オフセットを式から計算することができる場合、性能を改善するために、この 3 番目の引数を使用する必要があります。</p> <p>例えば、@OFFSET(Foo, Month, 12) のような式では、システムが Foo の最後の 12 個の値だけを保持していればよいと判断できます。3 番目の引数がなければ、安全のため、すべての値を保管しておく必要があります。オフセットの値が定数の場合 (定数である必要がある負の「先読み」オフセットの場合を含む)、3 番目の引数は無意味なので、この関数の 2 引数バージョンを使用してください。前述の 2 引数バージョンの自己参照関数に関する注意事項を参照してください。</p> <p>注: 1 つのノード内で「先読み」と「後読み」の両方を使用することはサポートされていません。</p>
@SDEV(FIELD)	Real	指定された <i>FIELD</i> または <i>FIELDS</i> に対して、値の平均値を返します。
@SDEV(FIELD, EXPR)	Real	現在のレコードを含めて、現在のノードが受け取った最後の <i>EXPR</i> レコードまでの、 <i>FIELD</i> の値の標準偏差を返します。 <i>FIELD</i> は数値型フィールドの名前でなければなりません。 <i>EXPR</i> は、0 より大きい整数を評価する任意の式を使用できます。 <i>EXPR</i> を省略した場合、または <i>EXPR</i> の値が、現在までに受け取ったレコード数を超えている場合は、現在までに受け取ったすべてのレコードの標準偏差を返します。
@SDEV(FIELD, EXPR, INT)	Real	現在のレコードを含めて、現在のノードが受け取った最後の <i>EXPR</i> レコードまでの、 <i>FIELD</i> の値の標準偏差を返します。 <i>FIELD</i> は数値型フィールドの名前でなければなりません。 <i>EXPR</i> は、0 より大きい整数を評価する任意の式を使用できます。 <i>EXPR</i> を省略した場合、または <i>EXPR</i> の値が、現在までに受け取ったレコード数を超えている場合は、現在までに受け取ったすべてのレコードの標準偏差を返します。 <i>INT</i> には、参照する値の最大数を指定します。この方法は、2 つの引数だけを使用するよりも効率的です。
@SINCE(EXPR)	Any	任意の CLEM 式が真 (true) の場合に、 <i>EXPR</i> から過ぎたレコード数を返します。
@SINCE(EXPR, INT)	Any	2 番目の引数 <i>INT</i> には、前方参照するレコードの最大数を指定します。 <i>EXPR</i> が一度も真 (true) にならなかった場合、 <i>INT</i> は @INDEX+1 です。
@SINCE0(EXPR)	Any	現在のレコードも考慮します。一方、@SINCE は現在のレコードは考慮しません。@SINCE0 は、現在のレコードについて <i>EXPR</i> が真 (true) の場合に 0 を返します。
@SINCE0(EXPR, INT)	Any	2 番目の引数 <i>INT</i> には、前方参照するレコードの最大数を指定します。

表 33. CLEM シーケンス関数 (続き)

関数	結果	説明
@SUM(FIELD)	Number	指定された <i>FIELD</i> または <i>FIELDS</i> に対して、値の合計値を返します。
@SUM(FIELD, EXPR)	Number	現在のレコードを含めて、現在のノードが受け取った最後の <i>EXPR</i> レコードまでの、 <i>FIELD</i> の値の合計値を返します。 <i>FIELD</i> は数値型フィールドの名前でなければなりません。 <i>EXPR</i> には、0 より大きい整数として評価される任意の式を使用できます。 <i>EXPR</i> を省略した場合、または <i>EXPR</i> の値が現在までに受け取ったレコード数を超えている場合は、現在までに受け取ったすべてのレコードの合計が返されます。
@SUM(FIELD, EXPR, INT)	Number	現在のレコードを含めて、現在のノードが受け取った最後の <i>EXPR</i> レコードまでの、 <i>FIELD</i> の値の合計値を返します。 <i>FIELD</i> は数値型フィールドの名前でなければなりません。 <i>EXPR</i> には、0 より大きい整数として評価される任意の式を使用できます。 <i>EXPR</i> を省略した場合、または <i>EXPR</i> の値が現在までに受け取ったレコード数を超えている場合は、現在までに受け取ったすべてのレコードの合計が返されます。 <i>INT</i> には、参照する値の最大数を指定します。この方法は、2 つの引数だけを使用するよりも効率的です。
@THIS(FIELD)	Any	現在のレコードの <i>FIELD</i> で指定された名前のフィールドの値を返します。@SINCE 式でのみ使用されます。

グローバル関数

関数 @MEAN、@SUM、@MIN、@MAX、および @SDEV は、最も広範囲の場合、現在までに読み取られたすべてのレコード (現在のレコードを含む) で機能します。しかし、現在のレコードの値とデータセット全体での値とを比較できると便利な場合もあります。グローバル・ノードを使用してデータ・セット全体の値を生成したら、CLEM 式でグローバル関数を使用してこれらの値にアクセスすることができます。

例えば、

@GLOBAL_MAX(Age)

は、データセット内で最も大きい Age の値を返します。一方、

$(Value - @GLOBAL_MEAN(Value)) / @GLOBAL_SDEV(Value)$

は、このレコードの Value とグローバル平均との差を標準偏差として示します。グローバル・ノードによりグローバル値が算出されないと、グローバル値を使用することはできません。現在のすべてのグローバル値は、「ストリームのプロパティ」ダイアログ・ボックスの「グローバル」タブにある「グローバル値の消去」ボタンをクリックしてキャンセルすることができます。

表 34. CLEM グローバル関数 :

関数	結果	説明
@GLOBAL_MAX(FIELD)	Number	以前にグローバル・ノードで生成されたように、データ・セット全体の FIELD の最大値を返します。FIELD は、数値型フィールド、日付/時刻/日時フィールド、または文字列フィールドの名前でなければなりません。対応するグローバル値が設定されていない場合は、エラーが発生します。
@GLOBAL_MIN(FIELD)	Number	以前にグローバル・ノードで生成されたように、データ・セット全体の FIELD の最小値を返します。FIELD は、数値型フィールド、日付/時刻/日時フィールド、または文字列フィールドの名前でなければなりません。対応するグローバル値が設定されていない場合は、エラーが発生します。
@GLOBAL_SDEV(FIELD)	Number	以前にグローバル値の設定ノードで生成されたデータ・セット全体の FIELD の値の標準偏差を返します。FIELD は数値型フィールドの名前でなければなりません。対応するグローバル値が設定されていない場合は、エラーが発生します。
@GLOBAL_MEAN(FIELD)	Number	以前にグローバル・ノードで生成されたように、データ・セット全体の FIELD の平均値を返します。FIELD は数値型フィールドの名前でなければなりません。対応するグローバル値が設定されていない場合は、エラーが発生します。
@GLOBAL_SUM(FIELD)	Number	以前にグローバル・ノードで生成されたように、データ・セット全体の FIELD の値の合計を返します。FIELD は数値型フィールドの名前でなければなりません。対応するグローバル値が設定されていない場合は、エラーが発生します。

空白値とヌル値処理関数

CLEM を使用して、フィールド内の特定の値を「空白」、つまり欠損値と見なすように指定することができます。空白値を処理する関数を次に示します。

表 35. CLEM の空白値関数とヌル値関数 :

関数	結果	説明
@BLANK(FIELD)	Boolean	上流のデータ型ノードまたは入力ノードで設定された空白処理規則 (「データ型」タブ) にしたがって、値が空白のレコードに対して真を返します。
@LAST_NON_BLANK(FIELD)	Any	上流の入力ノードまたはデータ型ノードで定義されるように、空白でない FIELD の最後の値を返します。それまでに読み込んだレコードの FIELD の値がすべて空白である場合は、\$null\$ を返します。ユーザー欠損値とも呼ばれる空白値は、各フィールドに個別に定義することができることに注意してください。
@NULL(FIELD)	Boolean	FIELD の値がシステム欠損値 \$null\$ の場合に true を返します。ユーザー定義の空白値など、他のすべての値については false を返します。両方を確認する場合は、@BLANK(FIELD) および @NULL(FIELD) を使用します。
undef	Any	一般的に、CLEM で \$null\$ 値を入力するために用いられます。例えば、置換ノードで空白値にヌルを挿入するために用いられます。

空白フィールドが、置換ノードで書き込まれる場合もあります。置換ノードおよびフィールド作成ノード (複数モードの場合) の両方で、特殊 CLEM 関数の @FIELD は、調査対象の現在のフィールドを表します。

特殊フィールド

特殊関数は、調査対象の特定のフィールドを表したり、フィールドのリストを入力として生成したりするために用いられます。例えば、複数のフィールドを一度に作成する場合、@FIELD を使用して「このフィールド作成操作を選択したフィールドに対して行う」ことを指示します。式 $\log(@FIELD)$ を使用すると、選択した各フィールドに対して、新しいログ・フィールドが作成されます。

表 36. CLEM 特殊フィールド :

関数	結果	説明
@FIELD	Any	式のコンテキスト中に指定されているすべてのフィールドに対して処理を行います。
@TARGET	Any	ユーザー定義の分析関数で CLEM 式を使用する場合、@TARGET は対象のフィールドを表すか、または分析される対象と予測のペアに対する「正しい値」を表します。通常この関数は、精度分析ノードで使用されます。
@PREDICTED	Any	ユーザー定義の分析関数で CLEM 式を使用する場合、@PREDICTED は、分析される対象と予測のペアに対して、予測される値を表します。通常この関数は、精度分析ノードで使用されます。
@PARTITION_FIELD	Any	現在のデータ区分フィールドの名前を置き換えます。
@TRAINING_PARTITION	Any	現在の学習用データ区分の値を返します。例えば、条件抽出ノードを使用して学習レコードを選択するには、CLEM 式 ($@PARTITION_FIELD = @TRAINING_PARTITION$) を使用します。この方法は、データ内の各データ区分を表すためにどの値が使用されているかに関係なく、常に条件抽出ノードが正しく動作することを保証します。
@TESTING_PARTITION	Any	現在のテスト用データ区分の値を返します。
@VALIDATION_PARTITION	Any	現在の検証用データ区分の値を返します。
@FIELDS_BETWEEN(start, end)	Any	データ中のフィールドの普通の順序 (つまり、挿入) にもとづく、指定された開始フィールドと最終フィールドの間 (開始、最終フィールドを含む) のフィールド名のリストを返します。

表 36. CLEM 特殊フィールド (続き) :

関数	結果	説明
@FIELDS_MATCHING(pattern)	Any	<p>指定したパターンに一致するフィールド名のリストを返します。クエスチョン・マーク (?) をパターンに含めて正確に 1 つの文字に一致させることができます。アスタリスク (*) は 0 かそれ以上の文字数に一致します。リテラル・クエスチョン・マークまたはアスタリスクを (むしろ、ワイルドカードとして使用しないで) 一致させるために、バックスラッシュをエスケープ文字として使用することができます。</p> <p>注: 引数として文字列リテラルを使用する必要があります。引数を生成するために、入れ子になった式を使用することはできません。</p>
@MULTI_RESPONSE_SET	Any	<p>名前の付いた複数回答セットでフィールドのリストを返します。</p>

第 10 章 リポジトリでの IBM SPSS Modeler の使用

IBM SPSS Collaboration and Deployment Services Repository について

SPSS Modeler は IBM SPSS Collaboration and Deployment Services リポジトリと連動して使用でき、データ・マイニング・モデルや関連する予測オブジェクトのライフサイクルの管理し、これらのオブジェクトをエンタープライズ アプリケーション、ツール、ソリューションで使用することができます。このように共有できる IBM SPSS Modeler オブジェクトには、ストリーム、ノード、ストリーム出力、プロジェクト、およびモデルがあります。オブジェクトは中央リポジトリに保管され、ここでは、アプリケーションと共有したり、拡張されたバージョン管理、メタデータ、および検索機能を使用して追跡することができます。

SPSS Modeler をリポジトリで使用する前に、リポジトリ・ホストにアダプターをインストールする必要があります。このアダプターがない場合、特定の SPSS Modeler ノードまたはモデルからリポジトリのオブジェクトにアクセスしようとすると、次のメッセージが表示されます。

The repository may need updating to support new node, model and output types.

アダプターのインストールに関する詳細は、製品ダウンロードの一部として PDF ファイルで入手可能な「SPSS Modeler 展開インストール・ガイド」を参照してください。IBM SPSS Deployment Manager から IBM SPSS Modeler リポジトリ・オブジェクトにアクセスする方法については、「SPSS Modeler 展開ガイド」を参照してください。

次のセクションには、SPSS Modeler 内のリポジトリへのアクセスについての情報が記載されています。

拡張バージョン管理と検索のサポート

リポジトリには、包括的なオブジェクトのバージョン管理および検索機能があります。例えば、ストリームを作成して、そのストリームを リポジトリに格納して、他の部門の研究者と共有する場合を考えます。後で、SPSS Modeler のストリームを更新する場合、以前のバージョンを上書きしないで、更新されたバージョンをリポジトリに追加できます。すべてのバージョンが、アクセス可能な状態で維持され、名前、ラベル、使用されているフィールドまたはその他の属性で検索できます。例えば、純収益を入力フィールドと使用しているすべてのモデルのバージョンや、特定の作成者により作成されたすべてのモデルを検索できます(この処理を従来のファイル・システムで行うと、各バージョンごとに異なるファイル名で保存する必要があります。またバージョン間の関連がソフトウェアから認識できない場合があります)。

シングル・サインオン

シングル・サインオン機能を使用すると、接続するたびにユーザー名やパスワードを入力することなくリポジトリに接続することができます。ユーザーの既存のローカル・ネットワーク・ログインの詳細では、IBM SPSS Collaboration and Deployment Services への必要な認証が設定されています。この機能は、以下によって異なります。

- IBM SPSS Collaboration and Deployment Services を、シングル・サインオン・プロバイダを使用するよう設定する必要がある
- プロバイダと互換性のあるホストにログインする必要がある

詳しくは、164 ページの『リポジトリへの接続』を参照してください。

リポジトリ・オブジェクトの保存と展開

IBM SPSS Modeler で作成されるストリームは、拡張子 `.str` の付いたファイルとして、そのままリポジトリに保存できます。このように、全社規模の複数のユーザーが単一のストリームにアクセスできます。詳しくは、トピック 165 ページの『リポジトリのオブジェクトの保存』を参照してください。

リポジトリでストリームを展開することもできます。展開ストリームは、追加メタデータのあるファイルとして保存されます。展開ストリームは、自動化スコアリングやモデル更新など、IBM SPSS Collaboration and Deployment Services のエンタープライズ レベルの機能を十分に活用できます。例えば、新規データが利用可能になると、モデルを定期的なスケジュールで自動的に更新することができます。またあるいは一連のストリームを Champion-Challenger 分析を目的として展開することができ、ストリームを比較して最も効果的な世族モデルを含むシナリオを決定します。

注: アソシエーション・ルール、STP、TCM の各モデル作成ノードは、IBM SPSS Collaboration and Deployment Services のモデル評価または Champion Challenger のステップをサポートしません。

ストリーム (拡張子 `.str`) を展開できます。ストリームとして展開すると、ストリームをシンクライアント・アプリケーション IBM SPSS Modeler Advantage によって使用できます。詳しくは、トピック 183 ページの『IBM SPSS Modeler Advantage でストリームを開く』を参照してください。

詳しくは、177 ページの『ストリーム展開のオプション』を参照してください。

その他の展開のオプション

IBM SPSS Collaboration and Deployment Services は、企業コンテンツ管理に次のような広範の機能を提供し、ストリームを展開またはエクスポートするための多くのメカニズムを利用することができます。

- 後で IBM SPSS Modeler Solution Publisher ランタイムで利用するために、ストリームとモデルをエクスポートします。
- 1 つ以上のモデルを、モデル情報をエンコードするための XML 形式の PMML としてエクスポートします。詳しくは、トピック 184 ページの『PMML としてのモデルのインポートおよびエクスポート』を参照してください。

リポジトリへの接続

1. リポジトリに接続するには、IBM SPSS Modeler メニューから次の項目をクリックします。

「ツール」 > 「リポジトリ」 > 「オプション...」

2. 「リポジトリ URL」フィールドで、アクセスしたいリポジトリ・インストール環境のディレクトリー・パス (つまり、URL) を入力または選択します。一度に接続できるレポジトリは 1 つだけです。

設定は、各サイトまたはインストール環境によって変わってきます。特定のログイン詳細については、各サイトのシステム管理者に問い合わせてください。

「資格情報の設定」。このボックスのチェック・マークを外した状態にして、シングル・サインオン機能を有効にします。これにより、ローカル・コンピューターのユーザー名およびパスワードの詳細を使用してログインを試みます。シングル・サインオンを使用できない場合、またはこのボックスをチェックしてシングル・サインオンを無効にした場合 (例えば、管理者アカウントにログインした場合)、資格情報を入力するための詳細画面が表示されます。

リポジトリの資格情報の入力

設定によっては、「リポジトリ: 資格情報」ダイアログ・ボックス内で以下のフィールドを指定する必要があります。

「ユーザー ID とパスワード」。ログオンに使用する有効なユーザー名とパスワードを指定します。必要があれば、詳細についてローカル・システムの管理者にお問い合わせください。

「プロバイダー」。認証用のセキュリティー・プロバイダーを選択します。異なるセキュリティー・プロバイダーを使用するよう、リポジトリを設定することができます。必要な場合、詳しい情報を各サイトの管理者にお問い合わせください。

「リポジトリおよびユーザー ID を記憶」。毎回、接続時に同じ情報を再入力しなくても済むように、現在の設定をデフォルトとして保存します。

リポジトリ資格情報の参照

Analytic Server、Cognos、ODBC、または TM1 の各入力ノードからリポジトリに接続する場合、以前に記録した資格情報を選択してリポジトリに接続できます。そのような資格情報は「リポジトリ資格情報の選択」ダイアログ・ボックスにリストされます。このダイアログ・ボックスを選択するには、「資格情報」フィールドの横にある「参照」をクリックします。

「リポジトリ資格情報の選択」ダイアログ・ボックスに表示されたリストで資格情報を強調表示して「OK」をクリックします。リストが大き過ぎる場合は、「フィルター」フィールドを使用して名前またはその一部を入力し、必要な資格情報を検索します。

リポジトリ・コンテンツを参照

リポジトリにより、Windows Explorer のように格納した内容を参照することができます。また格納された各オブジェクトのバージョンも参照することができます。

1. IBM SPSS Collaboration and Deployment Services Repository ウィンドウを開くには、SPSS Modeler のメニューで次の各項目をクリックします。

「ツール」 > 「リポジトリ」 > 「探索...」

1. 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者にお問い合わせください。

エクスプローラ ウィンドウには、最初にフォルダー階層のツリー・ビューが表示されます。フォルダー名をクリックしてコンテンツを表示します。

現在の選択または検索基準に一致するオブジェクトが右側の領域にリストされ、選択されているバージョンの詳細な情報が右下の領域に表示されます。表示される属性値は、最も新しいバージョンに適用されます。

リポジトリのオブジェクトの保存

リポジトリにストリーム、ノード、モデル、モデル・パレット、プロジェクト、出力オブジェクトを格納できます。他のユーザーやアプリケーションはそれらにアクセスすることができます。

ストリーム出力をリポジトリに、その他のユーザーが IBM SPSS Collaboration and Deployment Services Deployment Portal を使用してインターネットで表示できる形式で公開することもできます。

オブジェクトの特性の設定

オブジェクトを保存する場合、「リポジトリ: 保存」ダイアログボックスが表示され、オブジェクトのさまざまなプロパティの値を設定できます。以下を行うことができます。

- オブジェクトを保存する名前およびリポジトリ・フォルダーを選択する
- バージョン・ラベルやその他の検索可能なプロパティ例など、オブジェクトに関する情報を追加する
- 1 つまたは複数の分類トピックをオブジェクトに割り当てる
- オブジェクトのセキュリティー・オプションを設定する

次の項では、設定できるプロパティについて説明します。

オブジェクトを保存する場所の選択

「リポジトリ: 保存」ダイアログ・ボックスで、以下を入力します。

ファイルの保存場所: オブジェクトが保存される場所、現在のフォルダーが表示されます。リストでフォルダー名をダブルクリックすると、そのフォルダーが現在のフォルダーに設定されます。親フォルダーを選択するには、「上位フォルダー・ボタン」を使用します。現在のレベルにフォルダーを作成するには、「新規フォルダー」ボタンを使用します。

ファイル名: オブジェクトを保存する名前。

保存: オブジェクトを現在の場所に保存します。

保存するオブジェクトに関する情報の追加

「リポジトリ: 保存」ダイアログ・ボックスの「情報」タブ上のフィールドは、いずれもオプションです。

「作成者」: リポジトリのオブジェクトを作成したユーザーのユーザー名。デフォルトでは、リポジトリ接続に使用されたユーザー名が表示されますが、ここでこの名前を変更できます。

「バージョン・ラベル」: オブジェクトのバージョンを示すためにリストからラベルを選択するか、「追加」をクリックして、新しいラベルを作成します。ラベルには「[」文字を使用しないでください。このオブジェクト・バージョンにラベルを割り当てない場合は、チェック・ボックスがオンになっていないことを確認してください。詳しくは、トピック 175 ページの『オブジェクトのプロパティの表示および編集』を参照してください。

説明: オブジェクトの説明です。説明によって、オブジェクトを検索できます (注意事項を参照)。

キーワード: オブジェクトに関連し、検索に使用できる 1 つまたは複数のキーワード (注記を参照)。

有効期限: その日以降、オブジェクトが一般ユーザーに表示されなくなる日付。ただし、この日以降も所有者やリポジトリ管理者には表示されます。有効期限の日付を設定するには、「日付」オプションを選択して日付を入力するか、カレンダーのボタンを使用して日付を選択します。

保存: オブジェクトを現在の場所に保存します。

注: 「説明」フィールドおよび「キーワード」フィールドの情報は、オブジェクトの「注釈」タブの SPSS Modeler に入力したものと異なるものとして扱われます。説明またはキーワードによるリポジトリ検索では、「注釈」タブからの情報は返されません。詳しくは、トピック 172 ページの『リポジトリ内のオブジェクトの検索』を参照してください。

保存したオブジェクトへのトピックの割り当て

トピックは、リポジトリに保存されているコンテンツの階層分類システムです。オブジェクトを保存するときに使用可能なトピックから選択でき、またユーザーはトピックによってオブジェクトを検索できます。使用できるトピックのリストは、適切な権限を持つリポジトリ・ユーザーによって設定されます (詳細は、『*Deployment Manager ユーザー・ガイド*』を参照してください)。

「リポジトリ: 保存」ダイアログ・ボックスの「トピック」タブで、トピックをオブジェクトに割り当てる手順は、次のとおりです。

1. 「追加」ボタンをクリックします。
2. 使用できるトピックのリストのトピック名をクリックします。
3. 「OK」をクリックします。

トピックの割り当てを削除する手順は、次のとおりです。

4. 割り当てたトピックのリストでトピックを選択します。
5. 「削除」をクリックします。

保存したオブジェクトのセキュリティー・オプションを設定

「リポジトリ: 保存」ダイアログ・ボックスの「セキュリティー」タブで、保存したオブジェクトのさまざまなセキュリティー・オプションを設定または変更できます。1 つまたは複数の主成分 (ユーザーまたはユーザーのグループ) について、次のことを実行できます。

- オブジェクトへのアクセス権限を割り当てる
- オブジェクトへのアクセス権限を変更する
- オブジェクトへのアクセス権限を削除する

主成分: このオブジェクトにアクセス権限があるユーザーまたはグループのリポジトリ・ユーザー名。

アクセス権。 このユーザーまたはグループがオブジェクトに対して持つアクセス権限。

追加: 1 つまたは複数のユーザーまたはグループをこのオブジェクトのアクセス権限を持つユーザーまたはグループのリストに追加できます。詳しくは、トピック『アクセス権リストへのユーザーの追加』を参照してください。

変更: このオブジェクトに対する選択したユーザーまたはグループのアクセス権限を変更することができます。デフォルトでは、読み取りアクセス権が付与されています。所有者、書き込み、読み取り、削除、変更という追加のアクセス権限を与えることができます。

削除: このオブジェクトに対するアクセス権リストから選択したユーザーまたはグループを削除します。

アクセス権リストへのユーザーの追加: 「リポジトリ: 保存」ダイアログ・ボックスの「セキュリティー」タブで、「追加」を選択すると、以下のフィールドを使用できます。

プロバイダーの選択: 認証用のセキュリティー・プロバイダーを選択します。異なるセキュリティー・プロバイダーを使用するよう、リポジトリを設定することができます。必要な場合、詳しい情報を各サイトの管理者に問い合わせてください。

検索: 追加するユーザーまたはグループのリポジトリ・ユーザー名を入力し、「検索」を検索するとユーザー・リストに名前が表示されます。一度に複数のユーザー名を追加するには、このフィールドを空白にし、「検索」をクリックすると、すべてのリポジトリ・ユーザー名のリストが表示されます。

ユーザー・リスト：リストから 1 つまたは複数のユーザー名を選択し、「OK」をクリックすると、そのユーザー名がアクセス権リストに追加されます。

オブジェクトに対するアクセス権限の変更：「リポジトリ：保存」ダイアログ・ボックスの「セキュリティー」タブで、「変更」を選択すると、以下のフィールドを使用できます。

所有者：このユーザーまたはグループに、オブジェクトに対する所有者のアクセス権限を与えます。所有者は、削除および変更のアクセス権限など、オブジェクトに対して完全なコントロール権を持ちます。

読み取り：デフォルトでは、オブジェクトの所有者ではないユーザーまたはグループには、オブジェクトに対する読み取りアクセス権限のみが与えられます。該当するチェック・ボックスをオンにすると、このユーザーまたはグループの書き込み、削除および権限の変更のアクセス権限が追加されます。

ストリームの保存

リポジトリにストリームを `.str` ファイルとして保存できます。リポジトリから他のユーザーはそれらにアクセスすることができます。

注：リポジトリ機能をさらに利用するための、ストリーム展開の詳細については、177 ページの『ストリームの展開』を参照してください。

現在のストリームを保存するには

1. メイン・メニューで次の各項目をクリックします。

「ファイル」 > 「保存」 > 「ストリームとして保存...」

2. 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者に問い合わせてください。
3. Repository の「保存」ダイアログ・ボックスで、オブジェクトを保存したいフォルダーを選択し、記録したいその他の情報を指定して「保存」ボタンをクリックします。詳しくは、トピック 166 ページの『オブジェクトの特性の設定』を参照してください。

プロジェクトの格納

リポジトリに完全な IBM SPSS Modeler プロジェクトを `.cpj` ファイルとして保存できます。他のユーザーもプロジェクトにアクセスすることができます。

プロジェクト・ファイルがその他の IBM SPSS Modeler オブジェクトのコンテナであるため、リポジトリのプロジェクトのオブジェクトを格納することを IBM SPSS Modeler に知らせる必要があります。この操作では、「プロジェクト・プロパティー」ダイアログ・ボックスでの設定を使用します。詳しくは、トピック 190 ページの『プロジェクトのプロパティーの設定』を参照してください。

プロジェクトをリポジトリに格納するよう設定すると、新規のオブジェクトをそのプロジェクトに追加するたびに、IBM SPSS Modeler はそのオブジェクトを格納するよう要求するメッセージを自動的に出力します。

IBM SPSS Modeler とのセッションを終了するときに、プロジェクト・ファイルの新しいバージョンを格納して、その追加を記憶する必要があります。プロジェクト・ファイルは、自動的にこのオブジェクトの最新バージョンを含めます（そして取得します）。IBM SPSS Modeler とのセッション中にオブジェクトを追加しなかった場合は、このプロジェクト・ファイルを復元してはいけません。ただし、変更したプロジェクトのオブジェクト（ストリーム、出力など）の新しいバージョンを格納する必要があります。

プロジェクトを格納する手順は、次のとおりです。

1. IBM SPSS Modeler のマネージャ パレットの「CRISP-DM」タブまたは「クラス」タブでプロジェクトを選択し、メニューから次の項目をクリックします。

「ファイル」 > 「プロジェクト」 > 「プロジェクトを格納...」

2. 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者に問い合わせてください。
3. Repository の「保存」ダイアログ・ボックスで、オブジェクトを保存したいフォルダーを選択し、記録したいその他の情報を指定して「保存」ボタンをクリックします。詳しくは、トピック 166 ページの『オブジェクトの特性の設定』を参照してください。

ノードの格納

リポジトリに現在のストリームの各ノード定義を .nod ファイルとして保存できます。リポジトリから他のユーザーはそのファイルにアクセスすることができます。

ノードを格納する手順は、次のとおりです。

1. ストリーム キャンバスにあるノードを右クリックして、「ノードを格納」 をクリックします。
2. 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者に問い合わせてください。
3. Repository の「保存」ダイアログ・ボックスで、オブジェクトを保存したいフォルダーを選択し、記録したいその他の情報を指定して「保存」ボタンをクリックします。詳しくは、トピック 166 ページの『オブジェクトの特性の設定』を参照してください。

出力オブジェクトの格納

リポジトリに現在のストリームの出力オブジェクトを .nod ファイルとして保存できます。リポジトリから他のユーザーはそのファイルにアクセスすることができます。

出力オブジェクトを格納する手順は、次のとおりです。

1. SPSS Modeler のマネージャー・パネルの「出力」タブでオブジェクトを選択し、メイン・メニューで次の項目をクリックします。

「ファイル」 > 「出力」 > 「出力を格納...」

2. または、「出力」タブにあるオブジェクトを右クリックして、「格納」 をクリックします。
3. 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者に問い合わせてください。
4. Repository の「保存」ダイアログ・ボックスで、オブジェクトを保存したいフォルダーを選択し、記録したいその他の情報を指定して「保存」ボタンをクリックします。詳しくは、トピック 166 ページの『オブジェクトの特性の設定』を参照してください。

モデルおよびモデル・パレットの格納

リポジトリに各モデルを .gm ファイルとして保存できます。リポジトリから他のユーザーはそれらにアクセスすることができます。また、モデル・パレットのすべての内容を .gen ファイルとしてリポジトリに保存することもできます。

モデルの格納。

1. SPSS Modeler の「モデル」パレットのオブジェクトをクリックし、メイン・メニューで次の項目をクリックします。

「ファイル」 > 「モデル」 > 「モデルを格納...」

2. または、モデル・パレットにあるオブジェクトを右クリックして、「モデルを格納」をクリックします。
3. 「格納手順の完了」から先に進みます。

モデル・パレットの格納。

1. モデル・パレットの背景を右クリックします。
2. ポップアップ・メニューで「パレットを格納」をクリックします。
3. 「格納手順の完了」から先に進みます。

格納手順の完了。

1. 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者に問い合わせてください。
2. Repository の「保存」ダイアログ・ボックスで、オブジェクトを保存したいフォルダーを選択し、記録したいその他の情報を指定して「保存」ボタンをクリックします。詳しくは、トピック 166 ページの『オブジェクトの特性の設定』を参照してください。

リポジトリからのオブジェクトの取得

リポジトリに格納されたストリーム、モデル、モデル・パレット、ノード、プロジェクト、出力オブジェクトを取得できます。

注：ここで説明したようにメニュー・オプションを使用するほか、SPSS Modeler ウィンドウの右上にあるマネージャー・パネルの該当するタブを右クリックして、ストリーム、出力オブジェクト、モデルおよびモデル・パレットを取得することができます。

1. ストリームを取得するには、IBM SPSS Modeler のメインメニューの次の項目をクリックしてください。

「ファイル」 > 「ストリームを取得...」

2. モデル、モデル・パレット、プロジェクト、または出力オブジェクトを取得するには、IBM SPSS Modeler のメイン・メニューで次の項目をクリックします。

「ファイル」 > 「モデル」 > 「モデルを取得...」

または

「ファイル」 > 「モデル」 > 「モデル・パレットを取得...」

または

「ファイル」 > 「プロジェクト」 > 「プロジェクトを取得...」

または

「ファイル」 > 「出力」 > 「出力を取得...」

3. または、マネージャ またはプロジェクト・パネル内で右クリックし、ポップアップ・メニューから「取得」をクリックします。
4. ノードを取得するには、IBM SPSS Modeler のメインメニューの次の項目をクリックしてください。

「挿入」 > 「リポジトリからノード (またはスーパーノード)...」

- a. 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者に問い合わせてください。
5. 「リポジトリ: 取得」ダイアログで、オブジェクトを参照し、オブジェクトを選択して「取得」ボタンをクリックします。詳しくは、トピックを参照してください。

取得するオブジェクトの選択

「リポジトリ: 読み込み/検索」ダイアログ・ボックスでは、以下のフィールドを使用できます。

検索対象: 現在のフォルダーのフォルダー階層が表示されます。異なるフォルダーに移動するには、このリストから選択して直接移動するか、このフィールドの下にあるオブジェクト・リストを使用して移動します。

「上位フォルダー」ボタン: 階層内の現在のフォルダーの 1 つ上のレベルに移動します。

「新規フォルダー」ボタン: 階層内の現在のフォルダーに新しいフォルダーを作成します。

ファイル名: 選択したオブジェクトのリポジトリ・ファイル名。そのオブジェクトを取得するには、「取得」をクリックします。

ファイルの種類: 取得するオブジェクトの種類。フォルダーと共に、この種類のオブジェクトのみが、オブジェクト・リストに表示されます。取得する異なる種類のオブジェクトを表示するには、リストからオブジェクトの種類を選択します。

「ロック状態で開く」。デフォルトでは、オブジェクトを取得すると、リポジトリ内でロックされ、他のユーザーは更新できません。取得時にオブジェクトをロックしたくない場合、このボックスをオフにします。

説明、キーワード: オブジェクトを保存したとき、オブジェクトに関する詳細が定義されている場合、これらの詳細情報がここに表示されます。詳しくは、トピック 166 ページの『保存するオブジェクトに関する情報の追加』を参照してください。

「バージョン」。最新版以外のオブジェクトのバージョンを取得するには、このボタンをクリックします。すべてのバージョンについての詳細な情報が表示され、必要なバージョンを選択できます。

オブジェクトのバージョンの選択

「リポジトリ: バージョンの選択」ダイアログ・ボックスで特定のバージョンのリポジトリ・オブジェクトを選択する手順は次のとおりです。

1. (オプション) 該当する列の見出しをクリックして、バージョン、ラベル、サイズ、作成日、作成ユーザーによってリストをソートします。
2. 使用するオブジェクトのバージョンを選択します。
3. 「続行」をクリックします。

リポジトリ内のオブジェクトの検索

オブジェクトを、名前、フォルダー、タイプ、ラベル、日付、またはその他の基準で検索できます。

名前によるオブジェクトの検索

1. IBM SPSS Modeler のメイン・メニューで次の各項目をクリックします。

「ツール」 > 「リポジトリ」 > 「探索...」

- 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者にお問い合わせください。

2. 「検索」タブをクリックします。

3. 「検索するオブジェクトの名前」 フィールドで、検索したいオブジェクトの名前を指定します。

名前でもオブジェクトを検索する場合、ワイルドカード文字としてアスタリスク (*) を使用すると、任意の文字列、疑問符 (?) を使用すると、1 文字に相当します。例えば、*cluster* は、cluster という文字列を名前の一部に持っているすべてのオブジェクトに一致します。文字列 m0?_* を検索すると、M01_cluster.str と M02_cluster.str は一致しますが、M01a_cluster.str には一致しません。検索する場合は大文字と小文字の区別はされません (cluster は Cluster と CLUSTER にも一致します)。

注: オブジェクトが多数ある場合は、検索に時間がかかる場合があります。

他の基準による検索

タイトル、ラベル、日付、作成者、キーワード、インデックス付きコンテンツ、または説明に基づいて検索を実行できます。すべての指定された検索基準に一致するオブジェクトだけが検索されます。例えば、特定のラベルが付けられ、特定の日付以降に変更された、1 つ以上のクラスタリング・モデルを含むすべてのストリームを検索できます。

「オブジェクト・タイプ」。モデル、ストリーム、出力、ノード、スーパーノード、プロジェクト、モデル・パレット、またはその他の種類のオブジェクトに検索を制限することができます。

- モデル: カテゴリー (分類、近似、クラスタリング、その他)、または Kohonen などの特定のモデル作成アルゴリズムによりモデルを検索できます。

使用されているフィールド名で検索することもできます。例えば、入力フィールドまたは出力 (対象) フィールドとして、income という名前のフィールドを使用しているすべてのモデルを検索できます。

- ストリーム: ストリームについて、使用されるフィールド名またはストリームに含まれるモデル タイプ (カテゴリーまたはアルゴリズム) で検索を制限できます。

トピック: 適切な権限を持つリポジトリ・ユーザーによって設定されたリストから、特定のトピックに関連するモデルを検索することができます (詳細は「Deployment Manager ユーザーズ・ガイド」を参照)。リストを取得するには、このボックスをオンにし、表示される「トピックを追加」ボタンをクリックして、リストから 1 つまたは複数のトピックを選択し、「OK」をクリックします。

ラベル: 特定のオブジェクト・バージョンのラベルに検索を制限します。

「日付」。作成日または修正日を指定して、以前、以降、または指定された日付範囲の間にあるオブジェクトを検索できます。

「作成者」。特定のユーザーが作成したオブジェクトに検索を制限します。

キーワード: 指定されたキーワードの検索。IBM SPSS Modeler では、キーワードは、ストリーム、モデル、または出力オブジェクトの「注釈」タブで指定します。

説明: 説明フィールドにある特定の用語を検索します。IBM SPSS Modeler では、説明は、ストリーム、モデル、または出力オブジェクトの「注釈」タブで指定します。検索するフレーズが複数なら、セミコロンで区切ります (例: income; crop type; claim value)。 (検索するフレーズでは空白が重要です。例えば、スペースが 1 個の crop type とスペースが 2 個の crop type は同一ではありません。)

リポジトリのオブジェクトを変更

SPSS Modeler から直接、リポジトリの既存のオブジェクトを変更できます。以下を行うことができます。

- フォルダーの作成、名前の変更または削除
- オブジェクトのロックまたはロック解除
- オブジェクトの削除

フォルダーの編集、名前の変更、削除

1. リポジトリでフォルダーの操作を実行するには、SPSS Modeler のメニューで次の項目をクリックします。

「ツール」 > 「リポジトリ」 > 「探索...」

- a. 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者に問い合わせてください。
2. 「フォルダー」 タブがアクティブであることを確認してください。
 3. 新しいフォルダーを作成するには、親フォルダーを右クリックして、「新規フォルダー」 をクリックします。
 4. フォルダー名を変更するには、そのフォルダーを右クリックして 「フォルダー名の変更」 をクリックします。
 5. フォルダーを削除するには、そのフォルダーを右クリックして 「フォルダーの削除」 をクリックします。

リポジトリ・オブジェクトのロックおよびロック解除

オブジェクトをロックして、ほかのユーザーが既存のバージョンを更新したり新しいバージョンを作成しないようにすることができます。ロックされたオブジェクトは、オブジェクト・アイコン上に南京錠のシンボルで示されます。



図 15. ロックされたオブジェクト

オブジェクトをロックするには

1. リポジトリ・ウィンドウで目的のオブジェクトを右クリックします。
2. 「ロック」をクリックします。

オブジェクトをロックを解除するには

1. リポジトリ・ウィンドウで目的のオブジェクトを右クリックします。
2. 「ロック解除」をクリックします。

リポジトリのオブジェクトを削除

オブジェクトをリポジトリから削除する前に、全部のバージョンを削除するか、またはあるオブジェクトの特定のバージョンを削除するのかを決める必要があります。

オブジェクトの全部のバージョンを削除するには

1. リポジトリ・ウィンドウで目的のオブジェクトを右クリックします。
2. 「オブジェクトの削除」をクリックします。

オブジェクトの最新バージョンを削除するには

1. リポジトリ・ウィンドウで目的のオブジェクトを右クリックします。
2. 「削除」をクリックします。

オブジェクトの前のバージョンを削除するには

1. リポジトリ・ウィンドウで目的のオブジェクトを右クリックします。
2. 「バージョンの削除」をクリックします。
3. 削除するバージョンを選択して「OK」をクリックします。

リポジトリ・オブジェクトのプロパティの管理

SPSS Modeler からさまざまなオブジェクトのプロパティを制御できます。以下を行うことができます。

- フォルダーのプロパティを表示する
- オブジェクトのプロパティを表示および編集する
- オブジェクトのバージョン・ラベルを作成、適用、および削除する

フォルダーのプロパティの表示

リポジトリ・ウィンドウにあるフォルダーのプロパティを表示するには、目的のフォルダーを右クリックします。「フォルダーのプロパティ」をクリックします。

「全般」タブ

このタブには、フォルダー名、作成日時、および修正日時が表示されます。

「権限」タブ

このタブでは、フォルダーの読み取り権限と書き込み権限を指定します。親フォルダーへのアクセスを持つすべてのユーザーとグループがリストされます。権限は階層によって異なります。例えば、読み取り権限を持っていない場合、書き込み権限を持つことができません。書き込み権限を持っていない場合、削除権限を持つことはできません。

ユーザーとグループ：このフォルダーに対して少なくとも読み取りアクセス権を持っているリポジトリ・ユーザーおよびグループが表示されます。「書き込み」チェック・ボックスおよび「削除」チェック・ボックスをオンにすると、このフォルダーに対するこれらのアクセス権が、特定のユーザーまたはグループ

に割り当てられます。別のユーザーまたはグループにアクセスを割り当てるには、「権限」タブの右側にある「ユーザーとグループを追加します」アイコンをクリックします。利用可能なユーザーとグループのリストは、管理者により管理されています。

権限のカスケード：子フォルダーがある場合、現在のフォルダーに対する変更が、子フォルダーにどのように適用されるかを制御できます。

- すべての権限をカスケードする：現在のフォルダーからすべての子、および子孫のフォルダーから権限設定をカスケードします。この方法は、複数のフォルダーの権限を一度に素速く設定できます。親フォルダーで必要な権限を設定し、それを必要に応じてカスケードします。
- 変更のみカスケードする：前回の変更が適用されてから行われた変更のみをカスケードします。例えば、新しいグループが追加されたので、そのグループに「販売」ツリーにある全てのフォルダーへのアクセスを与えたい場合、グループ・アクセスをルートの「販売」フォルダーに与えて、その変更をすべてのサブフォルダーにカスケードできます。既存のサブフォルダーに対するすべてのそれ以外の権限は、以前のまま変更されません。
- カスケードしない：変更は、現在のフォルダーにのみ適用され、子フォルダーにカスケードされません。

オブジェクトのプロパティの表示および編集

「オブジェクト・プロパティ」ダイアログ・ボックスで、プロパティを表示し、編集することができます。プロパティには変更できないものもありますが、いつでも新しいバージョンを追加することで更新できます。

1. リポジトリ・ウィンドウで目的のオブジェクトを右クリックします。
2. 「オブジェクトのプロパティ」をクリックします。

「全般」タブ

「名前」。リポジトリで表示したときのオブジェクトの名前です。

作成日：オブジェクト (バージョンでない) が作成された日付です。

最終更新日：最新のバージョンが更新された日時です。

「作成者」。ユーザーのログイン名です。

説明: デフォルトでは、SPSS Modeler でオブジェクトの「注釈」タブで指定された説明が含まれています。

リンクされたトピック：リポジトリを使用すると、モデルおよび関連付けられたオブジェクトを、必要に応じてトピックごとに整理できます。使用できるトピックのリストは、適切な権限を持つリポジトリ・ユーザーによって設定されます (詳細は、「*Deployment Manager* ユーザーズ・ガイド」を参照)。

キーワード: ストリーム、モデル、または出力オブジェクトの「注釈」タブでキーワードを指定します。複数のキーワードは、スペースで区切る必要があります。また、最大で 255 文字 (半角) までです。(キーワードにスペースが含まれる場合は、キーワード間の区切りには疑問符を使用してください。)

「バージョン」タブ

リポジトリに格納されたオブジェクトには、複数のバージョンがあります。「バージョン」タブは、各バージョンについての情報を表示します。

格納されているオブジェクトの指定されたバージョンで、次のプロパティを指定または変更できます。

「バージョン」。バージョンの一意の識別子で、そのバージョンが格納された時間に基づいて生成されます。

ラベル: バージョンの現在のラベルです。バージョン識別子と異なり、ラベルはオブジェクトのあるバージョンから別のバージョンに移動できます。

各バージョンで、ファイル・サイズ、作成日時、および作成者も表示されます。

ラベルの編集: 「バージョン」 タブの右上にある「ラベルの編集」 アイコンを使用して、格納されたオブジェクトのラベルを定義、適用または削除します。詳しくは、トピック『オブジェクトのバージョン・ラベルの管理』を参照してください。

「権限」 タブ

「権限」 タブで、オブジェクトに対する読み取りと書き込みの権限を設定できます。現在のオブジェクトへのアクセスを持つすべてのユーザーとグループがリストされます。権限は階層によって異なります。例えば、読み取り権限を持っていない場合、書き込み権限を持つことができません。書き込み権限を持っていない場合、削除権限を持つことはできません。

ユーザーとグループ: このオブジェクトに対して少なくとも読み取りアクセス権を持っているリポジトリ・ユーザーおよびグループが表示されます。「書き込み」チェック・ボックスおよび「削除」チェック・ボックスをオンにすると、このオブジェクトに対するこれらのアクセス権が、特定のユーザーまたはグループに割り当てられます。別のユーザーまたはグループにアクセスを割り当てるには、「権限」タブの右側にある「ユーザーとグループを追加します」 アイコンをクリックします。利用可能なユーザーとグループのリストは、管理者により管理されています。

オブジェクトのバージョン・ラベルの管理

「バージョン・ラベルの編集」 ダイアログ・ボックスで以下のことができます。

- 選択したオブジェクトへのラベルの適用
- 選択したオブジェクトからのラベルの削除
- 新しいラベルの定義およびオブジェクトへの適用

ラベルをオブジェクトに適用するには

1. 「利用可能なラベル」 リストで 1 つ以上のラベルを選択します。
2. 右矢印のボタンをクリックして、選択したラベルを 「適用されるラベル」 リストに移動します。
3. 「OK」 をクリックします。

オブジェクトからラベルを削除するには

1. 「適用されるラベル」 リストで 1 つ以上のラベルを選択します。
2. 左矢印のボタンをクリックして、選択したラベルを 「利用可能なラベル」 リストに移動します。
3. 「OK」 をクリックします。

新しいラベルを定義しオブジェクトに適用するには

1. 「新規ラベル」 フィールドでラベル名を入力します。
2. 右矢印のボタンをクリックして、新しいラベルを 「適用されるラベル」 リストに移動します。
3. 「OK」 をクリックします。

ストリームの展開

ストリームをシニアアプリケーションの IBM SPSS Modeler Advantage で使用できるようにするには、リポジトリ内でストリーム (.str ファイル) として展開する必要があります。

注: スコアリングブランチ内に複数の入力ノードがあるストリームはデプロイできません。

ストリームは、IBM SPSS Collaboration and Deployment Services のエンタープライズ レベルの機能を十分に活用できます。詳しくは、トピック 164 ページの『リポジトリ・オブジェクトの保存と展開』を参照してください。

現在のストリームを展開するには (「ファイル」メニューを使用した方法)

1. メイン・メニューで次の各項目をクリックします。

「ファイル」 > 「保存」 > 「展開」

2. 展開の種類を選択し、必要に応じてダイアログ・ボックスの残りを入力します。
3. 「ストリームとして展開」 をクリックすると、IBM SPSS Modeler Advantage または IBM SPSS Collaboration and Deployment Services と使用するようストリームを展開します。
4. 「保存」 をクリックします。詳細は、「ヘルプ」 をクリックしてください。
5. 「展開プロセスの完了」 から先に進みます。

現在のストリームを展開するには (「ツール」メニューを使用した方法)

1. メイン・メニューで次の各項目をクリックします。

「ツール」 > 「ストリームのプロパティ」 > 「展開」

2. 展開の種類を選択し、必要に応じて「展開」タブの残りを入力し、「保存」 をクリックします。詳しくは、トピック『ストリーム展開のオプション』を参照してください。

展開プロセスの完了

1. 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者に問い合わせてください。
2. Repository の「保存」ダイアログ・ボックスで、オブジェクトを保存したいフォルダーを選択し、記録したいその他の情報を指定して「格納」を格納します。詳しくは、トピック 166 ページの『オブジェクトの特性の設定』を参照してください。

ストリーム展開のオプション

「ストリーム・オプション」ダイアログ・ボックスの「展開」タブを使用すると、ストリーム展開のオプションを指定できます。

ストリームを展開すると、シンククライアント・アプリケーション IBM SPSS Modeler Advantage でストリームを開いたり変更したりできます。ストリームは、拡張子 .str のファイアとしてリポジトリに保存されます。

ストリームの展開により、マルチユーザー・アクセス、自動化スコアリング、モデル更新、Champion Challenger 分析など、IBM SPSS Collaboration and Deployment Services で使用できる追加機能を利用できます。

注: アソシエーション・ルール、STP、TCM の各モデル作成ノードは、IBM SPSS Collaboration and Deployment Services のモデル評価または Champion Challenger のステップをサポートしません。

「展開」タブで、IBM SPSS Modeler がストリームに作成するストリームの説明をプレビューすることもできます。詳しくは、トピック 55 ページの『ストリームの説明』を参照してください。

展開タイプ：ストリームの展開方法を選択します。すべてのストリームは、展開のために指定されたスコアリング・ノードが必要です。追加の要件およびオプションは展開の種類によって異なります。

注: アソシエーション・ルール、STP、TCM の各モデル作成ノードは、IBM SPSS Collaboration and Deployment Services のモデル評価または Champion Challenger のステップをサポートしません。

- **<none>**。ストリームはリポジトリに展開されません。「ストリームの説明をプレビュー」を除くすべてのオプションが無効になります。
- スコアリングのみ：「保存」 ボタンをクリックすると、ストリームがリポジトリに展開されます。「スコアリング ノード」 フィールドで指定したノードを使用して、データをスコアリングできます。
- モデル リフレッシュ：「スコアリングのみ」と同じですが、「モデル作成ノード」 フィールドおよび「モデル ナゲット」 フィールドで指定したオブジェクトを使用して、モデルをリポジトリで更新できます。IBM SPSS Collaboration and Deployment Services では、自動モデル リフレッシュはデフォルトでサポートされないため、リポジトリからストリームを実行するときこの機能を使用する場合には、この展開タイプを選択する必要があることに注意してください。詳しくは、トピック 180 ページの『モデル・リフレッシュ』を参照してください。

スコアリング ノード：グラフ、出力、エクスポート、またはエクスポート・ノードを選択して、データのスコアリングに使用するストリームブランチを識別します。ストリームは実際いくつもの有効なブランチ、モデル、ターミナル・ノードを含んでいますが、唯一のスコアリングブランチが展開のために定義されている必要があります。ストリームを展開する最も基本的な要件です。

スコアリング パラメーター：パラメーターを指定することができ、スコアリングブランチが実行される際にそのパラメーターを変更することができます。詳しくは、トピック 179 ページの『スコアリング パラメーターおよびモデル作成パラメーター』を参照してください。

モデリング ノード：モデル・リフレッシュについて、リポジトリのモデルを再生成または更新するために使用されるモデル作成ノードを指定します。「モデル ナゲット」 に指定されたモデル ナゲットと同じ種類のモデル作成ノードである必要があります。

モデル構築のパラメーター：パラメーターを指定することができ、モデル作成ノードが実行される際にそのパラメーターを変更することができます。詳しくは、トピック 179 ページの『スコアリング パラメーターおよびモデル作成パラメーター』を参照してください。

モデル ナゲット：モデル・リフレッシュについて、リポジトリでストリームが更新されるたびに更新または再生成されるモデル ナゲットを指定します (通常はスケジュール化されたジョブの一環として)。モデルはスコアリングブランチにある必要があります。複数のモデルがスコアリングブランチに存在する場合がありますが、定義されるのは 1 つだけです。ストリームが最初に作成される場合、新規データが利用可能になると更新または再生成されるプレースホルダー・モデルとなる場合があります。

ストリームとして展開: IBM SPSS Modeler Advantage または IBM SPSS Collaboration and Deployment Services でストリームを使用する場合に、このオプションをクリックします。

検査 :展開するために有効なストリームかどうかを確認します。展開する前にすべてのストリームが指定されたスコアリング・ノードが必要です。これらの条件が満たされない場合、エラーメッセージが表示されません。

保存 :有効な場合ストリームを展開します。有効でない場合、エラー・メッセージが表示されます。「修正」ボタンをクリックしてエラーを修正し、もう一度実行してください。

ストリーム記述のプレビュー: IBM SPSS Modeler がストリームについて作成するストリームの説明の内容を表示できます。詳しくは、トピック 55 ページの『ストリームの説明』を参照してください。

注: アソシエーション・ルール、STP、TCM の各モデル作成ノードは、IBM SPSS Collaboration and Deployment Services のモデル評価または Champion Challenger のステップをサポートしません。

スコアリング パラメーターおよびモデル作成パラメーター

ストリームを IBM SPSS Collaboration and Deployment Services に展開する場合、モデルが更新またはスコアリングされるごとに、どのパラメーターを表示または編集するかを選択できます。例えば、最大値および最小値、またはジョブが実行されるごとに変更する必要があるその他の値を指定することができます。

1. ストリームの展開後に表示または編集できるようパラメーターを表示可能にするには、「スコアリングパラメーター」ダイアログ・ボックスのリストから、そのパラメーターを選択します。

使用可能なパラメーターのリストは、「ストリームのプロパティ」ダイアログ・ボックスの「パラメーター」タブで定義されます。詳しくは、トピック 51 ページの『ストリームとセッション・パラメーターの設定』を参照してください。

スコアリングブランチ

ストリームを展開している場合、ストリームのいずれかのブランチをスコアリングブランチ (スコアリング・ノードを含むブランチ) として指定する必要があります。ブランチをスコアリングブランチとして指定すると、そのブランチはストリーム領域で強調表示され、スコアリングブランチのナゲットへのモデル・リンクとなります。この視覚的表示は、スコアリングブランチがすぐに明らかになっていない、複数のブランチによる複雑なストリームで特に役立ちます。

注 : スコアリングブランチとして指定できるストリームブランチは 1 つだけです。

ストリームでスコアリングブランチが既に定義されている場合、新しくブランチを指定すると、それがスコアリングブランチとして置き換えられます。「カスタム カラー」オプションを使用して、スコアリングブランチを示す色を設定できます。詳しくは、トピック 198 ページの『表示オプションの設定』を参照してください。

「ストリーム・マークアップを表示/非表示」ツールバー・ボタンを使用して、スコアリングブランチの表示を表示または非表示にできます。



図 16. 「ストリーム・マークアップを表示/非表示」ツールバー・ボタン

展開するスコアリングブランチの特定

ターミナル・ノードのポップアップ・メニューまたは「ツール」メニューを使用して、スコアリングブランチを指定できます。ポップアップ・メニューを使用すると、スコアリング・ノードがストリーム・プロパティの「展開」タブで自動的に設定されます。

ブランチをスコアリングブランチとして指定するには (ポップアップ・メニュー)

1. モデル ナゲットをターミナル・ノード (ナゲットから下流の処理ノードまたは出力ノード) に接続します。
2. ターミナル・ノードを右クリックします。
3. メニューで、「スコアリング枝として使用」を選択します。

ブランチをスコアリングブランチとして指定するには (「ツール」メニュー)

1. モデル ナゲットをターミナル・ノード (ナゲットから下流の処理ノードまたは出力ノード) に接続します。
2. メイン・メニューで次の各項目をクリックします。

「ツール」 > 「ストリームのプロパティ」 > 「展開」

3. 「展開タイプ」リストで、必要に応じて「スコアリングのみ」または「モデル リフレッシュ」をクリックします。詳しくは、トピック 177 ページの『ストリーム展開のオプション』を参照してください。
4. 「スコアリング ノード」フィールドをクリックして、リストからターミナル・ノードを選択します。
5. 「OK」をクリックします。

モデル・リフレッシュ

モデル・リフレッシュは、新しいデータを使用するストリームの既存のモデルを再構築するプロセスです。リポジトリのストリーム自体は変わりません。例えば、アルゴリズムやストリーム固有の設定は同じままですが、モデルは新しいデータに保持され、新しいバージョンのモデルが古いモデルより優れている場合、更新されます。

リフレッシュに設定できるストリームのモデル ナゲットは 1 つだけです。このモデルは、リフレッシュモデルと呼ばれます。ストリームのプロパティの「展開」タブで「モデル リフレッシュ」オプションをクリックすると (177 ページの『ストリーム展開のオプション』参照)、そのとき指定したモデル ナゲットがリフレッシュ モデルとなります。モデル ナゲットのポップアップ・メニューからモデルをリフレッシュ モデルに指定することもできます。指定するには、ナゲットがあらかじめスコアリングブランチにある必要があります。

ナゲットの「リフレッシュ モデル」状態をオフにすると、ストリームの展開タイプを「スコアリングのみ」に設定することとなり、ストリームのプロパティのダイアログの「展開」タブがそれに従って更新されます。現在のスコアリングブランチにあるナゲットのポップアップ・メニューで「リフレッシュ モデルとして使用」オプションを使用して、この状態をオンにしたりオフにしたりすることができます。

スコアリングブランチのナゲットのモデル・リンクを削除すると、ナゲットの「リフレッシュ モデル」状態も削除されます。「編集」メニューまたはツールバーを使用して、モデル・リンクの削除を取り消すことができます。取り消した場合、ナゲットの「リフレッシュ モデル」の状態も元に戻ります。

リフレッシュ モデルの選択方法

スコアリングブランチのほか、リフレッシュ モデルへのリンクもストリーム内で強調表示されます。リフレッシュ モデルに選択されるモデル ナゲット、つまり強調表示されるリンクは、ストリーム内のナゲットの数によって異なります。

ストリームで 1 つのモデル

リンクしたモデルがスコアリングに 1 つあり、そのように特定されている場合、そのナゲットがストリームのリフレッシュ モデルとなります。

ストリームに複数のモデル

ストリームにリンクしたナゲットが複数ある場合、次のようにしてリフレッシュ モデルを選択します。

「ストリームのプロパティ」ダイアログ・ボックスの「展開」タブでモデル ナゲットが定義されており、それがストリーム内にある場合、そのナゲットがリフレッシュ モデルになります。

「展開」タブでナゲットが定義されていない場合、またはナゲットが定義されているがスコアリングブランチにない場合、ターミナル・ノードに最も近いナゲットがリフレッシュ モデルとなります。

リフレッシュ リンクであるすべてのモデル・リンクの選択を解除すると、スコアリングブランチのみが強調表示され、リンクは強調表示されなくなります。展開タイプは「スコアリングのみ」に設定されます。

注：一方のリンクを置換え状態に設定できますが、もう一方は置き換えることができません。この場合、リフレッシュ モデルに選択されているモデル ナゲットは、スコアリングブランチが指定されているときにリフレッシュ リンクがあり、ターミナル・ノードに最も近いモデル ナゲットです。

ストリームにモデルがない

ストリームにモデルがない場合、またはモデル・リンクのないモデルだけがある場合、展開タイプは「スコアリングのみ」に設定されます。

スコアリングブランチでのエラーのチェック

スコアリングブランチを指定すると、エラーがないかチェックが行われます。

エラーが見つかり、スコアリングブランチがスコアリングブランチエラーの色で強調表示され、エラー・メッセージが表示されます。「カスタム カラー」オプションを使用して、エラーの色を設定できます。詳しくは、トピック 198 ページの『表示オプションの設定』を参照してください。

エラーが見つかった場合、次の手順を実行してください。

1. エラー・メッセージの内容に従って、エラーを修正します。
2. メイン・メニューで次の各項目をクリックします。

「ツール」 > 「ストリームのプロパティ」 > 「展開」

「チェック」をクリックします。

3. 必要に応じて、エラーがなくなるまでこのプロセスを繰り返します。

第 11 章 外部アプリケーションへのエクスポート

外部アプリケーションへのエクスポートについて

IBM SPSS Modeler では、データマイニング・プロセス全体を、外部のアプリケーションにエクスポートするためのいくつかの機構を備えており、データの準備作業を行ってから、IBM SPSS Modeler の外部で、ユーザー独自のアプリケーションを使用してモデルを構築できます。

前の項では、マルチユーザー・アクセス、ジョブ・スケジュール、その他の機能を利用するための IBM SPSS Collaboration and Deployment Services リポジトリへのストリームの展開方法について説明しました。同じように、IBM SPSS Modeler ストリームも次のアプリケーションと組み合わせて使用できます。

- IBM SPSS Modeler Advantage
- PMML 形式のファイルをインポートおよびエクスポートできるアプリケーション

IBM SPSS Modeler Advantage でのストリーム使用の詳細は、『IBM SPSS Modeler Advantage でストリームを開く』を参照してください。

PMML 形式をサポートするアプリケーションとモデルを共有するための、PMML ファイルとしてのモデルのエクスポートおよびインポートの詳細は、184 ページの『PMML としてのモデルのインポートおよびエクスポート』を参照してください。

IBM SPSS Modeler Advantage でストリームを開く

IBM SPSS Modeler ストリームは、シンククライアント・アプリケーション IBM SPSS Modeler Advantage と組み合わせて使用できます。IBM SPSS Modeler Advantage 全体でカスタマイズされたアプリケーションを作成できますが、アプリケーション ワークフローの基本として、IBM SPSS Modeler で作成されたストリームを使用することもできます。

IBM SPSS Modeler Advantage でストリームを開く手順は次のとおりです。

1. 「ストリームとして展開」 オプションをクリックして、IBM SPSS Collaboration and Deployment Services リポジトリでストリームを展開します。詳しくは、トピック 177 ページの『ストリームの展開』を参照してください。
2. 「IBM SPSS Modeler Advantage で開く」 ツールバー・ボタンをクリックするか、メイン・メニューで次の項目をクリックします。

「ファイル」 > 「IBM SPSS Modeler Advantage で開く」

1. 必要に応じて、リポジトリへの接続の設定を行います。詳しくは、トピック 164 ページの『リポジトリへの接続』を参照してください。特定のポート、パスワード、およびその他詳細については、各サイトのシステム管理者に問い合わせてください。

注：リポジトリ・サーバーにも、IBM SPSS Modeler Advantage ソフトウェアをインストールする必要があります。

1. Repository の「保存」ダイアログ・ボックスで、オブジェクトを保存したいフォルダーを選択し、記録したいその他の情報を指定して「保存」ボタンをクリックします。詳しくは、トピック 166 ページの『オブジェクトの特性の設定』を参照してください。

インストールすると、ストリームがすでに開いている IBM SPSS Modeler Advantage を起動します。IBM SPSS Modeler のストリームは閉じます。

PMML としてのモデルのインポートおよびエクスポート

PMML (Predictive Model Markup Language) は、モデルへの入力、データ・マイニングのデータの準備に使用する返還、モデル自体を定義するパラメーターなど、データ・マイニングおよび統計モデルを説明する XML 形式です。IBM SPSS Modeler は PMML をインポートおよびエクスポートし、IBM SPSS Statistics など、この形式をサポートする他のアプリケーションとモデルを共有できるようにします。

PMML の詳細は、データ・マイニング・グループの Web サイト (<http://www.dmg.org>) を参照してください。

モデルをエクスポートするには

PMML エクスポートでは、IBM SPSS Modeler 内で生成されたほとんどの種類のモデルがサポートされます。詳しくは、トピック 185 ページの『PMML をサポートするモデルの種類』を参照してください。

1. モデル・パレットのモデル ナゲットを右クリックします(または、キャンバス上のモデル ナゲットをダブルクリックして、「ファイル」メニューを選択します)。
2. メニューで、「**PMML** をエクスポート」 をクリックします。
3. 「エクスポート」(または「保存」) ダイアログ・ボックスで、対象ディレクトリーとモデルの一意の名前を指定します。

注：「ユーザー オプション」ダイアログ・ボックスで、PMML エクスポートのオプションを変更できます。メイン・メニューで次の各項目をクリックします。

「ツール」 > 「オプション」 > 「ユーザー オプション」

そこで「PMML」タブをクリックします。

詳しくは、トピック 200 ページの『PMML エクスポート・オプション の設定』を参照してください。

PMML として保存されたモデルをインポートするには

IBM SPSS Modeler または別のアプリケーションから PMML としてエクスポートされたモデルは、生成済みモデル・パレットへインポートできます。詳しくは、トピック 185 ページの『PMML をサポートするモデルの種類』を参照してください。

1. モデル・パレット内で、パレットを右クリックし、メニューから「**PMML** をインポート」を選択します。
2. インポートするファイルを選択し、必要に応じて、変数のラベルに関するオプションを指定します。
3. 「開く」 をクリックします。

モデル内に存在すれば変数ラベルを使用：PMML が、データ・ディクショナリー内の変数に対して、変数名と変数ラベル (*RefID* に対する *Referrer ID* など) の両方を指定している場合があります。元のエクスポートされた PMML に変数ラベルが存在するときに変数ラベルを使用するには、このオプションを選択します。

変数ラベル・オプションを選択したにもかかわらず、PMML 内に変数ラベルがない場合、変数名は通常のように使用されます。

PMML をサポートするモデルの種類

PMML のエクスポート

IBM SPSS Modeler のモデル: IBM SPSS Modeler で作成された次のモデルは、PMML 4.0 としてエクスポートできます。

- C&R Tree
- QUEST
- CHAID
- 線型回帰
- ニューラル・ネットワーク
- C5.0
- ロジスティック回帰
- 一般化線型
- SVM
- Apriori
- Carma
- K-Means
- Kohonen
- TwoStep
- GLMM (サポート対象は固定効果のみの GLMM モデルのみです)
- ディジション・リスト
- Cox
- シーケンス (シーケンス PMML モデルのスコアリングはサポートされていません)
- Statistics モデル

データベース固有のモデル: データベース固有のアルゴリズムを使用して作成されたモデルの場合、PMML エクスポートは、使用できません。Microsoft または Oracle Data Miner の Analysis Services を使用して作成されたモデルをエクスポートすることはできません。

PMML のインポート

IBM SPSS Modeler では、すべての IBM SPSS Statistics 製品の現在のバージョンで作成された PMML モデルをインポートおよびスコアリングできます。IBM SPSS Statistics 17.0 で生成されたモデルまたは変換 PMML と同様に、IBM SPSS Modeler からエクスポートされたモデルもインポートおよびスコアリングできます。基本的には、次の例外を除いて、スコアリング エンジンはずべての PMML をスコアリングできます。

- Apriori、CARMA、異常検出、シーケンス、およびアソシエーション ルールのモデルをインポートすることはできません。
- スコアリングに使用できる場合でも、IBM SPSS Modeler へのインポート後に PMML を参照することはできません。(これには、初めに IBM SPSS Modeler からエクスポートされたモデルは含まれません。この制限を回避するには、モデルを PMML ではなく、生成されたモデル・ファイル (*.gm) としてエクスポートしてください。)

- インポート時には制限つき検証が行われますが、モデルのスコアリング試行時には完全検証が実行されます。そのため、インポートは正常に行われますが、スコアリングは失敗したり不正な結果が生成されます。

注: サード・パーティーの PMML を IBM SPSS Modeler にインポートした場合、IBM SPSS Modeler は、認識できてスコアリング可能である有効な PMML のスコアリングを試行します。ただし、すべての PMML がスコアリングされることや、PMML を生成したアプリケーションと同じ方法で PMML がスコアリングされることは保証されません。

第 12 章 プロジェクトとレポート

プロジェクトの概要

プロジェクトは、データ・マイニング作業に関連したファイルのグループです。プロジェクトには、データ・ストリーム、グラフ、生成されたモデル、レポート、およびその他 IBM SPSS Modeler で作成したすべてが含まれます。IBM SPSS Modeler プロジェクトの第一印象は、単に出力を編成するための手段のように思われますが、それ以外にもさまざまな用途に利用することができます。プロジェクトを使用して、次のような処理ができます。

- プロジェクト・ファイル内の各オブジェクトに注釈を付ける。
- CRISP-DM 手法を使用してデータ・マイニング作業のガイド役をする。プロジェクトには、詳細な説明と CRISP-DM によるデータ・マイニング作業の実例を含んだ、CRISP-DM ヘルプ・システムも用意されています。
- データ・マイニングの目標をプレゼンテーションするための PowerPoint スライド・ショーや、使用するアルゴリズムに関する白書などの、IBM SPSS Modeler 以外のオブジェクトを追加する。
- 注釈に基づいて総合レポート、および簡単な更新レポートの両方を作成する。これらのレポートは、社内イントラネットなどに公開しやすい HTML 形式で生成することができます。

注：プロジェクト・パネルが IBM SPSS Modeler ウィンドウに表示されていない場合は、「表示」メニューで「プロジェクト」をクリックします。

プロジェクトに追加したオブジェクトを表示するには、クラス・ビューと CRISP-DM ビューの 2 つの方法を使用できます。プロジェクトに追加したものはすべて両方のビューに追加されます。2 つのビューを切り替えながら作業を行って、最適な編成を作成できます。

CRISP-DM ビュー

CRISP-DM (Cross-Industry Standard Process for Data Mining) をサポートすることで、IBM SPSS Modeler プロジェクトには、業界で実証された標準のデータ・マイニング作業の編成手段が用意されます。CRISP-DM では、作業の開始 (ビジネス要件の収集) から完了 (結果の展開) までのプロセスを大きく 6 つのフェーズに分けて記述します。通常、一部のフェーズは Clementine での作業とは関係ありませんが、IBM SPSS Modeler のプロジェクト・パネルでは、プロジェクトに関連するすべての資料やリソースを集中管理、保存、および追跡できるように、6 つのフェーズがすべて用意されています。例えば、一般的に業務の理解フェーズでは、要件の収集と各作業担当者との打ち合わせにより、データ・マイニングの最終目標を決定します。IBM SPSS Modeler でデータ処理などの作業を行うわけではありません。プロジェクト・パネルを利用すれば、このような打ち合わせ内容などの資料や記録を「業務の理解」フォルダーに保存して、将来の作業の参考にしたり、レポートに記載したりすることができます。

プロジェクト・パネルの CRISP-DM ビューには、データ・マイニング作業のライフ サイクルを説明する独自のヘルプ・システムも用意されています。IBM SPSS Modeler から、このヘルプには「ヘルプ」メニューの「CRISP-DM ヘルプ」をクリックしてアクセスします。

注：プロジェクト・パネルがウィンドウに表示されていない場合は、「表示」メニューで「プロジェクト」をクリックします。

デフォルトのプロジェクト・フェーズの設定

プロジェクトに追加されたオブジェクトは、CRISP-DM のデフォルト・フェーズに追加されます。このため、オブジェクトを使用するデータ・マイニング・フェーズに応じて、手作業でオブジェクトを編成し直す必要があります。現在作業中のフェーズを、デフォルト・フォルダーに設定することが賢明です。

デフォルト・フォルダーとして使用するフェーズを選択する手順は、次のとおりです。

1. CRISP-DM ビューで、デフォルトとして設定するフェーズに対応するフォルダーを右クリックします。
2. メニューから、「デフォルトとして設定」 をクリックします。

デフォルト・フォルダーは太字で表示されます。

クラス・ビュー

プロジェクト・パネルのクラス・ビューでは、IBM SPSS Modeler で行った作業が作成したオブジェクトの種類ごとに分類、編成されます。保存したオブジェクトは、次の任意のカテゴリーに追加することができます。

- ストリーム
- ノード
- モデル
- テーブル、グラフ、およびレポート
- その他 (スライド・ショーやデータ・マイニング作業の関連資料などの非 IBM SPSS Modeler ファイル)

オブジェクトをクラス・ビューに追加すると、そのオブジェクトは CRISP-DM ビューのデフォルト・フェーズ・フォルダーにも追加されます。

注：プロジェクト・パネルがウィンドウに表示されていない場合は、「表示」メニューで「プロジェクト」をクリックします。

プロジェクトの作成

基本的にプロジェクトは、プロジェクトに関連付けたすべてのファイルへの参照を含んだファイルです。つまり、プロジェクト内の各項目は個別に保存されるとともに、プロジェクト・ファイル (.cpj) 内に参照として保存されます。このような参照構造を採用しているため、次の点に注意してください。

- プロジェクトに追加する各項目は、プロジェクトに追加する前に、個別に保存しておく必要があります。追加する項目がまだ保存されていない場合は、プロジェクトに追加する前に保存するよう要求するメッセージが表示されます。
- ストリームのように個別に更新されるオブジェクトは、プロジェクト・ファイル内でも更新されます。
- ファイル・システムから手作業でオブジェクト (ストリーム、ノード、出力オブジェクトなど) を移動または削除すると、プロジェクト・ファイル内のリンクが無効になります。

新規プロジェクトの作成

新しいプロジェクトは、IBM SPSS Modeler ウィンドウから簡単に作成できます。まだプロジェクトを開いていない場合は、プロジェクトの作成を開始することができます。または、すでにプロジェクトを開いている場合、それを閉じて、新しいプロジェクトを作成することもできます。

メイン・メニューで次の各項目をクリックします。

「ファイル」 > 「プロジェクト」 > 「新規プロジェクト...」

プロジェクトへの追加

プロジェクトを作成するか、または既存のプロジェクトを開いた後に、データ・ストリーム、ノード、およびレポートなどのオブジェクトを追加できます。オブジェクトを追加するために、さまざまな方法を利用できます。

マネージャーからのオブジェクトの追加

IBM SPSS Modeler ウィンドウの右上にあるマネージャーを使用して、ストリームまたは出力を追加することができます。

1. 適切なマネージャ タブから、テーブルやストリームなどのオブジェクトを選択します。
2. 右クリックして、「プロジェクトに追加」を選択します。

オブジェクトがすでに保存されたことがある場合は、それが適切なオブジェクト・フォルダー (クラス・ビュー) またはデフォルトのフェーズ・フォルダー (CRISP-DM ビュー) に自動的に追加されます。

3. 代わりに、オブジェクトをマネージャーからプロジェクト・パネルにドラッグ・アンド・ドロップすることもできます。

注：先にオブジェクトの保存を要求するメッセージが表示されることもあります。オブジェクトを保存するには、「保存」ダイアログ・ボックスで「ファイルをプロジェクトに追加」を忘れずに選択してください。このオプションを選択すると、オブジェクトを保存した後、そのオブジェクトが自動的にプロジェクトに追加されます。

ストリーム領域からのノードの追加

「保存」ダイアログ・ボックスを使用して、ストリーム領域から個別のノードを追加できます。

1. キャンバスでノードを選択します。
2. ノードを右クリックして、「ノードの保存」をクリックします。または、メイン・メニューで次の各項目をクリックします。

「編集」 > 「ノード」 > 「ノードの保存...」

3. 「保存」ダイアログ・ボックスで、「ファイルをプロジェクトに追加」を選択します。
4. ノード名を指定して、「保存」をクリックします。

ファイルが保存され、プロジェクトにノードが追加されます。ノードは、クラス・ビューの「ノード」フォルダーと、CRISP-DM ビューのデフォルトのフェーズ・フォルダーに追加されます。

外部ファイルの追加

IBM SPSS Modeler 以外のさまざまなオブジェクトをプロジェクトに追加することができます。これは、IBM SPSS Modeler 内でデータ・マイニング・プロセス全体を管理する場合に役立ちます。例えば、データへのリンク、注釈、プレゼンテーション、およびグラフィックなどをプロジェクトに保存できます。CRISP-DM ビューでは、外部ファイルを保存するフォルダーを選択できます。クラス・ビューの場合、外部ファイルは「その他」フォルダーにしか保存できません。

プロジェクトに外部ファイルを追加する手順は、次のとおりです。

1. デスクトップからプロジェクトにファイルをドラッグ・アンド・ドロップします。

または

2. CRISP-DM ビューまたはクラス・ビュー内の対象フォルダーを右クリックします。
3. メニューで、「フォルダーに追加」 をクリックします。
4. ダイアログ・ボックス内でファイルを選択して、「開く」 をクリックします。

IBM SPSS Modeler プロジェクト内の選択したオブジェクトへの参照が追加されます。

プロジェクト IBM SPSS Collaboration and Deployment Services Repository への転送

すべてのコンポーネント・ファイルも含めたプロジェクト全体を、IBM SPSS Collaboration and Deployment Services Repository へ一度に転送できます。すでに目的の場所にあるオブジェクトは、移動しません。これは逆にも機能します。プロジェクト全体を IBM SPSS Collaboration and Deployment Services Repository からローカル・ファイル・システムに移行できます。

プロジェクトの転送

転送するプロジェクトがプロジェクト・パネル内で開かれていることを確認します。

プロジェクトを転送する手順は、次のとおりです。

1. ルート プロジェクト・フォルダーを右クリックして、「プロジェクトの転送」 をクリックします。
2. メッセージが表示されたら、IBM SPSS Collaboration and Deployment Services Repository へログインします。
3. プロジェクトの新しい場所を指定して、「OK」 をクリックします。

プロジェクトのプロパティの設定

プロジェクトのプロパティのダイアログ・ボックスを使用して、プロジェクトの内容やドキュメントをカスタマイズできます。プロジェクトのプロパティにアクセスする手順は、次のとおりです。

1. プロジェクト・パネル内のオブジェクトまたはフォルダーを右クリックして、「プロジェクトのプロパティ」 をクリックします。
2. プロジェクトの基本情報を指定するには、「プロジェクト」 タブをクリックします。

作成日：プロジェクトの作成日を表示します (編集不可)。

「要約」。プロジェクト・レポートに表示される、データ・マイニング・プロジェクトの要約を入力できません。

内容：このプロジェクト・ファイルに参照されるコンポーネントの種類と数を一覧表示します (編集不可)。

未保存のオブジェクトを名前を付けて保存：未保存のオブジェクトをローカル・ファイル・システムに保存するか、またはリポジトリに格納するかを指定します。詳しくは、トピック 163 ページの『IBM SPSS Collaboration and Deployment Services Repository について』を参照してください。

プロジェクトのロード中にオブジェクト参照を更新：プロジェクトのコンポーネントへの参照を更新する場合に、このオプションを選択します。注：プロジェクトに追加したファイルは、そのプロジェクト・フ

ファイル内に保存されるわけではありません。プロジェクトには、各ファイルへの参照が保存されます。つまり、元のファイルを移動または削除すると、その参照もプロジェクトから削除されます。

プロジェクトの注釈

プロジェクト・パネルには、データ・マイニング作業に関する注釈を記述するための、さまざまな方法が用意されています。プロジェクト・レベルの注釈は、しばしば「大局的な」目標や決定などの追跡に使用され、フォルダーやノード レベルの注釈は、詳細な説明を伝えるために使用されます。「注釈」タブには、取得できない欠損データの除外や、データ探索の過程で確立された有望な仮説などの、プロジェクト・レベルの詳細を記録するために十分な領域が用意されています。

プロジェクトに注釈を付ける手順は、次のとおりです。

1. CRISP-DM ビューまたはクラス・ビューで、プロジェクト・フォルダーを選択します。
2. フォルダーを右クリックして、「プロジェクトのプロパティ」 をクリックします。
3. 「注釈」タブをクリックします。
4. プロジェクトを記述するキーワードとテキストを入力します。

フォルダーのプロパティと注釈

個別のプロジェクト・フォルダー (CRISP-DM ビューおよびクラス・ビューの両方の中) に注釈を付けることができます。CRISP-DM ビューでは、注釈機能がデータ・マイニングの各フェーズに対する組織目標を記録するための、最も効果的な手段になります。例えば、「ビジネスの理解」 フォルダーで注釈ツールを使用して、「今回の調査におけるビジネス目標は、高い価値を持つ顧客の顧客離れを減らすことにあ

る。」のような情報を記入することができます。ここに記入したテキストは、「レポートに表示」 オプションを選択すると、プロジェクト・レポートに自動的に記載されます。

フォルダーに注釈を付ける手順は、次のとおりです。

1. プロジェクト・パネル内でフォルダーを選択します。
2. フォルダーを右クリックして、「フォルダーのプロパティ」 をクリックします。

CRISP-DM ビューのフォルダーには、各フェーズの目的の概略や、関連するデータ・マイニング作業を行うためのガイドなどの注釈が付けられます。これらの注釈は、必要に応じて削除したり編集したりすることができます。

「名前」。この領域には、選択したフィールドの名前が表示されます。

ツールヒント テキスト : マウス・ポインタをプロジェクト・フォルダー上に移動した時に表示される、カスタムのツールヒントを作成します。この機能は、CRISP-DM ビュー内で各フェーズの目標を記載したり、「作業中」や「完了」のようにフェーズのステータスを表示したりする場合に役立ちます。

「注釈」フィールド : このフィールドは、プロジェクト・レポートに記載できるより長い注釈を記載するために利用されます。CRISP-DM ビューの注釈には、データ・マイニングの各フェーズの注釈が記載されていますが、プロジェクトの内容に応じて、自由に変更することができます。

レポートに表示 : この注釈をレポートに記載するには、「レポートに表示」 を選択します。

オブジェクトのプロパティ

オブジェクトのプロパティを表示したり、プロジェクト・レポートに個別のオブジェクトを記載するかどうかを選択できます。オブジェクトのプロパティにアクセスする手順は、次のとおりです。

1. プロジェクト・パネルで、オブジェクトを右クリックします。

2. メニューで「オブジェクトのプロパティ」をクリックします。

「名前」。この領域には、保存したオブジェクトの名前が一覧表示されます。

パス：この領域には、オブジェクトの保存場所が一覧表示されます。

レポートに表示：生成するレポートにオブジェクトの詳細を記載する場合に選択します。

プロジェクトを閉じる

IBM SPSS Modeler を終了したり、新しいプロジェクトを開いたりすると、既存のプロジェクト・ファイル (.cpj) が閉じます。

製品と関連するいくつかのファイル (ストリーム、ノード、またはグラフなど) は開いたままとなる場合があります。これらのファイルを開いたままにする場合、「... これらのファイルを保存して閉じますか?」というメッセージに「いいえ」と答えてください。

プロジェクトを閉じた後に関連するファイルを変更して保存した場合、それらの変更内容は、次にプロジェクトを開いた時に反映されます。変更内容を反映しないようにするには、プロジェクトからそのファイルを削除するか、それを別のファイル名で保存してください。

レポートの生成

プロジェクトのもっとも役立つ機能の 1 つとして、プロジェクト内の各項目や注釈に基づいてレポートを生成できる機能があります。この機能は、「CRISP-DM 手法」で説明しているように、効果的なデータ・マイニングを行うために必要不可欠です。レポートは、いくつかのファイル タイプの 1 つに直接生成することも、ただちに確認できるように画面上のウィンドウに出力することもできます。この出力から、印刷や保存、または Web ブラウザー内でのレポートの表示を行うことができます。保存されたレポートは、組織内の他の部門に配布できます。

レポートは、データ・マイニング・プロセス中にプロジェクト・ファイルから何度も生成され、プロジェクトの関係者に配布されます。レポートには、プロジェクト・ファイルで参照されるオブジェクト情報や、作成した注釈などから、必要な情報や重要な情報を記載します。レポートは、クラス・ビューまたは CRISP-DM ビューのどちらかに基づいて作成できます。

レポートを生成する手順は次のとおりです。

1. CRISP-DM ビューまたはクラス・ビューで、プロジェクト・フォルダーを選択します。
2. フォルダーを右クリックして、「プロジェクト レポート」をクリックします。
3. レポート・オプションを指定して、「レポートの生成」 ボタンをクリックします。

「レポート」ダイアログ・ボックスのオプションには、目的に応じたさまざまなレポートを生成するいくつかの方法が用意されています。

出力名。レポートの出力を画面に送ることを選択した場合は、出力ウィンドウの名前を指定します。独自の名前を指定することも、IBM SPSS Modeler に自動的にウィンドウの命名をさせることもできます。

画面に出力：レポートを生成し、出力ウィンドウに表示する場合に選択します。出力ウィンドウからさまざまな種類のファイルへレポートをエクスポートするオプションもあります。

ファイルに出力：レポートを生成して、「ファイルの種類」リストで指定された種類のファイルとして保存する場合に選択します。

ファイル名 :生成されたレポートのファイル名を指定します。ファイルは、デフォルトで IBM SPSS Modeler ¥bin ディレクトリーへ保存されます。別の場所を指定するには、反復記号の「...」 ボタンを使用します。

ファイルの種類 :指定できるファイルの種類を次に示します。

- **HTML** ドキュメント : レポートは、単一の HTML ファイルとして保存されます。レポートにグラフが含まれている場合は PNG ファイルとして保存され、HTML ファイルに参照されます。レポートをインターネットで公開する場合は、HTML ファイルと、そのファイルが参照する画像の両方を必ずアップロードしてください。
- **テキスト・ドキュメント** : レポートは、単一のテキスト・ファイルとして保存されます。レポートにグラフが含まれている場合は、ファイル名とパスの参照のみがレポートに含まれます。
- **Microsoft Word** ドキュメント : レポートは、単一のドキュメントとして保存され、グラフはドキュメント内に直接埋め込まれます。
- **Microsoft Excel** ドキュメント : レポートは、単一のスプレッドシートとして保存され、グラフはスプレッドシート内に直接埋め込まれます。
- **Microsoft PowerPoint** ドキュメント : 各フェーズが新しいスライドに表示されます。どのようなグラフであっても、PowerPoint スライドに直接埋め込まれます。
- **出力オブジェクト**: IBM SPSS Modeler 内で開くと、このファイル (.cou) は 「レポートの形式」 の 「画面に出力」 と同じになります。

注 : Microsoft Office ファイルをエクスポートするには、対応するアプリケーションがインストールされている必要があります。

表題。 レポートのタイトルを指定します。

レポート構造: 「**CRISP-DM**」または「クラス」のいずれかを選択します。CRISP-DM ビューは、「大局的な」観点からのステータス・レポートと、データ・マイニングの各フェーズの詳細情報を提供します。クラス・ビューはオブジェクトを基準にしたビューで、データやストリームの内部的な追跡を行う場合に適しています。

「作成者」。デフォルトのユーザー名が表示されますが、変更できます。

レポートに表示する項目 : オブジェクトをレポートに含める方法を選択します。プロジェクト・ファイルに追加されたすべての項目を記載するには、「すべてのフォルダーとオブジェクト」を選択します。オブジェクトのプロパティーで「レポートに表示」が選択されているかどうかを基準にして、項目を記載することもできます。代わりに、レポートに表示しない項目を確認するために、除外することを選択した項目だけを記載することもできます (「レポートに表示」の選択を解除する)。

選択 : レポート内で「最近使用した項目」のみを選択することで、プロジェクトの更新内容を提供できるようにします。代わりに「古い項目」のパラメーターを設定して、古い項目や未解決の問題を追跡することもできます。レポートで時間的なことを考慮しない場合は、「すべての項目」を選択します。

表示順。フォルダー内のオブジェクトの順序を決めるために、次のオブジェクト特性の組み合わせを選択できます。

- **タイプ**: データ型を基準にオブジェクトをグループ化します。
- **「名前」**。オブジェクトをアルファベット順に並べます。
- **追加日** : プロジェクトに追加された日付を使用して、オブジェクトをソートします。

生成されたレポートの保存とエクスポート

画面に生成されたレポートは、新規出力ウィンドウに表示されます。レポートに含まれるグラフは、インラインの画像として表示されます。

レポート用語集

各ストリームのノードの合計数がレポート内に表示されます。次の見出しに数が表示され、CRISP-DM 用語集ではなく IBM SPSS Modeler 用語集が使用されます。

- データ・リーダー：入力ノード。
- データ・ライター：エクスポート・ノード。
- モデル・ビルダー：構築ノードまたはモデル作成ノード。
- モデル・アプライヤー：ナゲットとも呼ばれる、生成されたモデル。
- 出力ビルダー：グラフ・ノードまたは出力ノード。
- その他: プロジェクトに関連するその他のノード。例えば、ノード・パレットの「フィールド設定」タブまたは「レコード設定」タブで使用できるノード。

レポートを保存する手順は次のとおりです。

1. 「ファイル」メニューで「保存」をクリックします。
2. ファイル名を指定します。

レポートは、出力オブジェクトとして保存されます。

レポートをエクスポートする手順は次のとおりです。

3. 「ファイル」メニューで「エクスポート」をクリックし、エクスポートするファイルの種類を選択します。
4. ファイル名を指定します。

レポートは、選択した形式で保存されます。

次の種類のファイルへエクスポートできます。

- HTML
- テキスト
- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint

注：Microsoft Office ファイルをエクスポートするには、対応するアプリケーションがインストールされている必要があります。

ウィンドウの上部にあるボタンを使用して、次の作業を実行します。

- レポートを印刷します。
- 外部の Web ブラウザー内で HTML としてレポートを表示します。

第 13 章 IBM SPSS Modeler のカスタマイズ

IBM SPSS Modeler オプションのカスタマイズ

必要に応じて IBM SPSS Modeler をカスタマイズするためには、さまざまな操作があります。このカスタマイズは、主にメモリーの割り当て、デフォルトのディレクトリー、音声および色の使用など、特定のユーザー・オプションの設定で構成されています。また、IBM SPSS Modeler ウィンドウの下部にあるノート・パレットをカスタマイズすることもできます。

IBM SPSS Modeler オプションの設定

IBM SPSS Modeler のオプションをカスタマイズしたり、設定するには、さまざまな方法があります。

- メモリー使用量やロケールなどのシステム・オプションを設定するには、「ツール > オプション」メニューの「システム オプション」をクリックします。
- 表示フォントや色などのユーザー・オプションを設定するには、「ツール > オプション」メニューの「ユーザー オプション」をクリックします。
- IBM SPSS Modeler と連携するアプリケーションの場所を指定するには、「ツール > オプション」メニューの「ヘルパー アプリケーション」をクリックします。
- IBM SPSS Modeler が使用するデフォルト・ディレクトリーを指定するには、「ファイル」メニューの「ディレクトリーの設定」または「サーバー ディレクトリーの設定」をクリックします。

また、ストリームの一部またはすべてに適用するオプションを設定することができます。詳しくは、トピック 42 ページの『ストリームのオプションの設定』を参照してください。

システム・オプション

IBM SPSS Modeler の使用言語またはロケールを指定するには、「ツール > オプション」メニューの「システム・オプション」をクリックします。ここでは、SPSS Modeler の最大メモリー使用量の設定や、ストリームを自動的に保存する頻度の指定もできます。このダイアログ・ボックスで行った変更内容は、SPSS Modeler を再起動しないと有効にはならないことに注意してください。

最大メモリー: IBM SPSS Modeler のメモリー使用量を MB 単位で制限する場合に選択します。一部のプラットフォームでは、リソースに制約のあるコンピューターや負荷の高いコンピューターの負担を軽減するために、SPSS Modeler はプロセスのサイズを制限しています。そのため、大量のデータを扱うと、「メモリー不足」のエラーが発生することがあります。ここに新しい値を指定して、メモリーの負荷を調整することができます。

例えば、非常に大規模なディシジョン ツリーを表示しようとする、メモリー エラーが生じる場合があります。この場合、メモリーを最大値の 4096Mb まで増やすことをお勧めします。こうしたケースのように非常に大量のデータを処理する可能性がある場合は、メモリー許容量を増やした後に SPSS Modeler をシャットダウンしてコマンド ラインから始動することで、データの処理時に確実に最大量のメモリーが使用されるようにしてください。

コマンド ラインから始動するには (SPSS Modeler がデフォルトの場所にインストールされていると想定した場合)、コマンド プロンプト ウィンドウに以下を入力します。

```
C:\Program Files\IBM\SPSS\Modeler\18.1\bin\modelerclient.exe" -J-Xss4096M
```

システム ロケールを使用：このオプションはデフォルトで選択され、「日本語」に設定されています。他の言語を指定する場合は、このオプションの選択を解除して、リストから適切な言語やロケールを選択してください。

ストリーム自動保存間隔 (分): SPSS Modeler でストリームを自動的に保存する頻度を指定します。最大値は 60 分、最小値は 1 分、デフォルト値は 5 分です。

メモリーの管理

「システム オプション」ダイアログ・ボックスの「最大メモリー」の設定以外にも、メモリーの使用を最適化するためのさまざまな方法があります。

- 「ストリームのプロパティ」ダイアログ・ボックスの「名義型フィールドの最大メンバー」オプションを調整する。このオプションには、名義型フィールドの最大メンバー数を指定します。これ以降のフィールドの測定尺度は、データ型不明になります。詳しくは、トピック 43 ページの『ストリームの一般的なオプションの設定』を参照してください。
- IBM SPSS Modeler のウィンドウの右下に表示されている、IBM SPSS Modeler が使用中のメモリーおよびメモリーの総割り当て量 (xxMB/xxMB) の領域をクリックして、メモリーを解放する。この領域をクリックすると、いったん領域が暗くなって押されたような状態になります。その後領域が元の色に戻ると、IBM SPSS Modeler で解放できるメモリーがすべて解放されます。

デフォルト・ディレクトリーの設定

ファイルの参照や出力に使われるデフォルト・ディレクトリーを指定するには、「ファイル」メニューの「ディレクトリーの設定」または「サーバー ディレクトリーの設定」を選択します。

- ディレクトリーの設定 :作業ディレクトリーを設定します。デフォルトの作業ディレクトリーは、IBM SPSS Modeler のインストール場所、または IBM SPSS Modeler を起動するために用いられたコマンド・ラインのパスをベースにしています。ローカル・モードの場合、作業ディレクトリーはクライアント側の操作や出力ファイルで使われたパスになります (相対パスで参照された場合)。
- サーバー ディレクトリーの設定 : リモート・サーバーへの接続がある場合、「ファイル」メニューの「サーバー ディレクトリーの設定」オプションは常に有効になっています。このオプションを使用して、入力や出力に使われるすべてのサーバー・ファイルやデータ・ファイルのデフォルト・ディレクトリーを指定します。デフォルトのサーバー・ディレクトリーは、`$CLEO/data` になります。ここで、`$CLEO` は IBM SPSS Modeler のサーバー・バージョンがインストールされたディレクトリーを表します。コマンド・ラインで `-server_directory` フラグと `modelerclient` コマンド・ライン引数を使用して、このデフォルトを上書きすることもできます。

ユーザー・オプションの設定

IBM SPSS Modeler の一般オプションを設定するには、「ツール > オプション」メニューの「ユーザーオプション」を選択します。これらのオプションは、IBM SPSS Modeler で使用するすべてのストリームに適用されます。

ここでは、次の種類のオプションを設定することができます。これらのオプションを設定するには、該当するタブをクリックしてください。

- 「通知」オプション : モデルの上書きやエラー・メッセージなどのオプションを設定します。
- 「表示」オプション : グラフや背景の色を設定します。
- シンタックスの色表示オプション。
- モデルを Predictive Model Markup Language (PMML) へエクスポートするときに使った PMML エクスポート・オプション。

- 名前、イニシャル、メール・アドレスなどのユーザまたは作成者の情報。この情報を、作成したノードまたはその他のオブジェクトの「注釈」タブに表示することができます。
- 従来型モードと Analytic Server モードの間での切り替え。

数字区切り文字、時刻や日付の形式、最適化、ストリーム・レイアウト、およびストリーム・スクリプトなどのストリーム固有のオプションの設定は、「ストリームのプロパティ」ダイアログ・ボックスから行います。このダイアログ・ボックスは、「ファイル」メニューと「ツール」メニューから利用できます。

通知オプションの設定

「ユーザー オプション」ダイアログ・ボックスの「通知」タブでは、IBM SPSS Modeler の警告および確認ウィンドウの表示や種類に関するさまざまな設定を指定できます。また、新しい出力やモデルが生成された時の、マネージャー・パネルの「出力」タブおよび「モデル」タブの動作も設定することができます。

ストリーム実行フィードバック・ダイアログの表示: このオプションを選択すると、ストリームが 3 秒間実行されている場合、進行状況表示を含むダイアログ・ボックスが表示されます。ダイアログ・ボックスには、ストリームで作成された出力オブジェクトの詳細も表示されています。

- 完了時にダイアログが閉じます: デフォルトでは、ストリームが実行を完了すると、このダイアログ・ボックスが閉じられます。ストリームの実行が完了した後もダイアログ・ボックスを表示したままにする場合は、このボックスの選択を解除 (チェックを外す) します。

ノードがファイルを上書きする時に警告: このオプションを選択すると、ノード操作により既存のファイルが上書きされる場合に、エラー メッセージで警告されます。

ノードがデータベース テーブルを上書きする時に警告: このオプションを選択すると、ノード操作により既存のデータベース テーブルが上書きされる場合に、エラー・メッセージで警告されます。

音声通知

下にあるリストを使用して、イベントやエラーが発生した時に通知するサウンドを指定します。さまざまな音声やサウンドを利用できます。選択したサウンドを再生する場合は、「再生」ボタン (スピーカー) を使用します。サウンド・ファイルを探して選択する場合は、「...」ボタンを使用します。

注: IBM SPSS Modeler のサウンドを作成するために使用される .wav ファイルは、ご使用のコンピュータの /media/sounds ディレクトリにあります。

- すべての音をミュート: すべてのイベントのサウンド通知をオフにする場合に選択します。

視覚通知

このグループのオプションを使用して、新規項目生成時のマネージャー・パネル右上の「出力」タブおよび「モデル」タブの動作を指定します。リストから「新規モデル」または「新規出力」を選択して、対応するタブの動作を設定してください。

「新規モデル」には、次のオプションを使用できます。

前のモデルを置換: このオプションを選択した場合 (デフォルト)、「モデル」タブおよびストリーム領域の、このストリームの既存モデルを上書きします。このボックスをオフにすると、「モデル」タブおよび領域の既存モデルにモデルが追加されます。この設定は、モデルリンクのモデル置換設定によって上書きされます。

「新規出力」には、次のオプションを使用できます。

出力超過時に警告 [n]: 「出力」タブの項目数が事前に指定された数を超えた場合に、警告を表示するかどうかを選択します。デフォルトの数は 20 です。必要に応じて、この数を変更することができます。

すべてのケースで次のオプションを利用できます。

タブを選択: ストリームの実行時に、対応するオブジェクトが生成された場合に、「出力」タブまたは「モデル」タブに切り替えるかどうかを選択します。

- マネージャー・パネルの該当するタブに切り替える場合は、「常時」を選択します。
- 現在領域に表示されているストリームによってオブジェクトが生成された時にだけタブに切り替える場合は、「現在のストリームが生成した場合」を選択します。
- 出力やモデルが生成されたことを、対応するタブに切り替えて知らせない場合は、「しない」を選択します。

タブ点滅: 新しい出力やモデルが生成された場合に、マネージャー・パネルにある「出力」タブまたは「モデル」タブを点滅させるかどうかを選択します。

- マネージャー・パネル中の対応するタブを点滅させる場合は (タブが選択されていない場合)、「選択されていない場合」を選択します。
- オブジェクトが生成されたことを、対応するタブを点滅させて知らせない場合は、「しない」を選択します。

パレットを見えるようにスクロール (「新規モデル」のみ): 最新のモデルが見えるように、マネージャー・パネルの「モデル」タブを自動的にスクロールするかどうかを選択します。

- スクロールする場合は、「常時」を選択します。
- 現在領域に表示されているストリームによってオブジェクトが生成された時にだけスクロールする場合は、「現在のストリームが生成した場合」を選択します。
- 「モデル」タブの自動スクロールを行わない場合は、「しない」を選択します。

ウィンドウを開く (「新規出力」のみ): 生成時に出力ウィンドウを自動的に開くかどうかを選択します。

- 常に新しく出力ウィンドウを開く場合は、「常時」を選択します。
- 現在領域に表示されているストリームによって出力が生成された時にだけ新しいウィンドウを開く場合は、「現在のストリームが生成した場合」を選択します。
- 生成された出力に対して自動的に新しいウィンドウを開かない場合は、「しない」を選択します。

このタブのシステム デフォルト設定に戻すには、「デフォルト値」を選択します。

表示オプションの設定

「ユーザー オプション」ダイアログ・ボックスの「表示」タブを使用して、IBM SPSS Modeler のフォントや色の表示を設定することができます。

よろこダイアログが起動時に表示: よろこダイアログが起動時に表示されることを選択します。よろこダイアログには、アプリケーションの例チュートリアルを起動するオプション、デモ ストリームまたは既存のストリームやプロジェクトを開くオプション、または新規ストリームを作成するオプションがあります。

ストリームおよびスーパーノードのマークアップを表示: このオプションを選択した場合、デフォルトで、ストリームおよびスーパーノードのマークアップ (ある場合) が表示されます。マークアップはストリームのコメント、モデル リンク、およびスコアリング ブランチの強調表示で構成されています。

標準のフォントと色 (再起動後に有効): このコントロール ボックス内のオプションを使用して、IBM SPSS Modeler の画面デザイン、カラー スキーム、および表示されるフォントのサイズを指定します。ここで選択したオプションは、IBM SPSS Modeler を終了し再起動しないと有効になりません。

- 外観: 標準のカラー スキームおよび画面デザインを選択します。以下から選択できます。

SPSS スタンダード (デフォルト): IBM SPSS 製品共通のデザイン。

SPSS クラシック: 以前のバージョンの IBM SPSS Modeler のユーザーにお馴染みのデザイン。

Windows: ストリーム領域やパレットのコントラストを強くする場合に役立つ Windows のデザイン。

- ノードのデフォルト フォント サイズ: ノード パレット内で使用され、ストリーム領域に表示されるノードに使用されるフォント サイズを指定します。
- 固定幅フォントを指定: スクリプトおよび CLEM 式コントロールに使用する、固定幅フォントおよび関連するフォント サイズを選択するには、このチェック ボックスを選択します。デフォルト フォントは Monospace plain です。選択可能な他のフォントのリストを表示するには、「変更...」をクリックします。

注: 「ストリームのプロパティ」ダイアログ ボックスの「オプション」タブの「レイアウト」ペインでストリームのノード アイコンのサイズを設定できます。メイン メニューから、「ツール」 > 「ストリームのプロパティ」 > 「オプション」 > 「レイアウト」を選択します。

カスタム カラー: このテーブルには、さまざまな表示項目に使用される、現在選択されている色がリストされています。テーブルにリストされた各項目について、「色」列の該当する行をダブルクリックし、リストから色を選択して、現在の色を変更できます。カスタム カラーを指定するには、リストの最後にスクロールして、「色...」 エントリーをクリックします。

グラフ カテゴリの色順序: このテーブルには、新しく作成したグラフの表示に使用される、現在選択されている色がリストされています。色の順序が、グラフで使用される順序を表しています。例えば、色のオーバーレイに使用されている名義型フィールドに 4 種類の固有値がある場合、ここにリストされている先頭から 4 つの色のみが使用されます。テーブルにリストされた各項目について、「色」列の該当する行をダブルクリックし、リストから色を選択して、現在の色を変更できます。カスタム カラーを指定するには、リストの最後にスクロールして、「色...」 エントリーをクリックします。ここで行った変更は、それまでに作成したグラフには適用されません。

このタブのシステム デフォルト設定に戻すには、「デフォルト値」を選択します。

シンタックス表示オプションの設定

「ユーザー オプション」ダイアログ・ボックスの「シンタックス」タブを使用して、IBM SPSS Modeler で作成するスクリプトのフォント属性や表示色のオプションを設定できます。

シンタックスの強調表示。 このテーブルには、さまざまなシンタックス項目 (フォントやそれが表示されるウィンドウの両方を含む) に使われる色の、現在の設定が表示されています。テーブルに表示された各項目について、行の該当するドロップダウン・リストをクリックし、リストから色を選択して色を変更できます。また、フォント項目については、太字や斜体の強調を選択できます。

プレビュー: このテーブルには、「シンタックスの強調表示」テーブルで選択した色やフォント属性が使用されたサンプルのシンタックス表示が示されます。このプレビューは、ユーザーが選択を行うとすぐに更新されます。

このタブのデフォルト設定に戻す場合は、「デフォルト値」を選択します。

PMML エクスポート・オプション の設定

「PMML」タブで、IBM SPSS Modeler から Predictive Model Markup Language (PMML) ヘモデルをエクスポートする方法を制御できます。詳しくは、トピック 184 ページの『PMML としてのモデルのインポートおよびエクスポート』を参照してください。

PMML をエクスポート: PMML を対象アプリケーションと最高の状態で動作させるための各種設定をここでを行います。

- 「拡張機能」を選択し、標準 PMML 相当する機能のない特殊なケースに PMML 拡張機能を使用できるようにします。ほとんどのケースで、標準 PMML と同じ結果が生成されます。
- 「標準 PMML として」を選択し、PMML 標準にできる限り準拠するように PMML をエクスポートします。

標準 PMML オプション: 「標準 PMML として...」オプションを選択した場合には、線型回帰モデルとロジスティック回帰モデルをエクスポートするのに有効な方法を、次の 2 つから選択できます。

- **PMML <GeneralRegression>** モデルとして
- **PMML <Regression>** モデルとして

PMML の詳細は、データ・マイニング・グループの Web サイト (<http://www.dmg.org>) を参照してください。

ユーザー情報の設定

ユーザーと作成者の情報: ここで入力した情報は、作成したノードおよびその他のオブジェクトの「注釈」タブに表示されます。

モードの設定

モデラー モード設定: 「モード」タブで、以下のモードから選択できます。

- 「従来 **SPSS Modeler** モード」は、ユーザー・インターフェースにすべての使用可能なノードおよび式を表示します。
- 「**Analytic Server** モード」は、Analytic Server でサポートされるノードおよび式のみを表示します。しかし、ノードおよび CLEM 式の中には、Analytic Server で完全にサポートされていなくても依然として表示されるものがあるということに注意してください。以下の表は、Analytic Server でサポートされるノード、部分的にサポートされるノード、およびサポートされないノードに関する一般情報を示します。

詳細情報は、サポートされるノードを参照してください。

Analytic Server について詳しくは、Analytic Server の資料を参照してください。

表 37. ノードのサポート

ノード・タイプ (パレット名)	Analytic Server でサポート	Analytic Server で部分サポート	Analytic Server で非サポート
入力	<ul style="list-style-type: none"> • Analytic Server ソース・ノード 		<ul style="list-style-type: none"> • データベース • 可変長ファイル • 固定長ファイル • Statistics ファイル • Data Collection • IBM Cognos • TWC インポート • TM1 インポート • SAS ファイル • Excel • XML ユーザー入力 • シミュレーション生成 • データ・ビュー • 地理空間 • オブジェクト ストレージ • 拡張インポート • R インポート • SNA (拡散分析とグループ分析)
レコード設定	<ul style="list-style-type: none"> • 条件抽出 • ソート • バランス • 重複レコード • RFM 集計 • レコード追加 • ストリーミング時系列分析 • 拡張変換 • ストリーミング TCM 	<ul style="list-style-type: none"> • サンプル (シンプルな方法の Random% のみサポートします。複雑な方法はサポートされません。) • レコード結合 (キーおよび条件による結合のみサポートします) • レコード集計 (第 1 四分位数、第 3 四分位数、および中央値は Analytic Server でサポートされません) 	<ul style="list-style-type: none"> • スペース タイム ボックス • CPLEX の最適化

表 37. ノードのサポート (続き)

ノード・タイプ (パレット名)	Analytic Server でサポート	Analytic Server で部分サポート	Analytic Server で非サポート
フィールド設定	<ul style="list-style-type: none"> • データ型 • フィルター • フィールド作成 • 置換 • データ分類 • アンサンブル • フラグ設定 • 再構成 • フィールド順序 • 再投影 • 時間間隔 	<ul style="list-style-type: none"> • 自動データ準備 (変換のみサポートします) • データ分割 (同順位設定の「現在のまま保持」が選択されている場合、データ分割手段 equalFreq はサポートされません) • RFM 分析 (同順位設定の「現在のまま保持」が選択されている場合、データ分割手段 Tiles はサポートされません) • データ区分 (一意のフィールドを使用してデータ区分に対して行を繰り返し割り当てない限り、Analytic Server でサポートされません) • 行列入替 (一般的な行列入替は Analytic Server でサポートされませんが、再構成は Python for Spark を使用して Analytic Server でサポートされます) 	<ul style="list-style-type: none"> • 匿名化 • 時系列
グラフ作成	<ul style="list-style-type: none"> • 作図 • マルチ散布図 • 時系列グラフ • 分布 • ヒストグラム • 集計棒グラフ • Web • 評価 • マップ視覚化 	<ul style="list-style-type: none"> • グラフボード (測定の尺度が離散、名義、順序、またはフラグであるフィールドに対する集計モード関数のみサポートします) 	

表 37. ノードのサポート (続き)

ノード・タイプ (パレット名)	Analytic Server でサポート	Analytic Server で部分サポート	Analytic Server で非サポート
モデリング	<ul style="list-style-type: none"> 時系列 TCM ランダム ツリー Tree-AS Linear-AS GLE LSVM 生存 STP TwoStep-AS アソシエーション・ルール 	<ul style="list-style-type: none"> 自動分類 自動数値 (この 2 つのノードは分割のみをサポートし、自動分類オプションの「Analytic Server で実行 (分割が有効)」を使用する場合は、分割役割を持つフィールドを指定する必要があります) 拡張 (R シンタックスのモデル構築は、Analytic Server でサポートされません) 分割および PSM をサポートするのは、以下のノードのみです。 <ul style="list-style-type: none"> C&R Tree 線型 ニューラル・ネットワーク CHAID QUEST 	<ul style="list-style-type: none"> 自動クラスター ディジション・リスト C5.0 回帰 因子分析 変数選択 判別 ロジスティック回帰 一般化線型 GLMM ベイズ・ネット Apriori Carma シーケンス K-Means Kohonen TwoStep 異常値 KNN R 以下のノードには ASL がありますが、ただの読み取り/書き込みの ASL です。 <ul style="list-style-type: none"> Cox SVM SLRM
出力	<ul style="list-style-type: none"> 表 クロス集計 分析 データ検査 変換 記述統計 平均値 レポート グローバル 	<ul style="list-style-type: none"> 拡張出力 R シンタックス出力 	<ul style="list-style-type: none"> シミュレーション評価 シミュレーションの当てはめ R 出力

表 37. ノードのサポート (続き)

ノード・タイプ (パレット名)	Analytic Server でサポート	Analytic Server で部分サポート	Analytic Server で非サポート
エクスポート	<ul style="list-style-type: none"> Analytic Server エクスポート・ノード 	<ul style="list-style-type: none"> 拡張エクスポート R シンタックス エクスポート 	<ul style="list-style-type: none"> データベース フラット・ファイル Statistics エクスポート Data Collection Excel IBM Cognos エクスポート TM1 エクスポート SAS XML エクスポート オブジェクト ストレージ R エクスポート
IBM SPSS Statistics			<ul style="list-style-type: none"> Statistics ファイル Statistics 変換 Statistics モデル Statistics 出力 Statistics エクスポート
IBM SPSS Text Analytics	<ul style="list-style-type: none"> テキスト リンク分析 テキスト・マイニング 言語ノード 		<ul style="list-style-type: none"> ファイル リスト Web フィード テキスト リンク分析 翻訳 テキスト・マイニング ファイル ビューアー
Python			<ul style="list-style-type: none"> SMOTE One-Class SVM XGBoost Tree XGBoost Linear

ノード・パレットのカスタマイズ

ストリームは、ノードを使用して構築されます。IBM SPSS Modeler ウィンドウの下部にあるノード・パレットには、ストリームの構築に使われるすべてのノードが用意されています。詳しくは、トピック 16 ページの『ノード・パレット』を参照してください。

ノード・パレットは、次の 2 つの方法で再編成することができます。

- パレット・マネージャーをカスタマイズする。詳しくは、トピック 205 ページの『パレット・マネージャーのカスタマイズ』を参照してください。
- サブパレットを含むパレットがノード・パレットにどのように表示されるかを変更する。詳しくは、トピック 207 ページの『サブパレットの作成』を参照してください。

パレット・マネージャのカスタマイズ

パレット・マネージャを、IBM SPSS Modeler の使用方法に合わせてカスタマイズすることができます。例えば、データベースの時系列データを頻繁に分析するような場合、データベース入力ノード、時間区分ノード、時系列ノード、時系列グラフ・ノードがすべて一意のパレットのタブで使用できるようにしておく便利です。パレット・マネージャを使用すると、これらのノードをノード・パレットのカスタム・パレット・タブを作成して容易に調整することができます。

パレット・マネージャを使用して、次のようなさまざまなタスクを実行することができます。

- ストリーム領域の下のノード・パレットに表示されるパレット・タブを制御する。
- パレット・タブがノード・パレットに表示される順序を変更する。
- パレット・タブおよび関連するサブ パレットを作成および編集する。
- タブのデフォルトのノード選択を編集する。

パレット・マネージャにアクセスするには、「ツール」メニューの「パレット管理」をクリックします。

パレット名: ノード・パレット上に表示されているかどうかに関係なく、使用できるパレット・タブが表示されます。ここには、作成したパレット・タブが表示されます。詳しくは、トピック『パレット・タブの作成』を参照してください。

ノード数: 各パレット・タブに表示されるノード数。数が多いほど、タブのノードを分割するサブ パレットを作成しやすくなります。詳しくは、トピック 207 ページの『サブパレットの作成』を参照してください。

表示 ?: ノード・パレット上にパレット・タブを表示するには、このフィールドを選択します。詳しくは、トピック 206 ページの『ノード・パレットのパレット・タブの表示』を参照してください。

サブ パレット: パレット・タブに表示するサブ パレットを選択するには、該当する「パレット名」強調表示し、このボタンをクリックすると「サブ パレット」ダイアログ・ボックスが表示されます。詳しくは、トピック 207 ページの『サブパレットの作成』を参照してください。

復元デフォルト。 このボタンをクリックすると、パレットおよびサブ パレットに行ったすべての変更および追加を完全に削除し、デフォルトのパレット設定に戻します。

パレット・タブの作成

カスタム・パレット・タブを作成する手順は、次のとおりです。

1. 「ツール」メニューからパレット・マネージャを起動します。
2. 「表示 ?」列の右側にある「パレットを追加」ボタンをクリックすると、「パレットの作成/編集」ダイアログ・ボックスが表示されます。
3. 一意の「パレット名」に入力します。
4. 「使用可能なノード」領域で、パレット・タブに追加されるノードを選択します。
5. 「ノードを追加」の右矢印ボタンをクリックして、選択強調表示されているノードを「選択されたノード」領域に移動します。該当するすべてのノードを追加するまで、この手順を繰り返します。

該当するノードをすべて追加した後、次のようにしてパレット・タブに表示される順序を変更することができます。

6. 単純な矢印ボタンをクリックすると、ノードが 1 行上または下に移動します。
7. 線の付いた矢印ボタンを使用すると、ノードがリストの最上位または最下位に移動します。

8. ノードをパレットから削除するには、「選択されたノード」領域の右側にある「削除」ボタンをクリックします。

ノード・パレットのパレット・タブの表示

IBM SPSS Modeler 内に使用しないオプションがある場合があります。この場合、パレット・マネージャーを使用して、これらのノードを含むタブを非表示にすることができます。

ノード・パレットに表示されるタブを選択する手順は、次のとおりです。

1. 「ツール」メニューからパレット・マネージャーを起動します。
2. 「表示？」列のチェック・ボックスを使用して、各パレット・タブを表示するか非表示にするかを選択します。

ノード・パレットからパレット・タブを完全に削除するには、ノードを強調表示して、「表示？」列の右側にある「削除」ボタンをクリックします。削除したパレット・タブを復元することはできません。

注：「お気に入り」タブを除き、IBM SPSS Modeler で提供されたデフォルトのパレット・タブを削除することはできません。

ノード・パレット上の表示順の変更

表示するパレット・タブを選択した後、次のようにしてノード・パレットに表示される順序を変更することができます。

1. 単純な矢印ボタンをクリックすると、パレット・タブが 1 行上または下に移動します。上に移動するとノード・パレットの左側に移動し、下に移動するとノード・パレットの右側に移動します。
2. 線の付いた矢印ボタンを使用すると、パレット・タブがリストの最上位または最下位に移動します。リストの最上位にあるタブ パレットは、ノード・パレットの左側に表示されます。

「パレット」タブのサブパレットの表示

ノード・パレットに表示されるパレット・タブを制御するのと同じように、親パレット・タブで使用可能なサブパレットを制御することができます。

パレット・タブで表示するサブパレットを選択する手順は、次のとおりです。

1. 「ツール」メニューからパレット・マネージャーを起動します。
2. 必要なパレットを選択します。
3. 「サブパレット」ボタンをクリックすると、「サブパレット」ダイアログ・ボックスが表示されます。
4. 「表示？」列のチェック・ボックスを使用して、各サブパレットを「パレット」タブに追加するかどうかを選択します。「すべて」のサブパレットは常に表示され、削除することはできません。
5. パレット・タブからサブパレットを完全に削除するには、サブパレットを強調表示して、「表示？」列の右側にある「削除」ボタンをクリックします。

注：「モデル作成パレット」タブで提供されているデフォルトのサブパレットを削除することはできません。

パレット・タブ上の表示順の変更

表示するサブパレットを選択した後、次のようにして親パレットタブに表示される順序を変更することができます。

1. 単純な矢印ボタンをクリックすると、サブパレットが 1 行上または下に移動します。
2. 線の付いた矢印ボタンを使用すると、サブパレットがリストの最上位または最下位に移動します。

親パレット・タブを選択すると、作成したサブパレットがノード・パレットに表示されます。詳しくは、トピック『パレット・タブの表示の変更』を参照してください。

サブパレットの作成

既存のノードを作成するカスタム・パレット・タブに追加することができるため、スクロールせずに容易に表示できないほど多くのノードを選択することができます。スクロールしなくてもすむようにするために、パレット・タブに選択するノードを配置するサブパレットを作成することができます。例えば、ストリーム作成に最も頻繁に使用するノードを含むパレット・タブを作成する場合、入力ノード、フィールド設定、モデル作成および出力で選択を分割するサブパレットを作成することができます。

注：親パレット・タブに追加されたサブパレット・ノードからのみ、選択することができます。

サブパレットを作成するには

1. 「ツール」メニューからパレット・マネージャーを起動します。
2. サブパレットを追加するパレットを選択します。
3. 「サブパレット」ボタンをクリックすると、「サブパレット」ダイアログ・ボックスが表示されます。
4. 「表示？」列の右側にある「サブパレットを追加」ボタンをクリックすると、「サブパレットの作成/編集」ダイアログ・ボックスが表示されます。
5. 一意の「サブパレット名」に入力します。
6. 「使用可能なノード」領域で、サブパレットに追加されるノードを選択します。
7. 「ノードを追加」の右の矢印のボタンをクリックして、選択されたノードを「選択したノード」に移動します。
8. 該当するノードを追加したら、「OK」をクリックして「サブパレット」ダイアログ・ボックスに戻ります。

親パレット・タブを選択すると、作成したサブパレットがノード・パレットに表示されます。詳しくは、トピック『パレット・タブの表示の変更』を参照してください。

パレット・タブの表示の変更

IBM SPSS Modeler で使用できるノード数が多すぎるため、ノード・パレットの左側または右側へのスクロールが必要ない小さい画面に表示することができません。これは、「モデル作成パレット」タブで特に注意する必要があります。スクロールを少なくするために、サブパレットに含まれるノードのみを表示することができます (該当する場合)。詳しくは、トピック『サブパレットの作成』を参照してください。

パレット・タブに表示されるノードを変更するには、パレット・タブを選択し、左側のメニューですべてのノードを表示するか、特定のサブパレットのノードのみを表示するかを選択します。

第 14 章 ストリームおよびノードのパフォーマンスの考察

ストリームは、最も効果的な構成でノードを配置したり、必要に応じてノード・キャッシュを有効にしたり、このセクションで説明するその他の検討事項に注意を払うことによって、パフォーマンスが最大になるように設計できます。

ここで説明される検討事項のほかに、データベースの効果的な使用、特に SQL 最適化を利用することによって、一般的に、より実質的な追加のパフォーマンス改善が実現します。

ノードの順序

SQL 最適化を使用していなくても、ストリーム内のノードの順序がパフォーマンスに影響を与えることがあります。一般的な目的は下流の処理を最小化することです、そのため、データの量を削減するノードがある場合、ストリームの初めあたりにそれらを配置します。IBM SPSS Modeler Server では、コンパイル時に自動的に順序変更ルールを適用し、安全であると判断した場合に特定のノードを進めることができます (この機能は、デフォルトで有効です。お使いのコンピューターでこの機能が有効であることをシステム管理者に確認してください)。

SQL 最適化を使用する場合は、可用性と効率性を最大限にするように設定します。データベースで実行できない操作がストリームに含まれているときに最適化が停止するので、ストリームの先頭に SQL 最適化の操作をグループ化するのが、もっともよい方法です。このストラテジによりデータベース内での処理が増えるので、IBM SPSS Modeler へ持ち込まれるデータが少なくなります。

ほとんどのデータベース内で、次の操作を実行できます。ストリームの「先頭」に次の操作をグループ化するようにしてください。

- キーによる結合
- 条件抽出
- レコード集計
- ソート
- サンプリング
- レコード追加
- すべてのフィールドが選択される *include* (含む) モードでの、識別の操作
- 置換操作
- 標準的な算術式または文字列操作を使用した、基本的な算出操作 (どの操作がデータベースにサポートされるかに依存)
- フラグ設定

以下の操作は、ほとんどのデータベース内で実行できません。上記リストの操作の「後」に、次の操作をストリームに配置する必要があります。

- フラット・ファイルのような、非データベースのデータに対する操作
- 順序による結合
- バランス

- *discard* モードでの識別の操作、またはフィールドのサブセットのみが異種として選択された場合の識別の操作
- 処理対象以外のレコードからのデータへのアクセスが必要な操作
- ステート型および度数型フィールドの導出
- 時系列ノードの操作
- "@" (時系列) 関数に関連する操作
- データ型チェック・モードの「警告」と「中止」
- モデル作成、アプリケーション、および分析

注：ディジション・ツリー、ルールセット、線型回帰、および因子生成のモデルで、SQL を生成し、データベースへプッシュバックできます。

- データを処理中の同じデータベース以外の任意の場所へのデータ出力

ノードのキャッシュ

ストリームの実行を最適化するために、ターミナル・ノード以外の任意のノードにキャッシュを設定することができます。ノードにキャッシュを設定すると、次のデータ・ストリームの実行時にそのノードを通過するデータがキャッシュされます。データがキャッシュされたら、以降のデータはデータ・ソースの代わりに、キャッシュ (ディスクの一時ディレクトリーに格納) から読み込まれます。

キャッシュは、ソート、結合、集計など、時間のかかる操作に最も有用です。例えば、データベースから販売データを読み込む入力ノードと、販売地域ごとに集計するレコード集計ノードがあるとします。この場合、データ・セット全体ではなく集計済みデータをキャッシュしたいので、入力ノードではなくレコード集計ノードにキャッシュを設定します。

注：入力ノードでキャッシュする場合、IBM SPSS Modeler に読み込まれるように元のデータのコピーを保存するため、多くの環境でパフォーマンスは改善されません。

キャッシュを有効にしたノードの右上には、小さい文書アイコンが表示されます。データがキャッシュされると、この文書アイコンが緑に変わります。

キャッシュを有効にするには

1. ストリーム領域でノードを右クリックし、メニューの「キャッシュ」をクリックします。
2. キャッシュ設定用のサブメニューから、「使用する」を選択します。
3. キャッシュを解除するには、ノードを右クリックして、キャッシュ設定用のサブメニューから「使用しない」を選択します。

データベース内でのノードのキャッシュ

データベース内で実行されるストリームの場合、ファイル・システムでなくデータベース内の一時テーブルへ、データを中流でキャッシュできます。この機能を SQL 最適化と組み合わせると、パフォーマンスが著しく向上する可能性があります。例えば、データ・マイニング・ビューを作成するために複数のテーブルを併合するストリームからの出力をキャッシュし、必要に応じて再使用できます。すべての下流ノードで自動的に SQL を生成するようにすると、パフォーマンスがさらに向上する可能性があります。

データベースのキャッシングを利用するには、SQL 最適化とデータベース・キャッシングの両方を有効にする必要があります。Client の設定は、Server の最適化設定によって無効になるので注意してください。詳しくは、トピック 45 ページの『ストリームの最適化オプションの設定』を参照してください。

データベースのキャッシングが有効にされると、任意の非ターミナル・ノードを右クリックするだけでその場所のデータがキャッシュされ、次にストリームが実行されるときに自動的に、データベース内にキャッシュが直接作成されます。データベース・キャッシングまたは SQL 最適化が有効にされていないと、キャッシュは、代わりにファイル・システムへ書き出されます。

注: 次のデータベースは、キャッシュのために一時テーブルをサポートします。

DB2、Netezza、Oracle、SQL Server および Teradata。その他のデータベースでは、データベース・キャッシングに通常のテーブルが使用されます。SQL コードを特定のデータベース向けにカスタマイズできません。サービスに連絡してください。

パフォーマンス : プロセス・ノード

ソート : ソート・ノードでは、ソートする前に入力データ・セット全体を読み込む必要があります。データは、一定の限度までメモリーに格納され、超過分はディスクへ書き出されます。ソート・アルゴリズムは組合せのアルゴリズムです。データは上限までメモリーに読み込まれ、高速で複合的なクイックソート・アルゴリズムを使用してソートされます。すべてのデータがメモリーに収まれば、ソートはこれで完了です。そうでない場合は、マージソートのアルゴリズムが適用されます。ソート済みのデータがファイルへ書き込まれ、次のデータの塊りがメモリーに読み込まれ、ソートされ、ディスクへ書き込まれます。この過程はすべてのデータが読み込まれるまで繰り返され、その後、ソート済みのデータの塊りが併合されます。この併合も、ディスクに格納されたデータに対して繰り返して行うことが必要な場合があります。処理のピーク時には、ソート・ノードが、ソート済みと未ソートの完璧な 2 つのデータ・セットをディスク上に持つこととなります。

このアルゴリズムの全体実行時間は約 $N \cdot \log(N)$ です。ここでの N はレコード数です。メモリー内でのソートはディスクからの併合よりも高速なので、ソートにメモリーを多く割り当てるほど、実際の実行時間を減らすことができます。このアルゴリズムは、IBM SPSS Modeler Server の構成オプションである「メモリー使用量の乗数」で制御される、物理 RAM の一部をアルゴリズム自体に割り当てます。ソートに使用されるメモリーを増やすには、物理 RAM を増やすか、この値を増やします。使用されるメモリー部分がプロセスの作業セットより大きくなり、メモリーの一部がディスクへページングされると、内部ソート・アルゴリズムのメモリーアクセス・パターンがランダムになり、過度のページングの原因となるので、パフォーマンスが減衰することを覚えておいてください。ソート・アルゴリズムはソート・ノード以外の複数のノードでも使用されますが、同じパフォーマンス・ルールが適用されます。

データ分割 : データ分割ノードは、ビンの境界を計算するために、レコードのビンへの割り当て前に入力データ・セット全体を読み込みます。データ・セットは、境界の計算中はキャッシュへ入れられ、その後、割り当て用に再スキャンされます。データ分割の手法が「固定幅」または「平均 + 標準偏差」の場合、データ・セットはディスクへ直接キャッシュされます。これらの手法には直線的に増加する実行時間が必要であり、また、データ・セット全体を格納するのに十分なディスク領域が必要です。データ分割手法が「ランク」または「分位」の場合、データ・セットは先に説明したソート・アルゴリズムを使用してソートされ、ソート済みのデータ・セットがキャッシュとして使用されます。これらの手法のソートの実行時間は $M \cdot N \cdot \log(N)$ です。ここでの M はビンに分割されるフィールド数、 N はレコード数です。また、実行にはデータ・セットのサイズの 2 倍のディスク領域が必要です。

生成されたビンに基づいてフィールド作成ノードを生成しておくこと、以後のパスのパフォーマンスが向上します。フィールド作成の速度は、データ分割に比べてかなり速くなります。

キーによる結合: 結合の手法がキー (データベース結合と同じ) であるレコード結合ノードでは、各入力データ・セットがキー・フィールドでソートされます。プロシーチャーのこの部分の実行時間は $M \cdot N \cdot \log(N)$ です。 M は入力の数、 N は最大の入力に含まれているレコードの数です。すべての入力データ・セットのほかに、最大のデータ・セットの 2 番目のコピーを格納できるだけの十分なディスク・スペースが必要で

す。結合自体の実行時間そのものは、出力データ・セットのサイズに比例し、キーが一致する頻度に左右されます。最悪の場合は出力が入力のデカルト積であり、実行時間が NM に近づくことがあります。これは稀な例ですが、ほとんどの結合では一致するキーが少ない場合があります。1 つのデータ・セットがほかにくらべて比較的大きい場合、または、入力データがすでにキー・フィールドでソートされている場合は、「最適化」タブを使用して、このノードのパフォーマンスを改善できます。

集計: 「連続キー」オプションが設定されていないと、このノードでは、集計済みの出力を作成する前に入力データ・セット全体を読みます (ただし、格納しません)。さらに極端な状況、つまり、集計されたデータのサイズが制限 (IBM SPSS Modeler Server 構成オプションの「メモリー使用量の乗数」で決定される) に達すると、データセットの残りの部分は、「連続キー」のオプションが設定されたようにしてソートされ、処理されます。このオプションが設定されていると、集計済みの出力レコードが入力データが読み込まれるごとに作成されるので、データは格納されません。

重複レコード: 重複レコード・ノードでは、入力データセット内の一意のキー・フィールドのすべてが格納されます。すべてのフィールドがキー・フィールドですべてのレコードが一意であるような場合、データセット全体が格納されます。デフォルトでは、重複レコード・ノードはキー・フィールドのデータをソートし、各グループの最初の重複レコードを選択 (または破棄) します。重複レコード・キーが少ない小さなデータセットの場合、または事前にソートされているデータセットの場合、処理のスピードおよび効率を改善するオプションを選択できます。

タイプ: このキャッシュは、下流の処理で使用されます。キャッシュでは、データ・セット全体を格納しても処理速度を上げるために、十分なディスク領域が必要です。

評価: 評価ノードでは、分位を計算するために入力データをソートする必要があります。スコアと結果のレコード順が各ケースで異なるので、評価済みの各モデルに対してソートが繰り返されます。実行時間は $M \times N \times \log(N)$ です。ここでの M はモデル数、 N はレコード数です。

パフォーマンス :モデル作成ノード

ニューラルおよび **Kohonen** : ニューラル・ネットワークの学習アルゴリズム (Kohonen アルゴリズムも含む) により、学習データに多くの経路が作成されます。データは、一定の限度までメモリーに格納され、超過分はディスクへ書き出されます。アクセス方法がランダムで、その結果ディスク活動が過度になるため、ディスクから学習データへアクセスすることはコスト的に高くなります。ノードのダイアログ・ボックスの「モデル」タブで「速度の最適化」オプションを選択することですべてのデータがメモリーに格納されるようにし、これらのアルゴリズムのディスク領域の使用を無効化することができます。データを格納するのに必要なメモリー量がサーバー・プロセスの作業セットより大きい場合、メモリーの一部がディスクへページングされて、それに従ってパフォーマンスが落ちることになります。

「メモリー優先で最適化」が有効にされると、IBM SPSS Modeler Server 構成オプションの「モデル作成メモリー制限の割合」の値に従って、物理 RAM の指定されたパーセンテージがアルゴリズムに割り当てられます。学習ニューラル・ネットワークにさらに多くのメモリーを使用するには、RAM をさらに増やすか、このオプションの値を増やすかですが、設定値が高すぎると、ページングの原因になります。

ニューラル・ネットワーク・アルゴリズムの実行時間は、精度のレベルに左右されます。実行時間は、ノードのダイアログ・ボックスで停止条件を設定することで制御できます。

K-Means : K-Means クラスタリング・アルゴリズムには、ニューラル・ネットワーク・アルゴリズムと同様の、メモリー使用量を制御するオプションがあります。データへのアクセスがシーケンシャルであるため、ディスクに格納されたデータへのパフォーマンスのほうが良好です。

パフォーマンス :CLEM 式

データ・ストリーム内を参照する CLEM シーケンス関数 (「@ 関数」) では、最長の参照を可能にするのに十分なデータを格納する必要があります。参照の程度に制限がない操作では、フィールドのすべての値が格納される必要があります。制限がない操作とは、例えば @OFFSET(Sales, Month) のように、オフセット値がリテラルの整数ではない操作です。オフセット値がフィールド名 *Month* なので、その値は実行されるまで不明です。正確な結果を保証するために、サーバーは *Sales* フィールドのすべての値を保存しておく必要があります。上限値がわかっている場合は、例えば @OFFSET(Sales, Month, 12) のように、上限値を追加の引数として指定する必要があります。この指定により、サーバーは *Sales* の 12 の最新値以降を保存しないようになります。制限付きか付きでないかにかかわらず、シーケンス関数によって、ほとんど常に SQL 生成が抑制されます。

第 15 章 IBM SPSS Modeler のアクセシビリティ

IBM SPSS Modeler のアクセシビリティの概要

IBM SPSS Modeler では、視覚障害やその他の機能障害のあるユーザーに対する特定のサポートの他に、すべてのユーザーにアクセシビリティ・サポートが提供されます。このセクションでは、画面読み上げソフトウェアやキーボード・ショートカットなどのアクセシビリティの概要、およびこれらの機能を使った作業方法を説明していきます。

アクセシビリティサポートの種類

視覚障害をお持ちの方や、キーボードを使った操作しかできないような方でも利用できるように、さまざまな代替手段が用意されています。例えば、ストリームの構築、オプションの設定、および出力の読み込みなどの作業を、すべてマウスを使わないで行うことができます。利用できるキーボード・ショートカットについては、後述します。また、IBM SPSS Modeler は、JAWS for Windows のような画面読み上げソフトウェアもサポートしています。さらに、コントラストを強くするために、カラー・スキーマを最適化することもできます。これらのサポート内容の詳細は、以降のトピックで説明していきます。

視力の弱いユーザーのためのアクセシビリティ

ソフトウェアを使いやすくするために、IBM SPSS Modeler にはさまざまなプロパティが用意されています。

表示オプション

グラフの表示色を選択することができます。また、ソフトウェア自体に対して特定の Windows 設定を利用することもできます。これは、コントラストを強くするような場合に役立ちます。

1. 表示オプションを設定するには、「ツール」メニューの「ユーザー オプション」をクリックします。
2. 「表示」タブをクリックします。このタブには、ソフトウェアの表示色、グラフの色、およびノードのフォント・サイズなどを設定できます。

注: 画面読み上げソフトウェアではグラフを読み取ることができないため、視力に障害のあるユーザーはグラフをご利用になれません。

サウンドを通知に使用

サウンドをオンまたはオフにして、ソフトウェアの特定の操作をサウンドで知らせるかどうかを制御できます。例えば、ノードの作成、削除、または新規出力やモデルの生成時に、サウンドを鳴らすことができます。

1. 通知オプションを設定するには、「ツール」メニューの「ユーザー オプション」をクリックします。
2. 「通知」タブをクリックします。

新規ウィンドウの自動起動

「ユーザー オプション」ダイアログ・ボックスの「通知」タブでは、テーブルやグラフなどの新しく生成された出力を、新しいウィンドウに表示するかどうかを指定することもできます。このオプションを無効にして、必要に応じて出力ウィンドウを表示することもできます。

1. これらのオプションを設定するには、「ツール」メニューの「ユーザー オプション」をクリックします。
2. 「通知」タブをクリックします。
3. ダイアログ・ボックスで、「視覚通知」のリストから「新規出力」を選択します。
4. 「ウィンドウを開く」で、「開かない」を選択します。

ノード・サイズ

ノードは、標準のサイズまたは小さいサイズで表示することができます。必要に応じて、大きさを調整してください。

1. ノード・サイズ・オプションを設定するには、「ファイル」メニューの「ストリームのプロパティ」をクリックします。
2. 「レイアウト」タブをクリックします。
3. 「アイコン サイズ」リストで、「標準」を選択します。

視覚障害ユーザーのためのアクセシビリティ

視覚障害者向けのサポート機能は、JAWS for Windows のような画面読み上げソフトウェアを使用しているかどうか大きく依存しています。IBM SPSS Modeler と一緒に使用する画面読み上げソフトウェアを最大限に活用できるように、多くの設定を指定できます。

表示オプション

スクリーン・リーダー (画面読み上げソフトウェア) は、画面のコントラストが強い方が、より正確に動作する傾向があります。Windows をすでに高コントラストの設定にしている場合は、それらの Windows 設定をソフトウェアで使用することもできます。

1. 表示オプションを設定するには、「ツール」メニューの「ユーザー オプション」をクリックします。
2. 「表示」タブをクリックします。

注: 画面読み上げソフトウェアではグラフを読み取ることができないため、視覚障害ユーザーはグラフをご利用になれません。

サウンドを通知に使用

サウンドをオンまたはオフにして、ソフトウェアの特定の操作をサウンドで知らせることができます。例えば、ノードの作成、削除、または新規出力やモデルの生成時に、サウンドを鳴らすことができます。

1. 通知オプションを設定するには、「ツール」メニューの「ユーザー オプション」をクリックします。
2. 「通知」タブをクリックします。

新規ウィンドウの自動起動

「ユーザー オプション」ダイアログ・ボックスの「通知」タブでは、新しく生成された出力を、新しいウィンドウに表示するかどうかを指定することもできます。このオプションを無効にして、必要に応じて出力ウィンドウを表示することもできます。

1. これらのオプションを設定するには、「ツール」メニューの「ユーザー オプション」をクリックします。
2. 「通知」タブをクリックします。
3. ダイアログ・ボックスで、「視覚通知」のリストから「新規出力」を選択します。
4. 「ウィンドウを開く」で、「開かない」を選択します。

キーボード・アクセシビリティ

製品の機能にキーボードを使用してアクセスできます。基本的には、Alt キーと他の適切なキーを同時に押してメニュー項目を選択したり (例: Alt + F キーで「ファイル」メニューを選択)、Tab キーを使用してダイアログ・ボックス中のコントロール間を移動することができます。しかし、それ以外に製品のメイン・ウィンドウのそれぞれに関連した固有な問題があり、ダイアログ・ボックス内での操作についての便利なヒントもあります。

このセクションでは、ストリームを開く、ノード・ダイアログ・ボックスの使用、および出力作業などを行うために利用できる、キーボードによるアクセシビリティを説明していきます。また、効率的な移動を行うために、キーボード・ショートカットの一覧も記載されています。

メイン・ウィンドウ内を移動するためのショートカット

大半の作業は IBM SPSS Modeler のメイン・ウィンドウで行われます。メイン領域はストリーム領域と呼ばれ、データ・ストリームの作成と実行に用いられます。ウィンドウの下部にあるノード・パレットには、利用できるノードが用意されています。パレットは、データ・マイニング操作の種類に応じたタブにわかれており、それぞれに適切なノードが配置されています。例えば、IBM SPSS Modeler にデータを取り込むために使われるノードは、「入力」タブに配置されています。また、フィールドの作成、フィルタリング、およびデータ型の設定に使われるノードは、「フィールド設定」タブに配置されています。

ウィンドウの右側には、ストリーム、出力、およびプロジェクトを管理するためのさまざまなツールが用意されています。ウィンドウ右側の上半分には、ストリーム、出力、および生成されたモデルを管理するために用いられるマネージャがあります。このマネージャには、3 つのタブがあります。ストリーム、出力、生成されたモデルなどのオブジェクトにアクセスするには、適切なタブを選択して、リストからオブジェクトを選択します。メイン・ウィンドウ右側の下半分には、プロジェクトの作業を行うためのプロジェクト・パネルがあります。このエリアには、プロジェクトに 2 の異なるビューが存在することを反映して 2 つのタブがあります。クラス・ビューは、プロジェクト・オブジェクトをタイプごとにソートしますが、**CRISP-DM** ビューは、データ準備およびモデル作成などの関連するデータ・マイニング・フェーズごとにオブジェクトをソートします。これらの IBM SPSS Modeler ウィンドウの各部の詳細は、ヘルプやマニュアルを参照してください。

IBM SPSS Modeler のメイン・ウィンドウで、およびストリームの構築に使用できるショートカットの一覧を次の表に示します。ダイアログ・ボックスや出力で用いられるショートカットは後述します。これらのショートカット・キーは、メイン・ウィンドウでしか利用できないことに注意してください。

表 38. メイン・ウィンドウのショートカット

ショートカット・キー	関数
Ctrl+F5	フォーカスをノード・パレットに移動します。

表 38. メイン・ウィンドウのショートカット (続き)

ショートカット・キー	関数
Ctrl+F6	フォーカスをストリーム領域に移動します。
Ctrl+F7	フォーカスをマネージャー・パネルに移動します。
Ctrl+F8	フォーカスをプロジェクト・パネルに移動します。

表 39. ノードとストリームのショートカット

ショートカット・キー	関数
Ctrl+N	ストリーム領域に新しい空のストリームを作成します。
Ctrl+O	既存のストリームを選択して開くことのできる「開く」ダイアログボックスを表示します。
Ctrl+数字キー	フォーカスをウィンドウまたはペインの対応するタブに移動します。例えば、タブのある領域やウィンドウで、Ctrl-1 を押すと左から 1 番目のタブが、Ctrl-2 を押すと 2 番目のタブが選択されます。
Ctrl+下矢印キー	ノード・パレットでフォーカスをパレット・タブからそのタブの最初のノードまで移動するのに使用します。
Ctrl+上矢印キー	ノード・パレットでフォーカスをノードからそのパレット・タブまで移動するのに使用します。
Enter	ノード・パレット (「生成されたモデル」パレットの調整済みモデルも含めて) でノードが選択されている場合に、このキーを押すとノードがストリーム領域に追加されます。キャンバスでノードがすでに選択されている場合に Enter を押すと、そのノードのダイアログ・ボックスが開きます。
Ctrl+Enter	パレットでノードが選択されている場合に、ノードをストリーム領域に追加します。ただし、追加したノードは選択されず、フォーカスはノード・パレットのまま変わりません。
Alt+Enter	パレットでノードが選択されている場合に、ノードをストリーム領域に追加して選択します。ただしフォーカスはノード・パレットのまま変わりません。
Shift+スペースキー	ノードまたはコンポーネントがパレットにフォーカスを置く場合、ノードまたはコンポーネントの選択と非選択を切り替えます。 ほかのノードまたはコンポーネントも選択されている場合、これらの選択は解除されます。
Ctrl+Shift+スペースキー	ノードまたはコメントがストリームにフォーカスを置く場合、またはノードまたはコメントがパレットにフォーカスを置く場合、ノードまたはコメントの選択と非選択を切り替えます。 そのほかの選択されたノードまたはコメントに影響はありません。
左右の矢印キー	ストリーム領域にフォーカスがある場合、ストリーム全体を画面上で平行に移動します。パレット・タブにフォーカスがある場合、タブの間を順次移動します。パレット・ノードにフォーカスがある場合、パレットのノード間を移動します。
上下の矢印キー	ストリーム領域にフォーカスがある場合、ストリーム全体を画面上で平行に移動します。パレット・ノードにフォーカスがある場合、パレットのノード間を移動します。サブパレットにフォーカスがある場合、このパレット・タブに対して他のサブパレット間を移動します。
Alt+左右の矢印キー	ストリーム領域上で選択されているノードまたはコメントを、矢印キーの方向に平行に移動します。

表 39. ノードとストリームのショートカット (続き)

ショートカット・キー	関数
Alt+上下の矢印キー	ストリーム領域上で選択されているノードまたはコメントを、矢印キーの方向で垂直方向に移動します。
Ctrl+A	ストリーム内のすべてのノードを選択します。
Ctrl+Q	ノードにフォーカスがある場合に、そのノードと下流にあるすべてのノードを選択し、上流にあるすべてのノードを非選択にします。
Ctrl+W	ノードにフォーカスがある場合に、そのノードと下流にあるすべての選択したノードを非選択にします。
Ctrl+Alt+D	選択したノードを複製します。
Ctrl+Alt+L	ストリームでモデル ナゲットが選択されている場合、「挿入」ダイアログを開いて .nod ファイルからストリームへ保存したモデルをロードできるようにします。
Ctrl+Alt+R	選択したノードの「注釈」タブを表示し、ノードの名前が変更できるようになります。
Ctrl+Alt+U	ユーザー入力ソース・ノードを生成します。
Ctrl+Alt+C	ノードのキャッシュのオンとオフを切り替えます。
Ctrl+Alt+F	ノードのキャッシュを消去します。
タブ	ストリーム領域で、現在のストリームのすべての入力ノードおよびコメントを順次移動します。ノード・パレットで、パレットのノード間を移動します。選択したサブパレットで、サブパレットの最初のノードに移動します。
Shift+Tab	Tab と同じ操作を実行しますが、順序が逆になります。
Ctrl+Tab	マネージャー・パネルまたはプロジェクト・パネルにフォーカスを置き、ストリーム領域にフォーカスを移動します。ノード・パレットにフォーカスを置き、ノードとパレット・タブ間でフォーカスを移動します。
アルファベット・キー	現在のストリームのノードにフォーカスがある状態から、名前が押したキーの文字で始まる次のノードにフォーカスを移動します。
[F1] キー	ヘルプのフォーカスに関連したトピックを開きます。
[F2] キー	ストリーム領域上で選択されているノードの接続処理を開始します。Tab キーを使用して接続先のノードを選択し、Shift + スペース・バーを押すと接続が行われます。
[F3] キー	ストリーム領域上で選択されているノードに対するすべての接続を削除します。
[F6] キー	マネージャー・パネル、プロジェクト・パネルおよびノード・パレット間でフォーカスを移動します。
F10	「ファイル」メニューを表示します。
Shift+F10	ノードまたはストリームのポップアップ・メニューを表示
Delete	キャンバスから選択したノードを削除します。
Esc	ポップアップ・メニューまたはダイアログ・ボックスを閉じます。
Ctrl+Alt+X	スーパーノードを展開します。
Ctrl+Alt+Z	スーパーノードをズーム・インします。
Ctrl+Alt+Shift+Z	スーパーノードからズーム・アウトします。
Ctrl+E	ストリーム領域上にフォーカスを置きながら、現在のストリームを実行します。

ほかにも、コピーの場合は **Ctrl-C** キーなど、IBM SPSS Modeler ではさまざまな標準のショートカットキーを利用することができます。詳しくは、トピック 22 ページの『ショートカット・キーの使用』を参照してください。

ダイアログ・ボックスや表のショートカット・キー

ダイアログ・ボックス、表、およびダイアログ・ボックス中の表に関する作業を行う際に役立つさまざまなショートカットキーや画像読み上げソフトウェア用キーが用意されています。これらのキーの一覧を次に示します。

表 40. ダイアログ・ボックスと *Clem* 式ビルダーのショートカット :

ショートカット・キー	関数
Alt+4	開かれているすべてのダイアログ・ボックスまたは出力ウィンドウを閉じるために用いられます。出力は、マネージャー・パネルの「出力」タブから取得することができます。
[Ctrl] + [End]	<i>Clem</i> 式ビルダー内のコントロールにフォーカスがある場合に、挿入ポイントを式の最後に移動します。
Ctrl+1	<i>Clem</i> 式ビルダーで、フォーカスを <i>CLEM</i> 式編集コントロールに移動します。
Ctrl+2	<i>Clem</i> 式ビルダーで、フォーカスを関数リストに移動します。
Ctrl+3	<i>Clem</i> 式ビルダーで、フォーカスをフィールド・リストに移動します。

テーブルのショートカット

テーブルのショートカットは、出力テーブルや、データ型、フィルター、およびレコード結合などのノードのダイアログ・ボックス中にあるテーブル・コントロールで用いられます。一般的には、テーブル中のセル間の移動には **Tab** キーを、テーブル・コントロールを終了するには **Ctrl-Tab** キーを使用します。注：場合によっては、画面読み上げソフトウェアがセルの内容をすぐに読み上げ始めないこともあります。矢印キーを 1 回または 2 回押すと、ソフトウェアがリセットされて、読み上げが開始されます。

表 41. テーブルのショートカット :

ショートカット・キー	関数
Ctrl+W	テーブルで選択されている行 (row) の簡単な説明を読み上げます。例えば、「選択されている行 2 の値は、性別、フラグ型、男/女…」のように読み上げます。
Ctrl+Alt+W	テーブルで選択されている行 (row) の長い説明を読み上げます。例えば、「選択されている行 2 の値は、フィールド = 性別、データ型 = フラグ型、性別 = 男/女…」のように読み上げます。
Ctrl+D	テーブルで選択されている領域の、簡単な説明 (description) を読み上げます。例えば、「列 6 の 1 行目が選択されています。」のように読み上げます。
Ctrl+Alt+D	テーブルで選択されている領域の、長い説明 (description) を提供します。例えば、「1 行 6 列が選択されています。選択されている列は、フィールド、データ型、欠損値です。選択された行は 1 です。」のようになります。
Ctrl+T	テーブルで選択されている領域の、簡単な説明を 'n 供します。例えば、「フィールド、データ型、欠損値」のようになります。

表 41. テーブルのショートカット (続き):

ショートカット・キー	関数
Ctrl+Alt+T	テーブルで選択されている領域の、長い説明を提供します。例えば、「選択されている列は、フィールド、データ型、欠損値です。」のようになります。
Ctrl+R	テーブルで、テーブル中のレコード数 (records) を提供します。
Ctrl+Alt+R	テーブルで、テーブル中のレコード数 (records)、および列名を提供します。
Ctrl+I	テーブルで、フォーカスがあるセルのセル情報 (information) または内容を読み上げます。
Ctrl+Alt+I	テーブルで、フォーカスがあるセルのセル情報 (information) の長い説明 (列名とセルの内容) を読み上げます。
Ctrl+G	テーブルで、簡単な一般選択情報 (general) を提供します。
Ctrl+Alt+G	テーブルで、長い一般選択情報 (general) を提供します。
Ctrl+Q	テーブルで、テーブル・セルを素早く (quick) 切り替えられるようにします。Ctrl-Q を押すと、テーブル内を矢印キーを使用して移動すれば、その詳細な説明が「性別 = 女性」のように読み上げられます。もう一度 Ctrl-Q を選択すると、簡単な説明に切り替わります (セルの内容)。
F8	テーブルで、フォーカスがテーブルにある場合に列見出しにフォーカスを設定します。
スペース・バー	テーブルで、フォーカスが列見出しにある場合に列をソートできるようにします。

コメントのショートカット

画面上のコメントを処理している場合、次のショートカットを使用できます。

表 42. コメントのショートカット

ショートカット・キー	関数
Alt+C	コメントの表示/非表示機能を切り替えます。
Alt+M	コメントが現在表示されている場合は新しいコメントを挿入し、非表示になっている場合はコメントを表示します。
タブ	ストリーム領域で、現在のストリームのすべての入力ノードおよびコメントを順次移動します。
Enter	コメントにフォーカスがある場合、編集の開始を指示します。
Alt+Enter または Ctrl+Tab	編集を終了し、編集の変更を保存します。
Esc	編集を中止します。編集中心に行われた変更は失われます。
Alt+Shift+上矢印キー	グリッドによる配置の設定がオン (またはオフ) の場合、1 グリッド・セル (または 1 ピクセル) ごとにテキスト領域の高さを縮小します。
Alt+Shift+下矢印キー	グリッドによる配置の設定がオン (またはオフ) の場合、1 グリッド・セル (または 1 ピクセル) ごとにテキスト領域の高さを拡大します。
Alt+Shift+左矢印キー	グリッドによる配置の設定がオン (またはオフ) の場合、1 グリッド・セル (または 1 ピクセル) ごとにテキスト領域の幅を縮小します。
Alt+Shift+右矢印キー	グリッドによる配置の設定がオン (またはオフ) の場合、1 グリッド・セル (または 1 ピクセル) ごとにテキスト領域の幅を拡大します。

クラスター・ビューアおよびモデル・ビューアのショートカット

ショートカット・キーは、クラスター・ビューアおよびモデル・ビューア・ウィンドウを移動するために使用できます。

表 43. 汎用ショートカット - クラスター・ビューアおよびモデル・ビューア

ショートカット・キー	関数
タブ	フォーカスを次の画集コントロールに移動します。
Shift+Tab	フォーカスを前の画面コントロールに移動します。
[↓]	ドロップダウン・リストにフォーカスがある場合、リストを開くか、リストの次の項目に移動します。 メニューにフォーカスがある場合、メニューの次の項目に移動します。 サムネイル・グラフにフォーカスがある場合、セット内の次のサムネイル (または最後のサムネイルにフォーカスがある場合は最初のサムネイル) に移動します。
[↑]	ドロップダウン・リストにフォーカスがある場合、リストの前の項目に移動します。 メニューにフォーカスがある場合、メニューの前の項目に移動します。 サムネイル・グラフにフォーカスがある場合、セット内の前のサムネイル (または最初のサムネイルにフォーカスがある場合は最後のサムネイル) に移動します。
Enter	使用中のドロップダウン・リストを閉じるか、または使用中のメニューで選択を行います。
[F6] キー	ウィンドウの左側および右側のパネルでフォーカスを切り替えます。
左右の矢印キー	タブにフォーカスがある場合、メニューの前後のタブに移動します。 メニューにフォーカスがある場合、前後のメニューに移動します。
Alt+文字	その名前が下線の付いた文字を含むボタンまたはメニューを選択します。
Esc	使用中のメニューまたはドロップダウン・リストを閉じます。

クラスター・ビューアのみ

クラスター・ビューアには、クラスター特徴グリッドを含むクラスター・ビューがあります。

モデルの要約ビューの代わりにクラスター・ビューを選択する手順は、次のとおりです。

1. 「表示」 ボタンが選択されるまで、Tab キーを繰り返し押します。
2. 下向き矢印を 2 回押して、「クラスター」 を選択します。

ここから、グリッド内の個別のセルを選択できます。

3. 視覚化ツールバーの最後のアイコンが選択されるまで、Tab キーを繰り返し押します。



図 17. 「視覚化ツリーを表示」 アイコン

4. Tab キーをもう一度押し、その後スペース、そして矢印キーを押します。

次のキーボード・ショートカットを使用できます。

表 44. クラスター・ビューアのショートカット

ショートカット・キー	関数
矢印キー	グリッド内の各セル間でフォーカスを移動します。右側のパネルのセルの分布表示が、フォーカスが移動するごとに変わります。
Ctrl+コンマ	セルにフォーカスがあるグリッドの列全体を選択または選択解除します。 列を選択部分に追加するには、矢印キーを使用して、その列のセルに移動し、もう一度 Ctrl キーを押しながらコンマを押します。
タブ	フォーカスをグリッドから次の画面コントロールに移動します。
Shift+Tab	フォーカスをグリッドから前の画面コントロールに移動します。
[F2] キー	編集モード (ラベルおよびセル名のセルのみ) に入ります。
Enter	編集した変更を保存し、編集モード (ラベルおよびセル名のセルのみ) を終了します。
Esc	変更を保存せずに編集モード (ラベルおよびセル名のセルのみ) を終了します。

ショートカット・キーの例 :ストリームの構築

キーボードや画面読み上げソフトウェアを使用した、ユーザーのストリームの構築作業を理解するために、マウスを使わないでストリームを構築する例を取り上げます。この例では、可変長ファイルノード、フィールド作成ノード、およびヒストグラム・ノードがあるストリームを、次の手順にしたがって作成していきます。

1. **IBM SPSS Modeler**を始動します。IBM SPSS Modeler が最初に始動される際、フォーカスはノード・パレットの「お気に入り」タブにあります。
2. **Ctrl+下矢印キー**: フォーカスをタブ自体から、タブの内容へと移動します。
3. 右矢印キー: フォーカスを可変長ファイル・ノードに移動します。
4. スペース・バー : 可変長ファイル・ノードを選択します。
5. **Ctrl+Enter**: 可変長ファイル・ノードをストリーム領域に追加します。このキーの組み合わせでは、可変長ファイル・ノードが選択されたままになるため、次のノードを追加すると可変長ノードに接続されます。
6. **Tab** : フォーカスをノード・パレットに移動します。
7. 右矢印キーを **4** 回: フィールド作成ノードを移動します。
8. スペース・バー : フィールド作成ノードを選択します。
9. **Alt+Enter**: フィールド作成ノードをストリーム領域に追加して、フィールド作成ノードを選択します。これで、このノードは次に追加されるノードと接続されることとなります。
10. **Tab** : フォーカスをノード・パレットに移動します。
11. 右矢印キーを **5** 回: フォーカスがパレット中のヒストグラム・ノードに移動します。
12. スペース・バー : ヒストグラム・ノードを選択します。
13. **Enter** : ノードをストリームに追加し、フォーカスをストリーム領域に移動します。

次の例に進むか、後で次の例を試す場合はストリームを保存します。

ショートカット・キーの例 : ノードの編集

この例では、前の例で作成したストリームを使用します。このストリームは、可変長ファイル・ノード、フィールド作成ノード、およびヒストグラム・ノードから成り立っています。この例では、ストリーム中の 3 番目にあるヒストグラム・ノードにフォーカスがある所から始まります。

1. **Ctrl+** 左矢印キーを 2 回 : フォーカスを可変長ファイル・ノードに移動します。
2. **Enter** : 「可変長ファイル」ダイアログ・ボックスを開きます。「ファイル」フィールドに Tab キーで移動し、テキスト・ファイルのパスおよび名前を入力して、そのファイルを選択します。Ctrl+Tab キーでダイアログの下部分に移動し、Tab キーで「OK」ボタンに移動して Enter キーを押すとダイアログ・ボックスが閉じます。
3. **Ctrl+**右矢印キー: フォーカスを 2 番目のフィールド作成ノードに移動します。
4. **Enter** : 「フィールド作成ノード」ダイアログ・ボックスが開きます。Tab キーを何回か押してフィールドを選択し、フィールド作成条件を指定します。Ctrl+Tab キーで「OK」ボタンまで移動し、Enter キーを押すとダイアログ・ボックスが閉じます。
5. **Ctrl+**右矢印キー: フォーカスを 3 番目のヒストグラム・ノードに移動します。
6. **Enter** : 「ヒストグラム・ノード」ダイアログ・ボックスが表示されます。Tab キーを何回か押してフィールドを選択し、グラフ・オプションを指定します。ドロップダウン・リストでは、下矢印キーを押してリストを開きリスト項目を強調表示してから、Enter キーを押してリスト項目を選択します。Tab キーで「OK」ボタンまで移動し、Enter キーを押すとダイアログ・ボックスが閉じます。

この時点で、他のノードを追加したり、現在のストリームを実行することができます。ストリームを構築するときは、忘れずに次のヒントに従ってください。

- 手作業でノードを接続する場合、F2 キーを使用して接続の始点を作成し、Tab キーを同時に使用して終点まで移動し、Shift キーとスペース・バーを使用して接続を完了します。
- ストリーム領域で選択されているノードのすべての接続を解除するには、F3 キーを使用します。
- ストリームを作成した後、Ctrl + E キーを押すと現在のストリームを実行します。

ショートカット・キーの完全な一覧が用意されています。詳しくは、トピック 217 ページの『メイン・ウィンドウ内を移動するためのショートカット』を参照してください。

画面読み上げソフトウェアの使用

多数の画面読み上げソフトウェアが市場で販売されています。IBM SPSS Modeler は IBM SPSS Modeler と一緒にインストールされる Java Access Bridge を使用して Windows JAWS に対応するように設定されています。JAWS がインストール済みであれば、IBM SPSS Modeler を起動する前に JAWS を起動するだけで利用できます。

注: SPSS Modeler で JAWS を実行するには、容量が少なくとも 6GB あることを推奨します。

IBM SPSS Modeler 独自のデータ・マイニング・プロセスのグラフィック表示のため、グラフは視覚的に使用するために最適化されています。ただし、画面読み上げソフトウェアを使用して「文字的に」表示できる出力やモデルを理解しながら、作業を行うことができます。

注: 64 ビットのクライアント・コンピューターでは、支援技術機能が機能しません。これは Java Access Bridge が 64 ビットの操作向けに設計されていないためです。

IBM SPSS Modeler デクシオナリー・ファイルの使用

IBM SPSS Modeler ディクショナリー・ファイル (*Awt.JDF*) は、JAWS に含めて利用することができません。このファイルを使用するには

1. IBM SPSS Modeler のインストール・ディレクトリー中の */accessibility* サブディレクトリーに移動して、ディクショナリー・ファイル (*Awt.JDF*) をコピーします。
2. コピーしたファイルを、JAWS スクリプトのあるディレクトリーに貼り付けます。

他の JAVA アプリケーションを稼働させている場合は、*Awt.JDF* という名前のファイルがすでに存在していることもあります。この場合、手作業でディクショナリー・ファイルを編集しないと、このファイルを使用できないこともあります。

HTML 出力での画面読み上げソフトウェアの使用

IBM SPSS Modeler で画面読み上げソフトウェアを使用して HTML 出力を表示する場合、いくつかの問題が発生する可能性があります。次のように、多数の出力タイプが影響を受けます。

- 線形回帰ノード、ロジスティック回帰ノード、および因子分析ノードの「詳細」タブに表示する出力。
- レポート・ノードの出力

これらの各ウィンドウまたはダイアログ・ボックスには、出力をデフォルトのブラウザに表示するためのツールがツールバーに用意されています。このデフォルトのブラウザは、標準の画面読み上げソフトウェアをサポートしています。ブラウザに表示した出力情報を、画面読み上げソフトウェアを使用して調査します。

「インタラクティブ・ツリー」ウィンドウのアクセス

「インタラクティブ・ツリー」ウィンドウ内でのディシジョン・ツリーの標準的な表示方法は、画面読み上げソフトウェアを使用する際に問題になることがあります。アクセス可能なバージョンにアクセスするには、インタラクティブ・ツリーのメニューから次のメニュー項目を選択してください。

「表示」 > 「アクセス可能なウィンドウ」

標準のツリー・マップと似ていますが、JAWS が正しく読み上げることができるビューを表示します。標準的な矢印キーを使用すると、上下左右に移動できます。アクセス可能なウィンドウ内を移動すると、それに応じて「インタラクティブ・ツリー」ウィンドウ内のフォーカスも移動します。選択を変更するにはスペース・バーを使用します。また、現在の選択を拡張するには、Ctrl+スペースを使用します。

使用上のヒント

IBM SPSS Modeler 環境をより使いやすくするための、さまざまなコツがあります。IBM SPSS Modeler で作業を行う際の一般的なヒントを次に示します。

- 拡張テキスト・ボックスの終了：拡張テキスト・ボックスを終了するには Ctrl+Tab を使用します。注：Ctrl-Tab は、テーブル・コントロールを終了する際にも用いられます。
- 矢印キーよりも **Tab** キーを利用する：ダイアログ・ボックスのオプションを選択する場合、オプション・ボタン間の移動には Tab キーを使用します。ここでは、矢印キーを移動に使用することはできません。
- ドロップダウン・リスト：ダイアログ・ボックスのドロップダウン・リストでは、Esc キーまたはスペース・バーを使用して項目を選択し、リストを閉じることができます。他のコントロールにタブで移動しており、閉じていないドロップダウン・リストを閉じるには、Esc キーを使用することもできます。
- 実行ステータス：大きいデータベースでストリームを実行している場合、JAWS ではストリーム・ステータスの読み上げを遅くすることができます。ステータス・レポートを更新するには、Ctrl キーを定期的に押してください。

- ノード・パレットの使用：ノード・パレットのタブを初めて参照する場合、JAWS がノード名の代わりに「グループボックス」と読み上げることがあります。この場合、Ctrl-右矢印キーまたは Ctrl-左矢印キーを使用して、画面読み上げソフトウェアをリセットしてから、ノード名を読み上げてください。
- メニューの読み上げ：最初にメニューを開く際に、JAWS が最初のメニュー項目を読み上げないことがあります。この問題が発生している可能性がある場合は、下矢印キーを押した後に上矢印キーを押して、最初のメニュー項目を確認してください。
- カスケード・メニュー：JAWS は、カスケード・メニューのメイン（第 1 レベル）メニューは読み上げません。メニュー内を移動している際に、何も読み上げられないことがあった場合は、右矢印キーを押して子メニューの項目を読み上げてください。

また、IBM SPSS Modeler Text Analytics をインストールした場合、次のヒントでインタラクティブ・ワークベンチ・インターフェースがより使いやすくなります。

- ダイアログ・ボックスの入力：ダイアログ・ボックスに入力直後、Tab キーを押して最初のコントロールに焦点を当てることが必要な場合があります。
- 拡張テキスト・ボックスの終了：Ctrl+Tab キーを押して、拡張テキスト・ボックスを終了し、次のコントロールに移動します。注：Ctrl-Tab は、テーブル・コントロールを終了する際にも用いられます。
- 最初の文字を入力してツリー・リストの要素を検索：カテゴリ領域、抽出された結果の領域、またはライブラリー・ツリーで要素を検索する場合、領域にフォーカスがある際に要素の最初の文字を入力することができます。これにより、入力した文字で始まる要素の次の出現を選択します。
- ドロップダウン・リスト：ダイアログ・ボックスのドロップダウン・リストでは、スペース・バーを使用して項目を選択し、リストを閉じることができます。

使用上のその他のヒントは、以降のトピックで詳細に説明します。

他のソフトウェアとの相性

弊社開発チームが IBM SPSS Modeler で JAWS などの画面読み上げソフトウェアをテストした際に、SMS (Systems Management Server) を使用していると、JAWS による IBM SPSS Modeler などの Java ベース・アプリケーションの読み上げが正常に行われない可能性があることがわかりました。SMS を無効化するとこの問題を解決できます。SMS の詳細は、Microsoft の Web サイトを参照してください。

JAWS および Java

JAWS のバージョンによって、Java ベースのソフトウェア・アプリケーションのサポート・レベルも異なります。IBM SPSS Modeler は最近のバージョンのすべての JAWS で利用できますが、バージョンによっては Java ベースのシステムで使用する際に小さな問題が発生する場合があります。詳細は、JAWS for Windows の Web サイト <http://www.FreedomScientific.com> を参照してください。

IBM SPSS Modeler でのグラフの使用

ヒストグラム、評価グラフ、線グラフ、および散布図のような、視覚的に表示された情報を、画面読み上げソフトウェアで確認することは困難です。ただし、Web グラフや棒グラフの場合は、出力ウィンドウで利用できるテキスト要約情報を参照することができます。

第 16 章 Unicode のサポート

IBM SPSS Modeler 内での Unicode のサポート

IBM SPSS Modeler は、IBM SPSS Modeler と IBM SPSS Modeler Server の両方で、Unicode を完全にサポートしています。これにより、多言語データベースなどの Unicode をサポートする他のアプリケーションと、ロケール固有のコード化スキーム間で変換を行ったときに生じる情報の欠落なしに、データの交換が可能になります。

- IBM SPSS Modeler は、内部でデータを Unicode で保存しており、データベースに Unicode で保存されている多言語データを、情報の失うことなしに読み取ったり、書き込んだりすることができます。
- IBM SPSS Modeler は、UTF-8 エンコード・テキスト・ファイルを読み書きできます。テキスト・ファイルのインポートとエクスポートのデフォルトは、ロケールのエンコードになっていますが、それ以外に UTF-8 もサポートされています。この設定は、ファイル・インポートおよびエクスポート・ノードで指定できます。また、「ストリームのプロパティ」ダイアログ・ボックスで、デフォルトのエンコードを変更できます。詳しくは、トピック 43 ページの『ストリームの一般的なオプションの設定』を参照してください。
- ロケールのエンコードで保存されている Statistics、SAS、およびテキスト・データ・ファイルは、インポート時に UTF-8 に変換され、エクスポート時に元に戻されます。任意のファイルに書き込む場合、そのロケール文字セットに対応する文字が存在しない Unicode 文字がある場合、別の文字で置換され、また、警告が表示されます。この問題は、データが、Unicode をサポートするデータ・ソース (データベースまたは UTF-8 テキスト・ファイル) からインポートされ、さらに、そのデータにロケールが異なる文字や、複数のロケールまたは文字セットからの文字が含まれている場合についてのみ発生します。
- IBM SPSS Modeler Solution Publisher のイメージは、UTF-8 でエンコードされており、プラットフォームおよびロケール間で完全に移植可能です。

Unicode について

Unicode 規格の目標は、多言語テキストの一貫性のあるエンコード法を提供することです。これにより、国、地域、およびアプリケーションを越えた共有が容易になります。Unicode 規格は現在、バージョン 4.0.1 ですが、世界中で一般的に使用されているすべての文字セットのスーパーセットの文字セットを定義しており、それぞれの文字に一意的な名前とコードポイントを割り当てています。文字とそのコードポイントは、ISO-10646 により定義された Universal Character Set (UCS) のものと同一です。詳細は、Unicode ホームページを参照してください。

特記事項

本書は米国 IBM が提供する製品およびサービスについて作成したものです。この資料は、IBM から他の言語でも提供されている可能性があります。ただし、これを入手するには、本製品または当該言語版製品を所有している必要がある場合があります。

本書に記載の製品、サービス、または機能が日本においては提供されていない場合があります。日本で利用可能な製品、サービス、および機能については、日本 IBM の営業担当員にお尋ねください。本書で IBM 製品、プログラム、またはサービスに言及していても、その IBM 製品、プログラム、またはサービスのみが使用可能であることを意味するものではありません。これらに代えて、IBM の知的所有権を侵害することのない、機能的に同等の製品、プログラム、またはサービスを使用することができます。ただし、IBM 以外の製品とプログラムの操作またはサービスの評価および検証は、お客様の責任で行っていただきます。

IBM は、本書に記載されている内容に関して特許権 (特許出願中のものを含む) を保有している場合があります。本書の提供は、お客様にこれらの特許権について実施権を許諾することを意味するものではありません。実施権についてのお問い合わせは、書面にて下記宛先にお送りください。

〒103-8510

東京都中央区日本橋箱崎町19番21号

日本アイ・ビー・エム株式会社

法務・知的財産

知的財産権ライセンス渉外

IBM およびその直接または間接の子会社は、本書を特定物として現存するままの状態を提供し、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任を負わないものとします。国または地域によっては、法律の強行規定により、保証責任の制限が禁じられる場合、強行規定の制限を受けるものとします。

この情報には、技術的に不適切な記述や誤植を含む場合があります。本書は定期的に見直され、必要な変更は本書の次版に組み込まれます。IBM は予告なしに、随時、この文書に記載されている製品またはプログラムに対して、改良または変更を行うことがあります。

本書において IBM 以外の Web サイトに言及している場合がありますが、便宜のため記載しただけであり、決してそれらの Web サイトを推奨するものではありません。それらの Web サイトにある資料は、この IBM 製品の資料の一部ではありません。それらの Web サイトは、お客様の責任でご使用ください。

IBM は、お客様が提供するいかなる情報も、お客様に対してなら義務も負うことのない、自ら適切と信ずる方法で、使用もしくは配布することができるものとします。

本プログラムのライセンス保持者で、(i) 独自に作成したプログラムとその他のプログラム (本プログラムを含む) との間での情報交換、および (ii) 交換された情報の相互利用を可能にすることを目的として、本プログラムに関する情報を必要とする方は、下記に連絡してください。

IBM Director of Licensing

IBM Corporation

North Castle Drive, MD-NC119

Armonk, NY 10504-1785

US

本プログラムに関する上記の情報は、適切な使用条件の下で使用することができますが、有償の場合もあります。

本書で説明されているライセンス・プログラムまたはその他のライセンス資料は、IBM 所定のプログラム契約の契約条項、IBM プログラムのご使用条件、またはそれと同等の条項に基づいて、IBM より提供されます。

記載されている性能データとお客様事例は、例として示す目的でのみ提供されています。実際の結果は特定の構成や稼働条件によって異なります。

IBM 以外の製品に関する情報は、その製品の供給者、出版物、もしくはその他の公に利用可能なソースから入手したものです。IBM は、それらの製品のテストは行っておりません。したがって、他社製品に関する実行性、互換性、またはその他の要求については確認できません。IBM 以外の製品の性能に関する質問は、それらの製品の供給者にお願いします。

IBM の将来の方向性および指針に関する記述は、予告なく変更または撤回される場合があります。これらは目標および目的を提示するものにすぎません。

本書には、日常の業務処理で用いられるデータや報告書の例が含まれています。より具体性を与えるために、それらの例には、個人、企業、ブランド、あるいは製品などの名前が含まれている場合があります。これらの名称はすべて架空のものであり、類似する個人や企業が実在しているとしても、それは偶然にすぎません。

商標

IBM、IBM ロゴおよび ibm.com は、世界の多くの国で登録された International Business Machines Corporation の商標です。他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。現時点での IBM の商標リストについては、<http://www.ibm.com/legal/copytrade.shtml> をご覧ください。

Adobe、Adobe ロゴ、PostScript、PostScript ロゴは、Adobe Systems Incorporated の米国およびその他の国における登録商標または商標です。

インテル、Intel、Intel ロゴ、Intel Inside、Intel Inside ロゴ、Centrino、Intel Centrino ロゴ、Celeron、Xeon、Intel SpeedStep、Itanium、および Pentium は、Intel Corporation または子会社の米国およびその他の国における商標または登録商標です。

Linux は、Linus Torvalds の米国およびその他の国における登録商標です。

Microsoft、Windows、Windows NT および Windows ロゴは、Microsoft Corporation の米国およびその他の国における商標です。

UNIX は The Open Group の米国およびその他の国における登録商標です。

Java およびすべての Java 関連の商標およびロゴは Oracle やその関連会社の米国およびその他の国における商標または登録商標です。

製品資料に関するご使用条件

これらの資料は、以下のご使用条件に同意していただける場合に限りご使用いただけます。

適用条件

IBM Web サイトの「ご利用条件」に加えて、以下のご使用条件が適用されます。

個人的使用

これらの資料は、すべての著作権表示その他の所有権表示をしていただくことを条件に、非商業的な個人による使用目的に限り複製することができます。ただし、IBM の明示的な承諾をえずに、これらの資料またはその一部について、二次的著作物を作成したり、配布（頒布、送信を含む）または表示（上映を含む）することはできません。

商業的使用

これらの資料は、すべての著作権表示その他の所有権表示をしていただくことを条件に、お客様の企業内に限り、複製、配布、および表示することができます。ただし、IBM の明示的な承諾をえずにこれらの資料の二次的著作物を作成したり、お客様の企業外で資料またはその一部を複製、配布、または表示することはできません。

権利

ここで明示的に許可されているもの以外に、資料や資料内に含まれる情報、データ、ソフトウェア、またはその他の知的所有権に対するいかなる許可、ライセンス、または権利を明示的にも黙示的にも付与するものではありません。

資料の使用が IBM の利益を損なうと判断された場合や、上記の条件が適切に守られていないと判断された場合、IBM はいつでも自らの判断により、ここで与えた許可を撤回できるものとさせていただきます。

お客様がこの情報をダウンロード、輸出、または再輸出する際には、米国のすべての輸出入 関連法規を含む、すべての関連法規を遵守するものとします。

IBM は、これらの資料の内容についていかなる保証もしません。これらの資料は、特定物として現存するままの状態を提供され、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任なしで提供されます。

索引

日本語, 数字, 英字, 特殊文字の順に配列されています。なお, 濁音と半濁音は清音と同等に扱われています。

[ア行]

アイコン
オプションの設定 22, 48
アウトライン 73
省略 73
展開 73
ビューア 73
レベルの変更 73
アクセシビリティ 215, 226
例 223, 224
IBM SPSS Modeler の機能 215
IBM SPSS Modeler のヒント 225
値
データ検査から表示 123
CLEM 式への追加 123
値のラベル 91, 104
アウトライン・ペイン 104
ピボット・テーブル 104
アプリケーション 25
アプリケーションの例 3
位置合わせ 72, 103
出力 72, 103
一時ディレクトリー 13
位置調整 72, 103
出力 72, 103
色
設定 198
印刷 24, 85, 86, 87, 95, 97, 101
印刷プレビュー 86
グラフ 85
グラフ・サイズ 87
出力項目間のスペース 87
ストリーム 22, 39
層 85, 95, 97
テーブルのスケールリング 95, 97
テーブル・ブレイクの制御 101
テキスト出力 85
ピボット・テーブル 85
ページ番号 87
ヘッダーとフッター 86
インタラクティブ出力 77
「インタラクティブ・ツリー」ウィンドウ
アクセシビリティ 225

インポート
PMML 184, 185
エクスポート
ストリームの説明 57
PMML 184, 185
エラー・メッセージ 50
エンコード 43, 227
演算子
文字列の結合 135
CLEM 式内 119
演算子の優先順位 130
オートメーション 111
オブジェクト
プロパティ 191
オブジェクトのプロパティ、IBM SPSS
Collaboration and Deployment
Services Repository 175
オプション 103, 104, 195
構文 199
出力ラベル 104
ストリームのプロパティ 42, 43, 44,
45, 47, 48, 49, 50
全般 103
データ・エディターの記述統計量 104
ピボット・テーブル・ルック 103
ビューア 103
表示 198
ユーザー 196
IBM SPSS Modeler の 195
PMML 200
折り返し 95
折り返されたテキストの列幅の制御 95

[カ行]

カイ 2 乗分布
確率関数 140
帰帰 225
概要 127
IBM SPSS Modeler 9, 195
科学的表記法 103
表示書式 45
確率関数 140
カスタム・パレットの作成 205
サブパレットの作成 207
画面読み上げソフトウェア 217, 220, 224,
225
例 223, 224
関数 129, 130, 133, 134, 152
欠損値の処理 110
データベース 119, 120

関数 (続き)
ユーザー定義関数 (UDF) 119
例 111
CLEM 式内 119
@BLANK 109
@FIELD 119, 160
@GLOBAL_MAX 158
@GLOBAL_MEAN 158
@GLOBAL_MIN 158
@GLOBAL_SDEV 158
@GLOBAL_SUM 158
@PARTITION 160
@PREDICTED 119, 160
@TARGET 119, 160
カンマ 43
キーボード・ショートカット 217, 220,
221, 222
キーワード
ノードの注釈設定 63
起動ダイアログ・ボックス 198
脚注 95, 98, 99, 100
再番号付け 99
マーカー 95
キャッシュ
キャッシュの設定 39
使用可能化 196
取り消し 39, 43
保存 39
キャッシュ・ファイル・ノード
ロード 66
キャプション 98
行 100
ピボット・テーブルでの選択 100
行と列の移動 90
行と列の入れ替え 90
行と列の再配列 90
行の色の変更
ピボット・テーブル 96
行または列をグループ化する 90
距離関数 140
切り取り 19
空間処理関数 140
空白 106, 116
空白の処理
CLEM 関数 159
空白文字
文字列から削除する 115, 143
クライアント
デフォルト・ディレクトリー 196
クラス 18, 187, 188
グラフ 71, 78, 102

グラフ (続き)

- エクスポート 78
- 出力の保存 65
- 非表示 71
- ピボット・テーブルからの作成 102
- プロジェクトへの追加 188
- グラフのエクスポート 78, 84, 85
- グラフの保存 78, 84, 85
 - メタファイル 78
 - BMP ファイル 78, 84
 - EMF ファイル 78
 - EPS ファイル 78, 85
 - JPEG ファイル 78, 84
 - PICT ファイル 78
 - PNG ファイル 85
 - PostScript ファイル 85
 - TIFF ファイル 85
- グリッド線 100
 - ピボット・テーブル 100
- グループ化記号
 - 数字の表示形式 43
- グループ・ラベル 90
- グループ・ラベルの削除 90
- グループ・ラベルを挿入する 90
- グループ・ラベルを追加する 90
- グローバル値
 - CLEM 式内 122
- グローバル関数 158
- 警告 50
 - オプションの設定 197
- 形式を選択してコピー 76
- 罫線 97, 100
 - 隠れた罫線を表示する 100
- 継続テキスト 97
 - ピボット・テーブルの 97
- 欠損値 106, 116
 - レコード 106
 - CLEM 式 110
 - system 107
- 欠損値検査ノード
 - 欠損値 106
- 言語
 - オプション 195
 - 出力言語の変更 92
- 検索
 - ストリーム内でのノード 54
- 検索および置換
 - ビューア文書 74
- 高速ピボット・テーブル 103
- コマンド・ライン
 - IBM SPSS Modeler の開始 10
- コメント
 - キーボード・ショートカット 221, 222
 - ストリームのすべてを一覧表示 62
 - ノードとストリームの 58

[サ行]

- サーバー
 - サーバーの COP の検索 12
 - 接続の追加 11
 - デフォルト・ディレクトリー 196
 - ログイン 10
- 最小化 21
- サイズ 73
 - アウトライン 73
- サイズの変更 21
- サブパレット
 - 作成 207
 - パレット上の表示 206
 - パレット・タブからの削除 206
- 三角関数 139
- シーケンス関数 152
- 資格情報
 - IBM SPSS Collaboration and Deployment Services Repository 165
- 時間関数 129, 130
 - time_before 136, 148
 - time_hours_difference 148
 - time_in_hours 148
 - time_in_mins 148
 - time_in_secs 148
 - time_mins_difference 148
 - time_secs_difference 148
- 時間と日付の関数 129, 130
- 時間の形式 44, 129, 130
- 時間フィールド
 - 変換 152
- 式 127
- 式ビルダー 220
 - アクセス 119
 - 概要 119
 - 使用 119
- 次数
 - 測定単位 45
- 指数関数 138
- 指数表記
 - 出力での抑制 103
- システム
 - オプション 195
- システム欠損値
 - レコード 107
- 実行時間、表示 50
- 実行を中止 19
- 実数 127, 128
- 実例
 - アプリケーション ガイド 3
 - 概要 4
- 出力 17, 71, 72, 76, 78, 87, 103
 - 暗号化 87
 - 位置合わせ 72, 103

出力 (続き)

- 移動 72
- インタラクティブ 77
- エクスポート 78
- コピー 72
- 削除 72
- 出力言語の変更 92
- 他のアプリケーションへの貼り付け 76
- 中央揃え 72, 103
- 非表示 71
- ビューア 71
- 表示 71
- 保存 87
- 出力オブジェクト
 - IBM SPSS Collaboration and Deployment Services Repository での格納 169
- 出力ノード 36
- 出力のエクスポート 78, 82, 83
 - Excel 形式 78, 81
 - HTML 79
 - HTML 形式 78
 - PDF 形式 78, 83
 - PowerPoint 形式 78
 - Web レポート 80
 - Word 形式 78, 80
- 出力の削除 72
- 出力の中央揃え 72, 103
- 出力の保存 78, 82, 83
 - テキスト形式 78, 83
 - Excel 形式 78, 81
 - HTML 78, 79
 - HTML 形式 78
 - PDF 形式 78, 83
 - PowerPoint 形式 78, 82
 - Web レポート 80
 - Word 形式 78, 80
- 出力ファイル
 - 保存 65
- ショートカット
 - 一般的な使用法 69
 - キーボード 22, 217, 220, 221, 222
- 条件 114
- 小数桁数
 - 表示形式 45
- 小数点記号
 - 数字の表示形式 43
- 情報関数 134
- 新機能 7
- シングル・サインオン 10
- シングル・サインオン、IBM SPSS Collaboration and Deployment Services Repository 163, 164
- ズーム 19
- 垂直ラベル・テキスト 90
- 数字の表示形式 45

数値関数 138
スクリプト 24, 111
スクリプトの色
設定 199
スケーリング
ピボット・テーブル 95, 97
スコアリング
ブランチ 58, 179, 180, 181
スタック オーバーフロー エラー 195
ステート
保存 64, 65
ロード 66
ストリーム 9, 15, 177
オプション 42, 43, 44, 45, 47, 48, 49
コメントの追加 58
作成 35
実行 57
実行時間を表示 50
注釈 58, 63
地理空間座標系 49
展開オプション 177
名前変更 55, 63
ノードの迂回 37
ノードの接続 37
ノードの追加 36, 38
ノードを無効化 38
バックアップ・ファイル 64
ビューへの調整 22
プロジェクトへの追加 188, 189
保存 64
ロード 66
IBM SPSS Collaboration and
Deployment Services Repository で
の格納 168
ストリームの実行 57
ストリームのすべてのコメントを一覧表示
62
ストリームの説明 55, 57
ストリームのデフォルトの文字コード 43
ストリームのビューへの調整 22
ストリームのプロパティ
Analytic Server 48
ストリーム名 63
ストリーム領域
設定 48
ストリーム・パラメーター 52
ストリーム・ファイルのバックアップ
復元 64
正規分布
確率関数 140
整数 127
生成されたモデル・パレット 17
セッション・パラメーター 52
接続
サーバー・クラスター (server
cluster) 12

接続 (続き)
IBM SPSS Analytic Server への 12
IBM SPSS Collaboration and
Deployment Services Repository へ
の 164, 165
IBM SPSS Modeler Server への 10,
11, 12
接続の COP の検索 12
セット 43
セル・プロパティ 97, 98
線型回帰
PMML としてのエクスポート 200
選択方法 100
ピボット・テーブルでの行と列の選択
100
ソース・ノード 36
データのマッピング 67
ソート
ピボット・テーブルの行 90
ソート・ノード
パフォーマンス 211
層 85, 92, 93, 95, 97
印刷 85, 95, 97
作成 92
ピボット・テーブル 92
表示 92, 93
属性 (attribute) 25
測定システム 103

[タ行]

ターミナル・ノード 36
ダイアログ・ボックス 103
変数の表示順 103
変数名の表示 103
変数ラベルの表示 103
タイトル
ビューアへの追加 74
他のアプリケーションへの出力のコピーと
貼り付け 76
他のアプリケーションへの出力の貼り付け
76
置換ノード
欠損値 109
知識の発見 25
注釈
コメントへの変換 62
ストリーム 58, 63
ノード 58, 63
プロジェクト 191
folder 191
重複レコード・ノード
パフォーマンス 211
地理空間座標
システムの選択 49
表示書式 49

地理空間座標系の選択 49
地理空間座標の形式 49
ツールチップ
ノードの注釈設定 63
ツールバー 19
追加
プロジェクトに 188
通貨の表示形式 45
通知
オプションの設定 197
ツリー・ベースの分析
典型的なアプリケーション 25
データ
preview 41
データ型ノード
欠損値 109
パフォーマンス 211
データ検査ノード
探索に使用 25
データ・マイニングでの使用 26
データのマッピング 68
データ分割ノード
パフォーマンス 211
データベース
関数 119, 120
データベース関数
ユーザー定義関数 (UDF) 120
CLEM 式内 120
データ・エディター
記述統計量のオプション 104
複数の開かれたデータ・ファイル 103
データ・ストリーム
作成 35
データ・タイプ 113
パラメーター 52
データ・ファイル
複数の開かれたデータ・ファイル 103
データ・マイニング 25
アプリケーションの例 34
戦略 27
データ・マイニングの用途 26
データ・マッピング・ツール 66, 67
テーブル
プロジェクトへの追加 188
テーブルの分割 101
テーブル・ブレイクの制御 101
テーブル・グラフ 102
テーブル・ブレイク 101
テーブル・ルック 94
作成 94
適用 94
ディクショナリー・ファイル 224
ディシジョン ツリー
アクセシビリティ 225
ディレクトリー
デフォルト 196

テキスト 74, 78, 83
出力をテキストとしてエクスポート
78, 83
ビューアにテキスト・ファイルを追加
する 74
ビューアへの追加 74
テキストの文字コード 43
テキスト・データ・ファイル
エンコード 227
テナント
IBM SPSS Analytic Server 12
デフォルト
プロジェクト・フェーズ 188
展開 164
展開オプション 177
展開タイプ 177
典型的なアプリケーション 25
テンプレート 67
テンプレート・フィールド 68
等号演算子 136
特殊関数 160
特殊文字
文字列から削除する 115
ドメイン名 (Windows)
IBM SPSS Modeler Server 10

[ナ行]

ナゲット 58
defined 17
ナビゲート
キーボード・ショートカット 217
名前変更
ストリーム 55
ノード 63
日をロールオーバー 44
ニューラル ネットワーク ノード
パフォーマンス 212
ラージ・セット 43
入力ノード
更新 43
ヌル 116
ノード 9
オプションの設定 39
概要 36
カスタム・サブパレットの作成 207
カスタム・パレットの作成 205
検索 54
コメントを追加 58
削除 36
実行時間 50
順序 209
使用可能化 38
使用不可化 38, 39
ストリーム内の迂回 37
ストリーム内の接続 37

ノード (続き)
ストリーム内の無効化 38
接続の削除 38
注釈 58, 63
重複 39
追加 36, 38
データのプレビュー 41
データ・プレビュー 41
パフォーマンス 211, 212
パレット上の表示 206
パレットからの削除 206
パレット・タブのカスタマイズ 207
プロジェクトへの追加 188, 189
編集 39
保存 64
ロード 66
ロック 41
IBM SPSS Collaboration and
Deployment Services Repository で
の格納 169
ノードとストリームに名前をつける 63
ノードの有効化 38
ノードのロック 41
ノード名 63
ノードを無効化 38, 39
ノード・パレットの選択 206
ノイズを含んだデータ 26

[ハ行]

バージョン・ラベル、IBM SPSS
Collaboration and Deployment
Services Repository オブジェクト 176
背景色 97
パスワード
IBM SPSS Analytic Server 12
バッチ機能 103
ジャーナル・ファイルのコマンド・シ
ンタックスの使用 103
幅広のテーブル
Microsoft Word への貼り付け 76
パフォーマンス
プロセス・ノード 211
モデル作成ノード 212
CLEM 式 213
パラメーター
CLEM 式内 122
貼り付け 19
パレット 15
カスタマイズ 205
パレット・タブのカスタマイズ 207
比較関数 136
ビジュアル・プログラミング 14
日付
操作 152
変換 152
日付関数 129, 130
date_before 136, 148
date_days_difference 148
date_in_days 148
date_in_months 148
date_in_weeks 148
date_in_years 148
date_months_difference 148
date_weeks_difference 148
date_years_difference 148
@TODAY 関数 148
日付と時間関数
datetime_date 148
datetime_day 148
datetime_day_name 148
datetime_day_short_name 148
datetime_hour 148
datetime_in_seconds 148
datetime_minute 148
datetime_month 148
datetime_month_name 148
datetime_month_short_name 148
datetime_now datetime_second 148
datetime_time 148
datetime_timestamp 148
datetime_weekday 148
datetime_year 148
日付の書式 44, 129, 130
日付/時刻の値 116
必須フィールド 66, 68
ビット単位関数 141
非等号演算子 136
非表示 71, 72, 93
脚注 98
キャプション 98
行と列 93
ディメンション・ラベル 93
表題 93
プロシージャータの結果 72
ピボット・テーブル 71, 76, 78, 85, 89,
90, 92, 93, 94, 95, 96, 97, 100, 101,
102, 103
アイコンの使用 89
値のラベル 91
新しいテーブルのデフォルト・ルック
103
位置合わせ 98
大きなテーブルの印刷 101
隠れた罫線を表示する 100
脚注 98, 99, 100
脚注プロパティ 95
キャプション 98
行と列の移動 90
行と列の入れ替え 90
行と列の選択 100
行と列の挿入 91

- ピボット・テーブル (続き)
 - 行の色の変更 96
 - 行のソート 90
 - 行または列のグループ化の解除 90
 - 行または列をグループ化する 90
 - グリッド線 100
 - グループ・ラベルの削除 90
 - グループ・ラベルを挿入する 90
 - 罫線 97
 - 継続テキスト 97
 - 高速ピボット・テーブル 103
 - セルの書式 96
 - セルの表示と非表示 93
 - セル幅 100
 - セル・プロパティ 97, 98
 - 全般的なプロパティ 95
 - 層 92
 - 操作 89
 - 層の印刷 85
 - 他のアプリケーションへの貼り付け 76
 - テーブルからのグラフの作成 102
 - テーブルとして貼り付け 76
 - テーブルをより早く表示 103
 - テーブル・ブレイクの制御 101
 - デフォルトの列幅の調整 103
 - 背景色 97
 - 非表示 71
 - ピボット 89
 - 表示順序の変更 89, 90
 - 表示する行数の制御 95
 - フォント 97
 - プロパティ 95
 - ページに合わせたスケールリング 95, 97
 - 変更を元に戻す 92
 - 編集 89
 - 変数ラベル 91
 - 余白 98
 - ラベルの回転 90
 - ルックの変更 94
 - レガシー・テーブル 102
 - HTML としてエクスポート 78
 - language 92
- ピボット・テーブルでの色 97
 - 罫線 97
- ピボット・テーブルのセル 93, 96, 100
 - 書式 96
 - 選択 100
 - 幅 100
 - 非表示 93
 - 表示 93
- ビューア 71, 72, 73, 74, 87, 103, 104
 - アウトライン 73
 - アウトラインのサイズ変更 73
 - アウトラインの展開 73
 - アウトラインを省略する 73
 - アウトライン・フォントの変更 74
- ビューア (続き)
 - アウトライン・ペイン 71
 - アウトライン・レベルの変更 73
 - 値のラベルの表示 104
 - 結果の非表示 71
 - 結果ペイン 71
 - 出力項目間のスペース 87
 - 出力の移動 72
 - 出力の削除 72
 - 情報の検索および置換 74
 - データ値の表示 104
 - 表示オプション 103
 - 文書の保存 87
 - 変数名の表示 104
 - 変数ラベルの表示 104
- ビューの表示時のエラー
 - メモリー不足 195
- 表 101, 220
 - 位置合わせ 98
 - 出力の保存 65
 - セル・プロパティ 97, 98
 - テーブル・ブレイクの制御 101
 - 背景色 97
 - フォント 97
 - 余白 98
- 評価ノード
 - パフォーマンス 211
- 表記方法 133
- 表示 71, 93
 - 脚注 98
 - キャプション 98
 - 行または列 93
 - 結果 71
 - ディメンション・ラベル 93
 - 表題 93
- 表示形式
 - グループ化記号 45
 - 小数桁数 45
 - 地理空間座標 49
 - currency 45
 - numbers 45
 - scientific 45
- 表示順 89, 90
- 表示する行数の制御 95
- 表題 74
- 開く
 - 出力 66
 - ステート 66
 - ストリーム 66
 - ノード 66
 - プロジェクト 188
 - モデル 66
- ピリオド 43
- ヒント
 - アクセシビリティ 225
 - 一般的な使用法 69
- ファイル 74
 - ビューアにテキスト・ファイルを追加する 74
- フィールド 25, 127, 129
 - 値の表示 123
 - CLEM 式内 122
- フィールドの関連付けの解除 66
- フィールドのマッピング 66
- フォルダー、IBM SPSS Collaboration and Deployment Services Repository 173, 174
- フォント 74, 97, 198, 199
 - アウトライン・ペイン 74
- 複数回答セット
 - CLEM 式内 118, 122
- 複数カテゴリー設定
 - CLEM 式内 118
- 複数二分法設定
 - CLEM 式内 118
- 複数の IBM SPSS Modeler セッション 14
- 複数の開かれたデータ・ファイル 103
- フッター 86
- ブランチ、モデル作成およびスコアリング 58, 179, 180, 181
- プロジェクト 18, 187
 - オブジェクトの追加 189
 - オブジェクトのプロパティ 191
 - クラス・ビュー 188
 - 作成 188
 - 新規作成 188
 - 注釈 191
 - デフォルト・フォルダーの設定 188
 - 閉じる 192
 - フォルダーのプロパティ 191
 - プロパティの設定 190
 - レポートの生成 192
 - CRISP-DM ビュー 187
 - IBM SPSS Collaboration and Deployment Services Repository での 190
 - IBM SPSS Collaboration and Deployment Services Repository での格納 168
- プロセス・ノード 36
 - パフォーマンス 211
- プロパティ 95
 - データ・ストリーム 42
 - ピボット・テーブル 95
 - 表 95
 - プロジェクト・フォルダー 191
 - レポート・フェーズ 192
- 文書 3
- 分布関数 140
- ページ設定 86, 87
 - グラフ・サイズ 87

- ページ設定 (続き)
 - ヘッダーとフッター 86
- ページの番号付け 87
- ヘッダー 86
- 変換関数 135
- 変数 25, 103
 - ダイアログ・ボックスでの表示順 103
- 変数選択ノード
 - 欠損値 106
- 変数名 103
 - ダイアログ・ボックス 103
- 変数ラベル 91, 103, 104
 - アウトライン・ペイン 104
 - ダイアログ・ボックス 103
 - ピボット・テーブル 104
- ポート番号
 - IBM SPSS Modeler Server 10, 11
- 報告書
 - 出力の保存 65
 - 生成 192
- ホスト名 (host name)
 - IBM SPSS Modeler Server 10, 11
- 保存
 - 出力オブジェクト 65
 - ステート 64, 65
 - ストリーム 64
 - ノード 64
 - 複数のオブジェクト 65
- ホット・キー 22

[マ行]

- マウス
 - IBM SPSS Modeler での使用 22, 37
- マウスの中央ボタン
 - シミュレート 22, 37
- マシン学習 25
- メイン・ウィンドウ 15
- メタファイル 78
 - グラフのエクスポート 78
- メッセージ
 - 生成された SQL の表示 47
- メモリー 103
 - 管理 195, 196
 - スタック オーバーフロー エラー 195
- メモリー不足 195
- 面
 - 空間処理関数 140
- 文字 127, 128
- 文字列 127, 128
 - 一致 115
 - 置換 115
 - CLEM 式内の操作 115
- 文字列関数 143
- 文字列の連結 135
- モデル 58

- モデル (続き)
 - エクスポート 200
 - 更新 180, 181
 - 置換 197
 - ブランチ 58
 - プロジェクトへの追加 188
- IBM SPSS Collaboration and Deployment Services Repository での格納 169
- モデル ナゲット 58
- 「モデル」パレット 169
- モデル作成ノード 36
 - パフォーマンス 212
 - モデル作成パレット・タブのカスタマイズ 207
- モデルの置換 197
- モデルのリフレッシュ 180, 181
- モデル・リフレッシュ 177
- 元に戻す 19

[ヤ行]

- ユーザー ID
 - IBM SPSS Modeler Server 10
- ユーザー定義関数 (UDF) 119, 120
- ユーザー・オプション 196
- 優先順位 130
- ようこそダイアログ・ボックス 198
- より小演算子 136
- より大演算子 136

[ラ行]

- ラジアン
 - 測定単位 45
- ラベル、IBM SPSS Collaboration and Deployment Services Repository オブジェクト 176
- ラベルの回転 90
- ランタイム・プロパティ 52
- リスト 127, 128
- 領域 15
- ルール・セット
 - 評価 43
- ルール・ノードの構築
 - ロード 66
- レガシー・テーブル 102
- レコード 25
 - 欠損値 106
 - システム欠損値 107
- レコード結合ノード
 - パフォーマンス 211
- 列 100
 - ピボット・テーブルでの選択 100
 - ピボット・テーブルでの幅の変更 100

- 列幅 95, 100, 103
 - 折り返されたテキスト幅の制御 95
 - 最大幅の制御 95
 - デフォルトの幅の制御 103
 - ピボット・テーブル 100
- レポート
 - プロジェクトへの追加 188
 - プロパティの設定 192
- ロード
 - ステート 66
 - ノード 66
- ログ・ファイル
 - 生成された SQL の表示 47
- ロケール
 - オプション 195
- ロジスティック回帰 225
 - PMML としてのエクスポート 200
- 論理関数 138

[数字]

- 508 準拠 215

A

- abs 関数 138
- Aggregate ノード
 - パフォーマンス 211
- allbutfirst 関数 143
- allbutlast 関数 143
- alphabefore 関数 143
- AND 演算子 138
- arccos 関数 139
- arccosh 関数 139
- arcsin 関数 139
- arcsinh 関数 139
- arctan 関数 139
- arctan2 関数 139
- arctanh 関数 139
- area 関数 140

B

- BMP ファイル 78, 84
 - グラフのエクスポート 78, 84

C

- case 25
- cdf_chisq 関数 140
- cdf_f 関数 140
- cdf_normal 関数 140
- cdf_t 関数 140
- Champion Challenger 分析 164, 177
- CLEM 119

CLEM (続き)
 概要 24, 111
 関数 119, 120
 式 114, 127
 式の検査 123
 データ型 127, 128, 129
 例 111
 CLEM 式の作成 119
 language 127

CLEM 関数
 空間 140
 空白とヌル 159
 グローバル 158
 欠損値 110
 三角関数 139
 使用可能なもののリスト 132
 情報 134
 数値 138
 特殊関数 160
 日時 148
 比較 136
 ビット単位 141
 変換 135
 ランダム 142
 論理 138
 probability 140
 sequence 152
 string 143

CLEM 式
 パフォーマンス 213

CLEM 式内の円記号 128

CLEM 式の検査 123

close_to
 空間処理関数 140

close_to 関数 140

Cognos Active Report 80

Coordinator of Processes 12

COP 12

copy 19

cos 関数 139

cosh 関数 139

count_equal 関数 117, 136

count_greater_than 関数 117, 136

count_less_than 関数 117, 136

count_non_nulls 関数 136

count_not_equal 関数 117, 136

count_nulls 関数 110, 117, 136

count_substring 関数 143

CRISP-DM 18, 187
 プロジェクト・ビュー 187

CRISP-DM プロセス・モデル 27, 28

crosses
 空間処理関数 140

crosses 関数 140

D

datetime_date 関数 135

date_before 関数 136

DIFF 関数 152

distance
 空間処理関数 140

div 関数 138

DTD 184

E

endstring 関数 143

EPS ファイル 78, 85
 グラフのエクスポート 78, 85

Excel 形式
 出力のエクスポート 78, 81

F

f 分布
 確率関数 140

factor 225

first_index 関数 118, 136

first_non_null 関数 118, 136

first_non_null_index 関数 118, 136

fracof 関数 138

H

hasendstring 関数 143

hasmidstring 関数 143

hasstartstring 関数 143

hassubstring 関数 143

HTML 78, 79
 出力のエクスポート 78, 79

HTML 出力
 スクリーン・リーダー 225

I

IBM SPSS Analytic Server
 接続 12
 複数の接続 12

IBM SPSS Collaboration and
 Deployment Services 164

IBM SPSS Collaboration and
 Deployment Services Repository 163,
 164
 オブジェクトおよびバージョンの削除
 174
 オブジェクトの格納 165
 オブジェクトの取得 170
 オブジェクトのプロパティ 175

IBM SPSS Collaboration and
 Deployment Services Repository (続き)
 オブジェクトのロックおよびロック解
 除 173
 検索 172
 参照 165
 資格情報 165
 シングル・サインオン 164
 接続 164, 165
 フォルダー 173, 174
 プロジェクトの転送 190

IBM SPSS Collaboration and
 Deployment Services Repository オブ
 ジェクトのロック 173

IBM SPSS Collaboration and
 Deployment Services Repository オブ
 ジェクトのロック解除 173

IBM SPSS Collaboration and
 Deployment Services Repository から
 のオブジェクトの取得 170

IBM SPSS Collaboration and
 Deployment Services Repository での
 オブジェクトの格納 165

IBM SPSS Collaboration and
 Deployment Services Repository のオ
 ブジェクトの検索 172

IBM SPSS Modeler 1, 14
 アクセシビリティ 215
 オプション 195
 概要 9, 195
 コマンド・ラインからの実行 10
 はじめに 9
 ヒントとショートカット 69
 文書 3

IBM SPSS Modeler Advantage 164

IBM SPSS Modeler Server 2
 ドメイン名 (Windows) 10
 ポート番号 10, 11
 ホスト名 (host name) 10, 11
 ユーザー ID 10
 password 10

IBM SPSS Modeler Server 接続の追加
 11, 12

IBM SPSS Modeler Server へのログイン
 10

if, then, else 関数 138

INDEX 関数 152

integer_bitcount 関数 141

integer_leastbit 関数 141

integer_length 関数 141

intof 関数 138

isalphacode 関数 143

isendstring 関数 143

islowercode 関数 143

ismidstring 関数 143

isnumbercode 関数 143

isstartstring 関数 143
issubstring 関数 143
issubstring_count 関数 143
issubstring_lim 関数 143
isuppercode 関数 143
is_date 関数 134
is_datetime 関数 134
is_integer 関数 134
is_number 関数 134
is_real 関数 134
is_string 関数 134
is_time 関数 134
is_timestamp 関数 134

J

Java 226
JAWS 215, 224, 226
JPEG ファイル 78, 84
 グラフのエクスポート 78, 84

K

Kohonen ノード
 パフォーマンス 212
 ラージ・セット 43
K-Means ノード
 パフォーマンス 212
 ラージ・セット 43

L

labels 90
 値 184
 グループ・ラベルを挿入する 90
 削除 90
 表示 43
 変数 (variable) 184
last_index 関数 118, 136
LAST_NON_BLANK 関数 152
last_non_null 関数 118, 136
last_non_null_index 関数 118, 136
length 関数 143
locchar 関数 143
locchar_back 関数 143
log 関数 138
log10 関数 138
lowertoupper 関数 143

M

managers 17
matches 関数 143
MAX 関数 152
max 関数 136

max_index 関数 118, 136
max_n 関数 117, 136
MEAN 関数 152
mean_n 関数 117, 138
member 関数 136
MIN 関数 152
min 関数 136
min_index 関数 118, 136
min_n 関数 117, 136
mod 関数 138

N

negate 関数 138
NOT 演算子 138
numbers 116, 127, 128
num_points
 空間処理関数 140
num_points 関数 140

O

OFFSET 関数 152
oneof 関数 142
OR 演算子 138
overlap
 空間処理関数 140
overlap 関数 140

P

parameters
 スコアリング 179
 ストリーム 52
 タイプ 52
 モデル構築 179
 ランタイム・プロパティ 52
 session 52
password
 IBM SPSS Modeler Server 10
PDF
 出力のエクスポート 78, 83
pi 関数 139
PMML
 エクスポート・オプション 200
 モデルのインポート 184, 185
 モデルのエクスポート 184, 185
PMML モデル
 線型回帰 200
 ロジスティック回帰 200
PNG ファイル 78, 85
 グラフのエクスポート 78, 85
PostScript ファイル (カプセル化) 78, 85
 グラフのエクスポート 78, 85
PowerPoint 82

PowerPoint (続き)
 出力を PowerPoint としてエクスポート 82
PowerPoint 形式
 出力のエクスポート 78
PowerPoint ファイル 188
power (指数) 関数 138
preview
 ノード・データ 41

R

random 関数 142
random0 関数 142
refresh
 ソース・ノード 43
rem 関数 138
replace 関数 143
replicate 関数 143
round 関数 138

S

SAS ファイル
 エンコード 227
scrolling
 オプションの設定 48
SDEV 関数 152
sdev_n 関数 117, 138
sign 関数 138
sin 関数 139
SINCE 関数 152
sinh 関数 139
skipchar 関数 143
skipchar_back 関数 143
soundex 関数 148
soundex_difference 関数 148
spaces
 文字列から削除する 115, 143
SQL 生成
 プレビュー 47
 ロギング 47
sqrt 関数 138
startstring 関数 143
Statistics ファイル
 エンコード 227
stripchar 関数 143
strmember 関数 143
subscrs 関数 143
substring 関数 143
substring_between 関数 143
SUM 関数 152
sum_n 関数 117, 138

T

t 分布

確率関数 140

tan 関数 139

tanh 関数 139

testbit 関数 141

THIS 関数 152

TIFF ファイル 85

グラフのエクスポート 78, 85

time_before 関数 136

to_date 関数 135, 148

to_dateline 関数 148

to_datetime 関数 135

to_integer 関数 135

to_number 関数 135

to_real 関数 135

to_string 関数 135

to_time 関数 135, 148

to_timestamp 関数 135, 148

trim 関数 143

trimend 関数 143

trim_start 関数 143

@FIELD 関数 110, 160

@FIELDS_BETWEEN 関数 110, 117, 160

@FIELDS_MATCHING 関数 110, 117,
160

@INDEX 関数 152

@LAST_NON_BLANK 関数 152, 159

@MAX 関数 152

@MEAN 関数 152

@MIN 関数 152

@MULTI_RESPONSE_SET 関数 118, 160

@NULL 関数 110, 134, 159

@OFFSET 関数 152
パフォーマンスに関する考慮事項 213

@PARTITION_FIELD 関数 160

@PREDICTED 関数 160

@SDEV 関数 152

@SINCE 関数 152

@SUM 関数 152

@TARGET 関数 160

@TESTING_PARTITION 関数 160

@THIS 関数 152

@TODAY 関数 148

@TRAINING_PARTITION 関数 160

@VALIDATION_PARTITION 関数 160

U

undef 関数 159

Unicode のサポート 227

unicode_char 関数 143

unicode_value 関数 143

uppertolower 関数 143

URL

IBM SPSS Analytic Server 12

UTF-8 エンコード 43, 227

V

values 113

value_at 関数 118, 136

W

Web レポート 80

出力のエクスポート 80

within

空間処理関数 140

within 関数 140

Word 形式

出力のエクスポート 78, 80

幅広のテーブル 78

[特殊文字]

@BLANK 関数 110, 134, 159

@DIFF 関数 152



Printed in Japan

日本アイ・ビー・エム株式会社

〒103-8510 東京都中央区日本橋箱崎町19-21