

IBM SPSS Modeler Server
16 管理および
パフォーマンス・ガイド

IBM

お願い

本書および本書で紹介する製品をご使用になる前に、75 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(r) SPSS(r) Modeler バージョン 16 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 Server 16
Administration and Performance 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

前書き	v
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 の IBM SPSS Collaboration and Deployment Services	2
IBM SPSS Modeler 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(E)	5
Demos フォルダー	5
第 2 章 アーキテクチャーとハードウェア に関する推奨事項	7
IBM SPSS Modeler アーキテクチャー	7
アーキテクチャーの説明	7
ハードウェアの要件	9
一時的なディスク領域と RAM の要件	10
データへのアクセス	11
データ・ファイルの参照	12
IBM SPSS Statistics データ・ファイルのインポ ート	13
インストール手順	13
第 3 章 IBM SPSS Modeler のサポート	15
IBM SPSS Modeler Server に接続中	15
シングル・サインオンの設定	16
IBM SPSS Modeler Server 接続の追加および編集	18
IBM SPSS Collaboration and Deployment Services のサーバーの検索	19
データとファイル・システム	19
ユーザー認証	19
ファイルの作成	20
結果の違い	20
第 4 章 IBM SPSS Modeler Server の 管理	21
IBM SPSS Modeler Server の開始および停止	21
Windows 上で起動、停止、またはステータスをチ ェックする手順	21
UNIX 上で起動、停止、またはステータスをチ ェックする手順	21
非応答性サーバー・プロセス (UNIX システム) の処 理	22

管理	23
IBM SPSS Modeler Administration Consoleの使用	23
Modeler Administration Console の開始	23
Modeler Administration Console でのアクセスの構 成	24
SPSS Modeler Server への接続	24
SPSS Modeler Server の構成	25
SPSS Modeler Server の監視	29
options.cfg ファイルの使用	30
使用されていないデータベース接続の終了	30
安全にデータを転送するための SSL の使用	31
SSL の動作方法	31
SSL を使用したクライアント-サーバー間および サーバー/サーバー間通信の保護	31
第 5 章 パフォーマンスの概要	35
サーバーのパフォーマンスと最適化の設定	35
クライアントのパフォーマンスと最適化の設定	35
データベースの使用率と最適化	37
SQL 最適化	37
第 6 章 SQL 最適化	39
SQL 生成の動作原理	40
SQL 生成の例	41
SQL 最適化の設定	42
生成された SQL のプレビュー	43
モデル・ナゲットの SQL の表示	43
SQL 生成を最大限にするためのヒント	43
SQL 生成をサポートしているノード	44
SQL 生成をサポートする CLEM 式と演算子	48
CLEM 式での SQL 関数の使用	51
SQL クエリーの記述	51
付録 A. UNIX プラットフォームでの Oracle の設定	53
SQL 最適化用の Oracle の設定	53
付録 B. UNIX 起動スクリプトの構成	55
はじめに	55
スクリプト	55
IBM SPSS Modeler Server の自動的な起動と停止	56
IBM SPSS Modeler Server の手動による起動と停止	56
スクリプトの編集	56
ファイル作成時の権限の制御	57
IBM SPSS Modeler Server と データ・アクセス・パ ック	57
ODBC 設定のトラブルシューティング	59
ライブラリー・パス	62
エンタープライズ・ビュー・ノードのドライバーを 構成	62

**付録 C. SPSS Modeler Server を UNIX
上で root 以外のプロセスとして設定および実行 65**

はじめに 65
非公開のパスワード・データベースがない状態で
root 以外のユーザーとして設定 65
非公開のパスワード・データベースを使用して root
以外のユーザーとして設定 66
SPSS Modeler Server を非 root ユーザーとして実行
する 67
ユーザー認証エラーのトラブルシューティング . . . 68

**付録 D. Windows での非公開パスワード・ファイルを使用した SPSS Modeler
Server の構成および実行 71**
はじめに 71

プライベート・パスワード・データベースを使用し
た設定 71

**付録 E. サーバー・クラスターでの負荷バ
ランシング 73**

特記事項 75
商標 76

索引 77

前書き

IBM® SPSS® Modeler は、IBM Corp. が開発した企業強化用のデータ・マイニング・ワークベンチです。SPSS Modeler を使用すると、企業はデータを詳しく調べることで顧客および一般市民とのリレーションシップを強化することができます。企業は、SPSS Modeler を使用して得られた情報に基づいて利益をもたらす顧客を獲得し、抱き合わせ販売の機会を見つけ、新規顧客を引き付け、不正を発見し、リスクを減少させ、政府機関へのサービスの提供を改善することができます。

SPSS Modeler の視覚的インターフェースを使用すると、特定ビジネスの専門知識を適用し、より強力な予測モデルを実現し、解決までの時間を短縮します。SPSS Modeler では、予測、分類、セグメント化、および関連性検出アルゴリズムなど、さまざまなモデリング手法を提供しています。モデルを作成した後は、IBM SPSS Modeler Solution Publisher により、企業全体の意思決定者やデータベースにモデルを配布することが可能になります。

IBM Business Analytics について

IBM Business Analytics ソフトウェアは、正確かつ完全で一貫した情報を提供し、意思決定者はその情報を信頼して用いることで業績向上を図ることができます。ビジネス・インテリジェンス、予測分析、財務実績および戦略管理、分析アプリケーション の包括的なポートフォリオを利用することによって、現在の実績を明確、迅速に理解し、将来の結果を予測することができます。豊富な産業用ソリューション、証明された実践法、それに専門家によるサービスを組み合わせることにより、あらゆる規模の会社組織が、最高の生産性を推進し、信頼できる意志決定を自動化し、そして、よりよい結果を実現させることができます。

このポートフォリオの一部として、IBM SPSS Predictive Analytics ソフトウェアを使用する組織は、将来のイベントを予測し、その洞察に基づいて積極的に行動し、より優れた業績を実現することができます。全世界の企業、政府、学術分野のお客様が IBM SPSS の技術を活用し、不正行為を減少させ、リスクを軽減させながら、顧客の獲得、保持、成長において、競争優位を高めることができます。IBM SPSS ソフトウェアを日常業務に組み込むことで、組織は企業としての予測能力を向上させ、意思決定の監督と自動化によってビジネス目標を達成し、高い競争力を獲得できます。詳細な情報、または営業担当者へのお問い合わせ方法については、<http://www.ibm.com/spss> を参照してください。

技術サポート

お客様はテクニカル・サポートをご利用いただけます。IBM Corp. 製品の使用方法、または対応するハードウェア環境へのインストールについてサポートが必要な場合は、テクニカル・サポートにご連絡ください。テクニカル・サポートのご利用には、<http://www.ibm.com/support>のIBM Corp. Web サイトをご覧ください。支援を要請される場合は、事前にユーザー、会社組織、そして、サポート契約を明確にしておいていただくよう、お願いします。

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リング手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、「<http://www.ibm.com/software/analytics/spss/products/modeler/>」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 Modeler Batch
- IBM SPSS Modeler Solution Publisher
- IBM SPSS Modeler Server の 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、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 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、リソース主体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されるようになりました。その結果、規模が比較的大きいデータ・セットを処理するパフォーマンスを実現しました。

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、メモリー集中型の操作は、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することができるため、SPSS Modeler Server は大きなデータ・セットに対し優れたパフォーマンスを示すことができます。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 は多くの SPSS Modeler Server 設定オプションを管理し、オプション・ファイルによって設定可能なグラフィカルアプリケーションです。アプリケーションには、SPSS Modeler Server のインストールを監視、構成するコンソール・ユーザー・インターフェースが用意されており、しかも、現在の SPSS Modeler Server のお客様には無料で提供されます。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し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 ライセンスが必要です。

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。

IBM SPSS Modeler Server の IBM SPSS Collaboration and Deployment Services

さまざまな IBM SPSS Collaboration and Deployment Services アダプターを使用すると、SPSS Modeler および SPSS Modeler Server が 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、またはシンククライアントアプリケーション IBM SPSS Modeler Advantage からアクセスできます。リポジトリをホストするシステム上のアダプターをインストールします。

IBM SPSS Modeler 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、エンティティの分析やソーシャル・ネットワーキングなどの特化したデータ、又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 は次のコンポーネントで構成されています。

IBM SPSS Modeler Entity Analytics により、IBM SPSS Modeler の予測分析に新たな次元が追加されます。予測分析は過去のデータから将来の行動を予測しようとするのに対し、エンティティ分析ではレコードの中でアイデンティティの競合を解決することで現在のデータの干渉性と一貫性を改善することに焦点を当てます。アイデンティティは、個人、組織、オブジェクトまたは曖昧さの存在する他のエンティティとなります。アイデンティティの解決は、顧客関係の管理、不正行為の検出、マネーロンダリング防止、国内および国際的なセキュリティなどのさまざまなフィールドにおいて重要になります。

IBM SPSS Modeler Social Network Analysis は、関係に関する情報を、個人およびグループの社会的行動を特徴づけるフィールドに変換します。ソーシャル・ネットワークの基底となる関係を説明するデータを使用して、IBM SPSS Modeler Social Network Analysis はネットワークの他の人の行動に影響を与えるソーシャル・リーダーを識別します。また、他のネットワーク参加者に最も影響を受ける人を確認できます。これらの結果を他の指標と組み合わせることによって、予測モデルの基準となる個人の包括的なプロフィールを作成できます。この社会的情報を含むモデルは、含まないモデルに比べてパフォーマンスが高くなります。

IBM SPSS Modeler Text Analytics は、高度な言語技術と Natural Language Processing (NLP) を使用して、多様な未構築のテキスト・データを急速に処理し、重要なコンセプトを抽出および組織化、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

オンライン・ヘルプ形式のドキュメントは、SPSS Modeler の「ヘルプ」メニューから使用できます。SPSS Modeler、SPSS Modeler Server、および SPSS Modeler Solution Publisher のアプリケーション ガイドやその他サポート資料が含まれています。

各製品の PDF 形式の完全なドキュメント (インストール手順を含む) は、各製品 DVD の *Documentation* フォルダーにもあります。インストール・マニュアルは、Web サイト <http://www-01.ibm.com/support/docview.wss?uid=swg27038316> からダウンロードできます。

これらの形式のドキュメントは、SPSS Modeler インフォメーション・センター <http://publib.boulder.ibm.com/infocenter/spssmodl/v16r0m0/> から入手できます。

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 の使用方法の一般的な概要です。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操作方法のほか、IBM SPSS Collaboration and Deployment Services、Predictive Applications、または IBM SPSS Modeler Advantage にストリームを展開するために、ストリームをパッケージ化する方法が記載されています。
- 「**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リング手法が用意されています。
- **IBM SPSS Modeler アルゴリズム・ガイド:** IBM SPSS Modeler で使用されている手法の数学的な基礎の説明があります。このガイドは、PDF 形式のみです。
- 「**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 5 ページの『アプリケーションの例(E)』を参照してください。
- 「**IBM SPSS Modeler スクリプトとオートメーション**」。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 Modeler のストリームやシナリオを IBM SPSS Collaboration and Deployment Services Deployment Manager のジョブを処理するステップとしての実行についての情報。
- **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製のプログラムを統合します。
- 「**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 の設定と管理の方法について説明します。
- **IBM SPSS Modeler Administration Console ユーザー・ガイド:** IBM SPSS Modeler Server を監視して設定するためのコンソール・ユーザー・インターフェースのインストールおよび使用に関する情報。コンソールは、Deployment Manager アプリケーションへのプラグインとして実装されます。
- 「**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。

- 「**IBM SPSS Modeler Entity Analytics ユーザー・ガイド**」。SPSS Modeler でエンティティ分析を使用する場合の情報。リポジトリのインストールと構成、エンティティ分析ノード、および管理用タスクについて説明します。
- 「**IBM SPSS Modeler Social Network Analysis ユーザー・ガイド**」。SPSS Modeler でソーシャル・ネットワーク分析を行うためのガイド。グループ分析、拡散分析などについて説明します。

- 「**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。
- **IBM SPSS Modeler Text Analytics Administration Console ユーザー・ガイド**：SPSS Modeler Text Analytics と使用するために IBM SPSS Modeler Server を監視して設定するためのコンソール・ユーザー・インターフェースのインストールおよび使用に関する情報。コンソールは、Deployment Manager アプリケーションへのプラグインとして実装されます。

アプリケーションの例(E)

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、特定のモデリング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た巨大なデータ・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に対して大規模です。

SPSS Modeler の「ヘルプ」メニューから「**アプリケーションの例**」を選択すると、例にアクセスすることができます。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* フォルダーにインストールされています。詳しくは、トピック『*Demos* フォルダー』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 フォルダー

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* フォルダーにインストールされています。このフォルダーには、Windows 「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「ファイルを開く」ダイアログ・ボックスの最近使ったディレクトリーの一覧から「*Demos*」を選択してアクセスすることもできます。

第 2 章 アーキテクチャーとハードウェアに関する推奨事項

IBM SPSS Modeler アーキテクチャー

このセクションでは、サーバー・ソフトウェア、クライアント・ソフトウェア、およびデータベースも含めた、IBM SPSS Modeler Server のアーキテクチャーを説明します。最適なパフォーマンスを達成するために IBM SPSS Modeler Server がどのようにして設計されたかについて説明し、適切なサイズのハードウェアを選択することでこのパフォーマンスを最大化するための推奨事項を挙げます。また、データ・アクセスについてのセクションもあり、必要な ODBC ドライバーのセットアップ場所も説明します。

アーキテクチャーの説明

IBM SPSS Modeler Server では、三層の分散型アーキテクチャーが使用されます。ソフトウェア操作は、クライアント・コンピューターとサーバー・コンピューターとで共有されます。スタンドアロンの IBM SPSS Modeler に比べて IBM SPSS Modeler Server をインストールして使用する場合は、特に大規模データ・セットを処理するときに顕著です。

- IBM SPSS Modeler Server は、Windows だけでなく、UNIX 上でも動作するため、インストール先をより柔軟に決定することができます。任意のプラットフォームで、データ・マイニング・プロセスに対するより速く、より大規模なサーバー・コンピューターであることに専念できます。
- IBM SPSS Modeler Server は高速パフォーマンス用に最適化されています。操作がデータベースへプッシュバックできない場合、IBM SPSS Modeler Server では、中間結果が一時ファイルとして、RAM ではなく、ディスクに格納されます。サーバーには通常、利用可能なかなりのディスク領域があるので、IBM SPSS Modeler Server により、非常に大きなデータ・セットに対して、ソート、レコード結合、およびレコード集計の操作を実行できます。
- クライアント/サーバーのアーキテクチャーを使用すると、組織内のデータ・マイニング・プロセスを集中化できます。この集中化によって、ビジネス・プロセスにおけるデータ・マイニングの役割を正式なものにすることを支援できます。
- IBM SPSS Modeler Administration Console (IBM SPSS Modeler Server に付属) や IBM SPSS Collaboration and Deployment Services (別売) のような管理ツールを使用すると、データ・マイニングのプロセスを監視でき、適切なコンピューティング リソースを確実に利用できます。IBM SPSS Collaboration and Deployment Services を使用して、一定のデータ・マイニング・タスクを自動化し、データ・モデルへのアクセスを管理し、組織全体で結果を共有することができます。

図「IBM SPSS Modeler Server アーキテクチャー」は、IBM SPSS Modeler の分散型アーキテクチャーのコンポーネントを示しています。

- **IBM SPSS Modeler.** クライアント・ソフトウェアは、エンド・ユーザーのコンピューターにインストールされます。このソフトウェアでユーザー・インターフェースが提供され、データ・マイニングの結果が表示されます。クライアントは IBM SPSS Modeler ソフトウェアの完全インストールですが、分散分析のために IBM SPSS Modeler Server に接続されると、その実行エンジンが活動停止になります。IBM SPSS Modeler は、Windows オペレーティング・システムでのみ動作します。
- **IBM SPSS Modeler Server.** IBM SPSS Modeler およびデータベースにネットワーク接続している、サーバー・コンピューターにインストールされたサーバー・ソフトウェア。IBM SPSS Modeler Server

は、サービス (Windows 上) またはデーモン・プロセス (UNIX 上) として実行され、クライアントからの接続を待機します。IBM SPSS Modeler を使用して作成されたストリームおよびスクリプトの実行を処理します。

- **データベース・サーバー。** データベース・サーバーは、稼動中のデータウェアハウス (大規模 UNIX サーバー上の Oracle など) である場合と、他の運用システムへの影響を抑えるためのローカル/部門サーバー上のデータマート (Windows 上の SQL Server など) である場合があります。

IBM SPSS Modeler Server アーキテクチャー

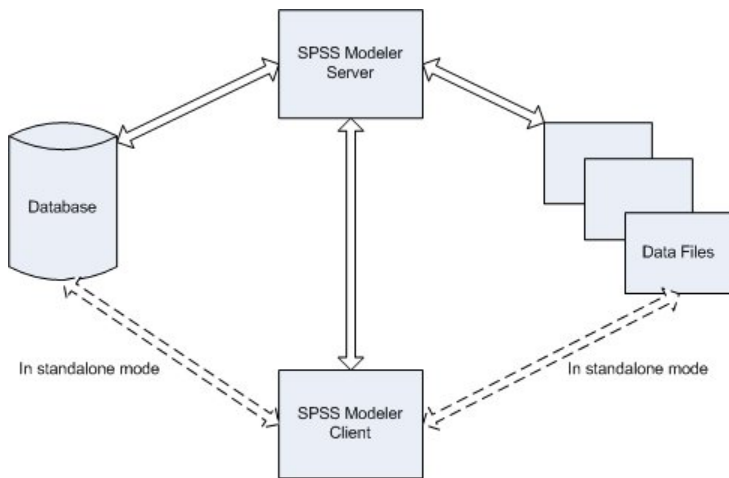


図 1. IBM SPSS Modeler Server アーキテクチャー

分散型アーキテクチャーでは、ほとんどの処理がサーバー・コンピューターで発生します。エンド・ユーザーがストリームを実行すると、IBM SPSS Modeler により、ストリームの記述がサーバーへ送られます。サーバーは、SQL で実行できる操作を判断し、適切なクエリーを作成します。これらのクエリーはデータベース内で実行され、結果のデータは、SQL を使用して表現できない処理を行うために、サーバーへ渡されます。処理が完了すると、関係のある結果だけがクライアントに戻されます。

必要な場合には、IBM SPSS Modeler Server により、すべての IBM SPSS Modeler 操作がデータベースの外側で実行されます。操作のためのデータを保持するために、RAM とディスク メモリーとの使用率のバランスが自動的にとられます。このプロセスで、IBM SPSS Modeler Server はフラット・ファイルとの完全な互換性がとられます。

サーバーのクラスターを使用して処理することにより、負荷バランシングも可能です。クラスタリングは、Coordinator of Processes プラグインを使用した場合 IBM SPSS Collaboration and Deployment Services 3.5 以降で可能です。詳しくは、トピック 73 ページの『付録 E. サーバー・クラスターでの負荷バランシング』を参照してください。IBM SPSS Modeler のサーバー・ログイン・ダイアログで直接 Coordinator of Processes で管理されたサーバーまたはクラスターに接続できます。詳しくは、トピック 15 ページの『IBM SPSS Modeler Server に接続中』を参照してください。

スタンドアロン・クライアント

IBM SPSS Modeler は、下図のように、自己完結型のデスクトップ・アプリケーションとして実行するように構成することもできます。詳しくは、15 ページの『第 3 章 IBM SPSS Modeler のサポート』を参照してください。

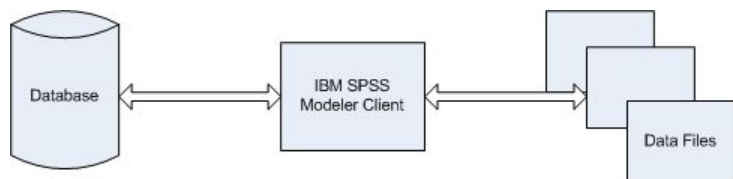


図2. IBM SPSS Modeler スタンドアロン

ハードウェアの要件

IBM SPSS Modeler Server のインストールを計画するときに、使用するハードウェアを考慮する必要があります。IBM SPSS Modeler Server は高速であるように設計されていますが、目的のデータ・マイニング・タスクに適したサイズのハードウェアを使用することで、効率を最大にすることができます。ハードウェアのアップグレードは、全体的なパフォーマンスを向上させるためにもっとも単純でもっとも経済的な方法であることが多い、と言えます。

専用サーバー：IBM SPSS Modeler Server が接続する可能性のあるデータベースも含めて、ほかのアプリケーションとリソースを競合することのない、専用のサーバーに IBM SPSS Modeler Server をインストールします。モデル作成の操作は特にリソース集約型の操作であり、ほかのアプリケーションと競合がない場合にパフォーマンスがよくなります。

注：IBM SPSS Modeler Server をデータベースと同じコンピューターにインストールすると、ネットワークの負荷を避けることでデータベースとサーバー間のデータ転送時間を減少できますが、ほとんどの場合の最良の構成は、リソースの競合を避けるために、サーバーとデータベースを別々のコンピューターに分けることです。2 つのコンピューター間的高速接続を提供することで、データ転送のコストを最小限に抑えます。

プロセッサ：コンピューターのプロセッサ数が、定期的に行うことが予想される並行タスクの数（同時に実行されるストリーム）より少なくならないようにする必要があります。一般的に、プロセッサ数が多くなるほど、パフォーマンスがよくなります。

- IBM SPSS Modeler Server の 1 つのインスタンスで複数のクライアント（ユーザー）からの接続を受け入れ、各クライアント接続で複数のストリームの実行を開始できます。したがって、1 つのサーバーには同時に進行中の複数の実行タスクがあります。
- 経験上の法則として、1 つのプロセッサには 1 人か 2 人のユーザー、2 つのプロセッサには 4 人までのユーザー、4 つのプロセッサには 8 人までのユーザーを許可します。2 人から 4 人までのユーザーごとに、作業の混合状態に応じて、追加のプロセッサを 1 つ加えます。
- 処理によっては SQL 最適化によってデータベースへプッシュバックされる可能性があることを考慮して、パフォーマンスの減少を最小にして、2 人以上のユーザーで CPU を共有することが可能です。
- マルチスレッド機能により 1 つのタスクが複数プロセッサを利用できるようにするので、プロセッサの追加で、一度にただ 1 つのタスクを実行中のような場合でも、パフォーマンスを向上させることができます。通常、マルチスレッド化は C5.0 モデル構築と、ソート、レコード集計、レコード結合など、一定のデータ準備の操作で使用されます。

64-bit のプラットフォーム：非常に大量のデータを処理したりそれでモデルを構築したりすることを計画している場合は、IBM SPSS Modeler Server プラットフォームとして 64-bit の Solaris、Windows、または Linux を使用し、コンピューターの RAM の量を最大にしてください。大規模なデータ・セットの場合、サーバーでは、32-bit のプラットフォームに与えられたプロセスごとのメモリー制限をすぐに使い果たして

しまい、データがディスクへこぼれ出し、実行時間が著しく増加します。64-bit のサーバーの実装では、追加の RAM を活用できます。最低 8 ギガバイト (GB) を推奨します。64-bit のサポートは、Solaris、Windows、および Linux プラットフォームで利用できます。

将来のニーズ： 実現可能であれば常に、サーバー・ハードウェアがメモリーと CPU の両面で拡張可能であるようにしておきます。つまり、使用率の増加に適合し (例えば、同時処理のユーザー数の増加や既存ユーザーの処理要件の増加)、将来の IBM SPSS Modeler Server のマルチスレッド機能の増加に対処できるようにしておきます。

一時的なディスク領域と RAM の要件

IBM SPSS Modeler Server では、大量のデータを処理するために一時的なディスク領域が使用されます。必要な一時領域の量は、処理するデータの量とデータ型、および実行する操作の種類によって異なります。データ量は、行の数と列の数の両方に比例します。処理する行と列が増えるほど、必要なディスク領域も増えます。

このセクションでは、一時的なディスク領域と追加の RAM が必要な条件と、必要な量の見積もり方法を説明します。データベース内で起きる処理のための一時的なディスク領域の要件は、そのような要件が各データベースによって異なるため、ここでは扱いません。

一時ディスク領域が必要な条件

IBM SPSS Modeler Server の強力な SQL 最適化機能により、可能なかぎり、処理はサーバー上でなく、データベース内で行われます。ただし、以下の条件が真 (true) の場合は、SQL 最適化を使用できません。

- 処理するデータが、データベースでなくフラット・ファイル内にある。
- SQL 最適化機能がオフにされている。
- 処理中の操作を SQL を使用して最適化できない。

SQL 最適化が使用できない場合は、次のデータ操作ノードと CLEM 関数により、データの一部またはすべての一時的なディスク・コピーが作成されます。サイトで使用されるストリームにこれらの処理コマンドや関数が含まれている場合は、追加のディスク領域をサーバー上に設定する必要がでてくる可能性があります。

- Aggregate ノード
- 重複レコード・ノード
- データ分割ノード
- キー単位の結合オプションを使用する場合のレコード結合ノード
- すべてのモデル作成ノード
- ソート・ノード
- テーブル出力ノード
- @THIS を使用する検索条件がある、@OFFSET 関数
- @MIN、@MAX、および @AVE など、オフセット・パラメーターが計算される @ 関数。

一時ディスク領域の使用量の計算

通常、IBM SPSS Modeler Server では、元のデータ・セットの少なくとも「3 倍の大きさ」の、一時ファイルを書き出すことが必要です。例えば、データ・ファイルが 2GB で SQL 生成が使用されない場合、IBM SPSS Modeler Server では、データを処理するために 6GB のディスク領域が必要です。同時並行ユーザー・アカウントのそれぞれが自分の一時ファイルを作成するので、同時並行のユーザー数に応じて、ディスク領域を増やす必要があります。

頻繁に大規模な一時ファイルを使用するサイトでは、IBM SPSS Modeler の一時ファイル用に、別のディスク上に作成される別個のファイル・システムを使用することを考慮してください。最高の結果を求めるには、RAID 0 または複数の物理ディスクに広がるストライプ化されたデータ・セットが、ディスク操作の高速化に使用できます。理想的には、別々のディスク・コントローラのストライプ化されたファイル・システム内の各ディスクを使用します。

RAM の要件

データベースで実行できないほとんどの操作に対しては、IBM SPSS Modeler Server により、中間結果が一時ファイルとして、メモリー (RAM) にではなく、ディスクに格納されます。ただし、モデル作成ノードの場合は、可能であれば RAM が使用されます。ニューラル、Kohonen、および K-Means ノードでは、大量の RAM が必要です。これらのノードがサイトで頻繁に使用される場合は、サーバーにさらに多くの RAM を搭載することを考慮してください。

一般的に、必要な RAM のバイト数は、次のように見積もることができます。

$$(\text{number_of_records} * \text{number_of_cells_per_record}) * \text{number_of_bytes_per_cell}$$

上記の `number_of_cells_per_record` は、名義型フィールドがある場合に、非常に大きな値になることがあります。

RAM に関する最新の推奨事項については、サーバー・インストール・ガイドのシステム要件に関するセクションを参照してください。4 人以上の同時並行ユーザーがいる場合は、さらに多くの RAM をお勧めします。メモリーは同時並行タスク間で共有する必要があるため、タスクの数に応じて必要量も増えます。通常、メモリーの追加は、全体的なパフォーマンスを向上させるための、おそらくもっとも費用対効果のよい方法の 1 つです。

データへのアクセス

データベースを読み書きするには、ODBC データ・ソースがインストールされていて、該当するデータベースに対して必要に応じて読み取り権限や書き込み権限が設定されている必要があります。IBM SPSS Data Access Pack には、この目的で使用できる ODBC ドライバーが含まれています。また、これらのドライバーは、IBM SPSS Data Access Pack DVD またはダウンロード・サイトから入手できます。ODBC データ・ソースの権限の作成または設定についてわからないことがある場合は、データベース管理者に問い合わせてください。

サポートされている ODBC ドライバー

IBM SPSS Modeler 16 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>) にある製品互換性マトリクスを参照してください。

ドライバーをインストールする場所

ODBC ドライバーは、処理が行われる各コンピューターにインストールして設定する必要があることに注意してください。

- ローカル (スタンドアロン) モードで IBM SPSS Modeler を実行する場合は、ドライバーをローカル・コンピューターにインストールする必要があります。
- IBM SPSS Modeler をリモートの IBM SPSS Modeler Server に対して分散モードで実行する場合、ODBC ドライバーは IBM SPSS Modeler Server がインストールされたコンピューターにインストールす

る必要があります。UNIX システムの IBM SPSS Modeler Server を使用している場合は、このセクションの『UNIX システムの ODBC ドライバーの設定』も参照してください。

- IBM SPSS Modeler と IBM SPSS Modeler Server の両方から同じデータ・ソースにアクセスする必要がある場合、ODBC ドライバーは両方のコンピューターにインストールする必要があります。
- 端末サービスを介して IBM SPSS Modeler を実行する場合、ODBC ドライバーは IBM SPSS Modeler がインストールされた端末サービス・サーバーにインストールする必要があります。

重要: UNIX 上の IBM SPSS Modeler Server を使用して Teradata データベースにアクセスし、かつ Teradata 付属の ODBC ドライバーを使用する場合、Teradata ODBC ドライバーとともにインストールされた ODBC ドライバー・マネージャーを使用する必要があります。(注: SDAP Teradata ドライバーを使用する場合は、以下の変更を行う必要はありません。) この変更を IBM SPSS Modeler Server に対して行うには、コメントで示されている `modelersrv.sh` スクリプトの最上位近くの `ODBC_DRIVER_MANAGER_PATH` に値を指定してください。この環境変数は、Teradata ODBC ドライバーに付属する ODBC ドライバ マネージャの場所 (Teradata ODBC ドライバーのデフォルト・インストールの `/usr/odbc/lib`) に設定する必要があります。変更を有効にするには、IBM SPSS Modeler Server を再起動する必要があります。Teradata へのアクセスをサポートする IBM SPSS Modeler Server プラットフォーム、およびサポートされている Teradata ODBC ドライバー・バージョンの詳細は、当社サポート・サイト <http://www.ibm.com/support> を参照してください。

UNIX システムの ODBC ドライバーの設定

デフォルトでは、DataDirect Driver Manager は UNIX システムの IBM SPSS Modeler Server 向けには設定されていません。DataDirect Driver Manager ロードするよう UNIX を設定するには、次のコマンドを入力します。

```
cd modeler_server_install_directory/bin
rm -f libspssodbc.so
ln -s libspssodbc_datadirect.so libspssodbc.so
```

これにより、デフォルトのリンクが削除され、DataDirect Driver Manager へのリンクを作成します。

注：上記のルールは、特にデータベース内のデータへのアクセスに適用されます。ストリーム、プロジェクト、モデル、ノード、PMML、出力、およびスクリプト・ファイルの展開や保存などのその他のファイル操作は、常にクライアント側で行われ、クライアント・コンピューターのファイル・システムに常に関連して指定されます。さらに、IBM SPSS Modeler 内の「ディレクトリーの設定」コマンドで「ローカル」なクライアント・オブジェクト (ストリームなど) の作業ディレクトリーが設定されますが、このコマンドでサーバーの作業ディレクトリーが影響を受けるわけではありません。

UNIX。 UNIX で IBM SPSS Statistics データ・アクセス・テクノロジーと連動するように IBM SPSS Modeler Server を構成する方法については、55 ページの『付録 B. UNIX 起動スクリプトの構成』を参照してください。

データ・ファイルの参照

Windows。 IBM SPSS Modeler Server と同じコンピューターにデータを格納する場合は、サーバー・コンピューターの観点からデータへのパスを指定することをお勧めします (例えば `C:¥ServerData¥Sales 1998.csv`)。ファイルを見つけるのにネットワークが使用されないほうが、パフォーマンスがよくなります。

データが別のホストに格納される場合は、UNC ファイル参照 (例えば `¥¥mydataserver¥ServerData¥Sales 1998.csv`) を使用することをお勧めします。UNC 名は、パスに共有ネットワーク・リソースの名前が含まれ

ている場合にだけ機能します。コンピューターの参照には、指定されたファイルを読み取る権限が必要です。分散モードからローカルの分析モードへ頻繁に切り換える場合は、モードに関係なく機能するので、UNC ファイル参照を使用します。

UNIX。 UNIX サーバーにあるデータ・ファイルを参照するには、完全なファイル指定とスラッシュ (斜線) を使用します (例: `/public/data/ServerData/Sales 1998.csv`)。IBM SPSS Modeler Server とともに使用するデータの UNIX ディレクトリーとファイル名には、円記号 (¥) の使用を避けてください。テキスト・ファイルで UNIX フォーマットと DOS フォーマットのどちらが使用されているかは問題ではありません。どちらのフォーマットも自動的に処理されます。

IBM SPSS Statistics データ・ファイルのインポート

サイトで IBM SPSS Statistics Server も実行している場合は、ユーザーが、分散モード中に IBM SPSS Statistics データをインポートまたはエクスポートできます。IBM SPSS Modeler を分散モードで実行すると、サーバーのファイル・システムが提示されます。IBM SPSS Statistics クライアントも、同じように機能します。2 つのアプリケーション間でインポートとエクスポートが行われるようにするために、両方のクライアントは同じモードで稼動している必要があります。違うモードの場合はファイル・システムのビューが異なり、ファイルを共有することができません。IBM SPSS Modeler の IBM SPSS Statistics ノードでは、IBM SPSS Statistics クライアントを自動的に起動できますが、ユーザーはまず、IBM SPSS Statistics クライアントが IBM SPSS Modeler と同じモードで稼動していることを確認する必要があります。

インストール手順

IBM SPSS Modeler Server のインストールの詳細は、Server CD の `¥documentation¥installation¥` フォルダにある説明を参照してください。別々のドキュメントが、Windows と UNIX で利用できます。

IBM SPSS Modeler のインストールと使用法の完全な説明は、クライアントの CD を参照してください。

第 3 章 IBM SPSS Modeler のサポート

このセクションは、管理者と、IBM SPSS Modeler のユーザーをサポートするヘルプ デスク担当者を対象としています。ここには、以下のトピックが含まれています。

- IBM SPSS Modeler Server へのログイン方法（または、サーバーから接続解除してスタンドアロンで実行する方法）
- ユーザーに必要な可能性のあるデータおよびファイル・システム
- IBM SPSS Modeler Server に関するユーザー・アカウントとファイル・アクセス権限
- IBM SPSS Modeler Server と IBM SPSS Modeler 間の切り換え時に確認されることのある、結果の違い

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・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 18 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 19 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。サーバーへのログオンに使用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・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。

シングル・サインオンの設定

サポートされている任意のプラットフォームで実行中の IBM SPSS Modeler Server に、シングル・サインオンを使用して接続できます。シングル・サインオンを使用して接続するには、IBM SPSS Modeler のサーバー・マシンとクライアント・マシンをまず設定する必要があります。

IBM SPSS Modeler Server と IBM SPSS Collaboration and Deployment Services の両方に接続するためにシングル・サインオンを使用している場合、IBM SPSS Modeler に接続する前に IBM SPSS Collaboration and Deployment Services に接続する必要があります。

注: IBM SPSS Modeler のサーバー・マシンとクライアント・マシンをシングル・サインオン向けに設定する前に、それらのマシンがドメイン・コントローラー・サーバーへのアクセス権限を持っていることを確認する必要があります。

Windows でのシングル・サインオンの設定

IBM SPSS Modeler Server を設定するために、以下を実行します。

1. Windows サーバー・マシンが、Active Directory (AD) ドメインのメンバーであることを確認します。
2. IBM SPSS Modeler Server のインストール場所で、config という名前のフォルダーを見つけます。
3. config フォルダーに、sso という名前のサブフォルダーを作成します。
4. sso フォルダーに、krb5.conf ファイルを作成します。krb5.conf ファイルの作成方法については、http://web.mit.edu/kerberos/krb5-current/doc/admin/conf_files/krb5_conf.html を参照してください。krb5.conf ファイルの例を以下に示します。

```
[libdefaults]
    default_realm = MODELERSSO.COM
    dns_lookup_kdc = true
    dns_lookup_realm = false

[realms]
    MODELERSSO.COM = {
        kdc = <AD server IP>:88
        admin_server = <AD server IP>:749
```

```

        default_domain = MODELERSSO.COM
    }

[domain_realm]
.modelersso.com = MODELERSSO.COM

```

IBM SPSS Modeler Client を設定するために、以下を実行します。

1. IBM SPSS Modeler を実行中のローカルの Windows マシンが、AD ドメインのメンバーであることを確認します。
2. ローカル・マシンで、管理者としてドメイン・ユーザーを追加します。
3. 次のようにして、Windows が TGT セッション・キーにアクセスできるようにします。
 - a. 「スタート」メニューの「ファイル名を指定して実行」をクリックします。
 - b. regedit と入力し、「OK」をクリックして、「レジストリ エディター」を開きます。
 - c. 次のように、ローカル・マシンの該当するオペレーティング・システムのレジストリー場所に移動します。
 - Windows XP: マイ コンピュータ
%HKKEY_LOCAL_MACHINE%System%CurrentControlSet%Control%Lsa%Kerberos
 - Windows Vista または Windows 7: コンピューター
%HKKEY_LOCAL_MACHINE%System%CurrentControlSet%Control%Lsa%Kerberos%Parameters
 - d. フォルダーを右クリックし、「新規」 > 「DWORD 値」を選択します。新しい値の名前は、allowtgtsessionkey にする必要があります。
 - e. allowtgtsessionkey の値を、1 の 16 進値、つまり 0x0000001 に設定します。
 - f. 「レジストリ エディター」を閉じます。
 - g. ローカル・マシンのオペレーティング・システムが Windows Vista または Windows 7 の場合、
<IBM SPSS Modeler installation location>%jre%bin にある kinit.exe を実行する必要があります。
4. IBM SPSS Modeler のインストール場所の config フォルダーに、sso という名前のフォルダーを作成します。
5. krb5.conf ファイルをサーバーから sso フォルダーにコピーします。
6. ローカル・マシンとサーバー・マシンを再起動します。

UNIX でのシングル・サインオンの設定

UNIX サーバー・マシンに対してシングル・サインオンを設定するために、UNIX 追加を Windows AD ドメインに追加し、Windows でシングル・サインオンを設定するための手順に従うことができます。詳しくは、トピック 16 ページの『Windows でのシングル・サインオンの設定』を参照してください。別の方法として、次のステップを実行することもできます。

1. UNIX マシンのドメイン・ユーザー・アカウントを作成します。
2. ホスト名を変更します。RedHat Linux を使用している場合は、/etc/sysconfig/network ファイルを開き、HOSTNAME を <name>.<realm> の形式に変更します。これにより、AD がサーバーの資格情報を見つけることができるようになります。
3. DNS サーバーが UNIX マシンを見つけることができるようにするため、次のいずれかのステップを実行します。
 - %windows%/system32/drivers/etc/hosts ファイルを開き、IP とホストとのマッピングを追加します。例:
192.168.1.102 test.modelersso.com test

または

- 新しい逆引きゾーン・エントリーを追加します。これにより、DNS サーバーに IP とホストとのマッピングが追加されます。

UNIX マシンの DNS エントリーが正しくない場合、DNS サーバーに逆引きエントリーを手動で追加できます。

4. トピック 16 ページの『Windows でのシングル・サインオンの設定』にある IBM SPSS Modeler Client の設定の手順に従います。

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。

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・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問い合わせてください。

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・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 18 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・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。

データとファイル・システム

IBM SPSS Modeler Server を使用しているユーザーは、ネットワーク上のデータ・ファイルやその他のデータ・ソースにアクセスしたり、ネットワーク上にファイルを保存したりすることが必要になる可能性があります。適宜、次の情報が必要になる場合があります。

- **ODBC データ・ソース情報** :ユーザーが、サーバー・コンピューター上で定義された ODBC データ・ソースにアクセスする必要がある場合に、そのデータ・ソースの名前、説明、およびログイン情報（データベース・ログイン ID とパスワードも含めて）が必要になります。
- **データ・ファイル・アクセス** :ユーザーが、サーバー・コンピューターまたはネットワーク上のどこかにあるデータ・ファイルにアクセスする必要がある場合は、そのデータ・ファイルの名前と場所が必要です。
- **保存されたファイルの場所** :IBM SPSS Modeler Server へ接続中にユーザーがデータを保存する場合、サーバー・コンピューターにファイルを保存しようとします。ただし、書き込み保護がされている場所が多くあります。そのような場合、どこにデータ・ファイルを保存すべきかをユーザーに知らせる必要があります(通常、この場所はユーザーのホーム・ディレクトリーです)。

ユーザー認証

IBM SPSS Modeler Server では、サーバーに接続するユーザーを認証するのに、サーバー・コンピューターのオペレーティング・システムを使用します。ユーザーが IBM SPSS Modeler Server へ接続する場合、ユーザーのために実行されるすべての操作は、ユーザーのセキュリティのコンテキスト内で実行されます。データベース・テーブルへのアクセスは、データベース自体のユーザーやパスワード権限に従います。

Windows :Windows の場合、ホスト ネットワークに対する有効なアカウントを持つユーザーは、誰でもログオンできます。デフォルトの認証では、ユーザーは <modeler_server_install>%Tmp ディレクトリーへの変更アクセス権限が必要です。これらの権利がないと、Windows のデフォルトの認証を使用してクライアントから IBM SPSS Modeler Server にログオンすることはできません。

UNIX。UNIX では、IBM SPSS Modeler Server はデフォルトでルートとして実行されると想定されています。これにより、ホスト ネットワーク上の有効なアカウントを持つユーザーは誰でもログオンでき、ユーザーのファイル・アクセスは、自分のファイルとディレクトリーに制限されます。ただし、ルート権限なしに実行するように IBM SPSS Modeler Server を構成できます。この場合、認証に使用されるプライベート・パスワード・データベースを作成する必要があり、すべての IBM SPSS Modeler ユーザーは、1 つの UNIX ユーザー・アカウントを共有する必要があります (その結果、データ・ファイルへのアクセス権も共有します)。詳しくは、トピック 66 ページの『非公開のパスワード・データベースを使用して root 以外のユーザーとして設定』を参照してください。

Solaris、HP-UX、Linux および AIX のプラットフォームでのみ、IBM SPSS Modeler Server は PAM を使用して認証を行います。必要な場合は、IBM SPSS Modeler Server の PAM モジュールを構成するために、サーバー名 modelerserver を使用できます。

ファイルの作成

IBM SPSS Modeler Server でデータにアクセスして処理する場合、そのデータの一時的なコピーをしばしばディスク上に維持する必要があります。一時ファイルに使用されるディスク領域の大きさは、エンド・ユーザーが分析しているデータ・ファイルのサイズと、実行中の分析の種類によって異なります。詳しくは、トピック 10 ページの『一時的なディスク領域と RAM の要件』を参照してください。

UNIX。UNIX バージョンの IBM SPSS Modeler Server では、一時ファイルに対するファイル権限の設定に、UNIX の `umask` コマンドを使用します。サーバーのデフォルトの権限は、上書きできます。詳しくは、トピック 57 ページの『ファイル作成時の権限の制御』を参照してください。

結果の違い

IBM SPSS Modeler と IBM SPSS Modeler Server の両方で分析を実行したユーザーは、結果にわずかな違いがあることがわかる場合があります。この違いは通常、レコードの並び順や丸めの違いが原因で発生します。

レコードの並び順 :ストリームが順序を明示してレコードをソートしないかぎり、レコードが提示される順序は、ローカルで実行されたストリームとサーバーで実行されたストリームでは異なります。さらに、データベース内で実行される操作順序と IBM SPSS Modeler Server 内で実行される操作順序に違いがあることもあります。これらの違いは、レコード集計などのレコードを並べ替える関数を実装するために、それぞれのシステムで異なるアルゴリズムが使用されることが原因です。また、明示した順序付けの操作がない場合は、レコードがデータベースから返される順序を SQL で指定しません。

丸め操作の違い :ローカル・モードで実行される IBM SPSS Modeler では、浮動小数点値の格納に、IBM SPSS Modeler Server とは別の形式が使用されます。丸め操作の違いが原因で、結果は各バージョンで若干異なります。

第 4 章 IBM SPSS Modeler Server の管理

この章では、IBM SPSS Modeler Server の起動と停止、さまざまなサーバー・オプションの構成、およびログ・ファイルの解釈について説明します。サーバーの構成と監視を支援するアプリケーションの IBM SPSS Modeler Administration Console の使用方法も説明します。このコンポーネントのインストール方法は、製品に付属する IBM SPSS Modeler Server のインストール方法を参照してください。

IBM SPSS Modeler Server の開始および停止

IBM SPSS Modeler Server は、Windows 上ではサービスとして、UNIX 上ではデーモン・プロセスとして実行されます。

スケジューリング上の注意：IBM SPSS Modeler Server を停止すると、エンド・ユーザーが切り離され、セッションが終結するため、使用率が低い期間中にサーバーの再起動をスケジューリングするようにします。これが不可能な場合は、サーバーを停止する前に必ずユーザーに通知してください。

Windows 上で起動、停止、またはステータスをチェックする手順

Windows の場合、Windows のコントロール・パネルにある「サービス」ダイアログを使用して、IBM SPSS Modeler Server を制御します。

1. **Windows XP**：Windows の「スタート」メニューを開きます。「設定」→「コントロール・パネル」を選択します。「管理ツール」、次に「サービス」をダブルクリックします。

Windows 2003 または 2008：Windows の「スタート」メニューを開きます。「コントロール・パネル」→「管理ツール」→「サービス」を選択します。

2. 「IBM SPSS Modeler Server <nn.n>」サービスを選択します。これで、Clementine Server のステータスをチェックでき、起動または停止ができ、必要に応じて起動パラメーターを編集できるようになります。

デフォルトでは、サービスが自動的に起動するように構成されます。つまり、サービスを停止すると、コンピューターが再起動されたときに、サービスも自動的に再起動します。この方法で起動されると、サービスはユーザーの介入なしに実行され、サーバー・コンピューターから、サービスに影響を与えることなくログオフできます。

UNIX 上で起動、停止、またはステータスをチェックする手順

UNIX の場合、IBM SPSS Modeler Server インストール・ディレクトリーで *modelersrv.sh* スクリプトを実行して、IBM SPSS Modeler Server を起動または停止します。

1. IBM SPSS Modeler Server のインストール・ディレクトリーに変更します。例えば、UNIX のコマンド・プロンプトで、次のように入力します。

```
cd /usr/modelersrv
```

上の例の *modelersrv* は、IBM SPSS Modeler Server のインストール先ディレクトリーです。

2. サーバーを起動するには、コマンド・プロンプトで次のように入力します。

```
./modelersrv.sh start
```

3. サーバーを停止するには、コマンド・プロンプトで次のように入力します。

```
./modelersrv.sh stop
```

4. IBM SPSS Modeler Server の状態を点検するには、UNIX コマンド・プロンプトで次のように入力します。

```
./modelersrv.sh list
```

これで、出力を調べます。この出力は、UNIX の `ps` コマンドが作成する出力と似ています。このリストの最初のプロセスが IBM SPSS Modeler Server のデーモン・プロセスであり、残りのプロセスが IBM SPSS Modeler の各セッションです。

IBM SPSS Modeler Server のインストール・プログラムには、サーバー・デーモンが起動時に自動的に始動するようにシステムを構成するスクリプト (`auto.sh`) が含まれています。そのスクリプトを実行し、その後サーバーを停止すると、コンピューターが再起動されたときに、サーバー・デーモンも自動的に再起動されます。詳しくは、トピック 56 ページの『IBM SPSS Modeler Server の自動的な起動と停止』を参照してください。

UNIX カーネルの制限

システムのカーネルの制限が、IBM SPSS Modeler Server の操作に対して十分であることを確認する必要があります。データ、メモリー、ファイルの制限は特に重要で、IBM SPSS Modeler Server 環境内では制限なしに設定する必要があります。これを行うには、次のようにします。

1. 次のコマンドを `modelersrv.sh` に追加します。

```
ulimit -d unlimited
```

```
ulimit -m unlimited
```

```
ulimit -f unlimited
```

2. IBM SPSS Modeler Server を再起動します。

非応答性サーバー・プロセス (UNIX システム) の処理

IBM SPSS Modeler Server プロセスは、ブロックされる (コールが返ってこない、返るまでに非常に長い時間がかかる) システム コールまたは ODBC ドライバ コールを実行する状況など、いくつかの理由で非応答となる場合があります。UNIX プロセスがこの状態に入った場合、UNIX の `kill` コマンドを使用してクリーン・アップすることができます (IBM SPSS Modeler クライアントによって起動して中断または IBM SPSS Modeler クライアントが終了しましたが影響はありません)。`kill` コマンドは、通常の `stop` コマンドの代替として設定され、管理者は `modelersrv.sh` を使用して適切な `kill` コマンドを容易に発行できます。

利用できない (「ゾンビ」) サーバー・プロセスの累積による影響を受けやすいシステムでは、以下の一連のコマンドを使用して、IBM SPSS Modeler Server を一定の間隔で停止および再起動することをお勧めします。

```
cd modeler_server_install_directory  
./modelersrv.sh stop  
./modelersrv.sh kill
```

`modelersrv.sh kill` コマンドを使用して終了した IBM SPSS Modeler プロセスにより (一時ディレクトリーの) 一時ファイルは残り、手動で削除する必要があります。一時ファイルは、リソースの消耗によるアプリケーションのクラッシュ、ユーザーによる中断、システムのクラッシュやその他の理由など、別の状況でも残る場合があります。そのため、IBM SPSS Modeler Server の一定間隔における再起動プロセスの一環として、IBM SPSS Modeler 一時ディレクトリーの残りすべてのファイルを削除することをお勧めします。

すべてのサーバー・プロセスが終了し、一時ファイルが削除されると、IBM SPSS Modeler Server を安全に再起動できます。

管理

IBM SPSS Modeler Server には、その動作を制御する多数の構成可能なオプションがあります。このようなオプションは、次の 2 つの方法で設定できます。

- 現在の IBM SPSS Modeler が無料で使用できる、IBM SPSS Modeler Administration Console アプリケーションを使用します。詳しくは、トピック『IBM SPSS Modeler Administration Consoleの使用』を参照してください。
- [サーバーのインストール先へのパス]/config ディレクトリーにある *options.cfg* テキスト・ファイルを使用します。詳しくは、トピック 30 ページの『options.cfg ファイルの使用』を参照してください。

options.cfg ファイルを編集するよりも、管理ツールとして IBM SPSS Modeler Administration Console をインストールして使用することをお勧めします。ファイルの編集では IBM SPSS Modeler Server のファイル・システムにアクセスすることが必要ですが、IBM SPSS Modeler Administration Console では、これらのオプションを調整するために、ユーザー・アカウントがあれば誰でも認証できるようになりました。また、IBM SPSS Modeler Administration Console では、使用率やパフォーマンスを監視できるようにして、サーバー・プロセスについての追加情報が提供されます。構成ファイルの編集と異なり、ほとんどの構成オプションが、IBM SPSS Modeler Server を再起動しなくても変更できます。

IBM SPSS Modeler Administration Console と *options.cfg* ファイルの使用法の詳細は、次のセクションを参照してください。

IBM SPSS Modeler Administration Consoleの使用

Modeler Administration Console アプリケーションには、SPSS Modeler Server のインストールを監視、構成するコンソール・ユーザー・インターフェースが用意されており、しかも、現在の SPSS Modeler Server のお客様には無料で提供されます。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。

Modeler Administration Console で利用できる多くのオプションは、*options.cfg* ファイルで指定することができ、SPSS Modeler Server インストール・ディレクトリー内の */config* に置かれます。ただし、Modeler Administration Console アプリケーションは、複数のサーバーに接続、構築、モニターリングできる共有のグラフィカル・インターフェースを提供します。

Modeler Administration Console の開始

Windows の「スタート」メニューから、「すべてのプログラム」を選択し、「IBM SPSS Collaboration and Deployment Services」、「Deployment Manager」、「IBM SPSS Modeler Administration Console <nn.n>」の順に選択します。

アプリケーションを初めて実行する場合、「サーバー管理」パネルおよび「プロパティ」パネルが空白の状態が表示されます (IBM SPSS Collaboration and Deployment Services サーバー接続をあらかじめ設定した状態で Deployment Manager をインストールした場合を除く)。Modeler Administration Console を構成した後、左側の「サーバー管理者」ウィンドウ枠に、管理対象の各 SPSS Modeler Server のノードが表示されます。右側のウィンドウ枠には、選択したサーバーの構成オプションが表示されます。管理対象の各サーバーごとに、まず接続をセットアップする必要があります。

Modeler Administration Console でのアクセスの構成

Modeler Administration Console を通じた SPSS Modeler Server への管理者アクセスは、SPSS Modeler Server のインストール・ディレクトリーの /config にある *options.cfg* ファイルの *administrators* 行によって制御されます。この行はデフォルトでコメント・アウトされているため、この行を編集して特定のユーザーにアクセスできるようにするか、* を使用してすべてのユーザーにアクセスできるようにする必要があります。以下にその例を示します。

```
administrators, "*"
administrators, "jsmith,mjones,achavez"
```

- 各行は *administrators,* で始め、エントリーは二重引用符で囲む必要があります。エントリーでは、大文字小文字を区別します。
- 複数のユーザー ID はコンマで区切ります。
- Windows アカウントの場合、ドメイン名は使用しません。
- アスタリスクは、注意して使用してください。アスタリスクにより、IBM SPSS Modeler Server に対する有効なユーザー・アカウントを持つ誰でも（ほとんどの場合、ネットワーク上のすべての人々）が、ログインして構成オプションを変更できるようになります。

SPSS Modeler Server への接続

管理対象のネットワーク上の各 SPSS Modeler Server サーバーに対して、接続を指定する必要があります。その後、各サーバーにログインする必要があります。サーバー接続は Modeler Administration Console セッションを横断して行うことが推奨されますが、ログインの資格情報はそうではありません。Modeler Administration Console を起動するごとに毎回ログインする必要があります。

サーバー接続をセットアップする手順

1. IBM SPSS Modeler Server サービスが開始していることを確認してください。
2. 「ファイル」メニューから「新規」→「管理サーバーの接続」を選択します。
3. ウィザードの最初のページで、サーバー接続の名前を入力します。この名前は自分専用で、例えば「**自動サーバー**」のように、説明的なものにする必要があります。「種類」が「**管理**」**IBM SPSS Modeler Server** に設定されていることを確認し、「次へ」をクリックします。
4. 2 番目のページで、サーバーのホスト名または IP アドレスを入力します。デフォルトのポート番号を変更する場合は、そのポート番号を入力します。「終了」をクリックします。新しいサーバー接続が、「サーバー管理者」ウィンドウ枠に表示されます。

管理タスクを実行するには、ここでログインする必要があります。

サーバーへログインする手順

1. 「サーバー管理者」ウィンドウ枠で、ログインするサーバーをダブルクリックして選択します。
2. 「ログイン」ダイアログ・ボックスで、資格情報を入力します。(サーバー・ホストに対するユーザー・アカウントを使用します。)「OK」をクリックします。

ログインが失敗し、「**サーバーの管理者権限を取得できません**」というメッセージが表示された場合、管理者アクセス権限が正しく設定されていないと考えられます。詳しくは、トピック『Modeler Administration Console でのアクセスの構成』を参照してください。

ログインが失敗し、「**サーバー「<server>」への接続に失敗しました**」というメッセージが表示された場合は、ユーザー ID とパスワードが正しいことを確認してから、IBM SPSS Modeler Server サービスが実行中であることを確認します。例えば、Windows の場合、「コントロール・パネル」>「管理ツール」>「サ

ービス」に移動し、IBM SPSS Modeler Server のエントリーを確認します。「状態」列に「開始済み」と表示されない場合、画面のこの行を選択して「開始」をクリックし、ログインを再試行します。

IBM SPSS Modeler Server へログインした後、サーバー名の下に、「構成」と「監視」の 2 つのオプションが表示されます。これらのオプションの 1 つをダブルクリックします。

SPSS Modeler Server の構成

「構成」ペインには、SPSS Modeler Server の構成オプションが表示されます。このウィンドウ枠を使用して、必要に応じてオプションを変更します。変更を保存するには、ツールバーの「保存」をクリックします。(ポート番号を変更すると、効力を発揮するためにサーバーの再起動が必要です。)

以下のセクショントピックで、オプションを説明します。各オプションのかつこ内に、*options.cfg* 内の対応する行が示されています。*options.cfg* のみに表示されるオプションについては、この項の最後で説明します。

接続/セッション

最大接続数。(*max_sessions*) 同時に接続できる最大ユーザー数。-1 を設定すると、制限はありません。

Port numberポート番号。(port_number) 受信する SPSS Modeler Server のポート番号。別のアプリケーションがすでにデフォルトを使用している場合は変更します。エンド・ユーザーは、SPSS Modeler Server を使用するためにポート番号を知っている必要があります。

Analytic Server 接続

Analytic Server の URL。(*as_url*) Analytic Server の URL には、サーバーの IP アドレス、ポート、および SPSS Modeler Server インストール済み環境が属しているテナントが含まれます。

パスワードを求めるプロンプト。(*as_prompt_for_password*) SPSS Modeler Server が、Analytic Server で使用されているユーザーおよびパスワードの認証システムと同じ認証システムを使用して構成されている場合、N を指定してください。それ以外の場合には Y を指定してください。

データ・ファイル・アクセス

データ・ファイル・パスのみにアクセスを制限。(*data_files_restricted*) このオプションを「はい」に設定すると、データ・ファイルが、標準のデータ・ディレクトリー、および下記の「データ・ファイル・パス」にリストされたディレクトリーに制限されます。

データ・ファイル・パス。(*data_file_path*) クライアントがデータ・ファイルを読み書きできる追加のディレクトリーのリスト。「データ・ファイル・パスのみにアクセスを制限」がオンにされていない限り、このオプションは無視されます。すべてのパス名には、スラッシュ記号 (斜線) を使用する必要があります。例えば [サーバーのインストール先のパス]/data;c:/data;c:/temp のように、セミコロンを使用して複数のディレクトリーを指定します。

プログラム・ファイル・パスのみにアクセスを制限。(*program_files_restricted*) このオプションを「yes」に設定すると、プログラム・ファイルへのアクセスが、標準の *bin* ディレクトリー、および下記の「プログラム・ファイル・パス」にリストされたディレクトリーに制限されます。

プログラム・ファイル・パス。(*program_file_path*) クライアントがプログラムを実行できる追加のディレクトリーのリスト。「プログラム・ファイル・パスのみへのアクセス制限」がオンにされていない限り、このオプションは無視されます。すべてのパス名には、スラッシュ記号 (斜線) を使用する必要があります。セミコロンを使用して、複数のディレクトリーを指定します。

最大ファイル・サイズ (MB)。(max_file_size) ストリームの実行中に作成される一時データ・ファイルおよびエクスポート・データ・ファイルの最大サイズ (SAS および SPSS Statistics のデータ・ファイルには適用されません)。-1 を設定すると、制限はありません。

一時ディレクトリー。(temp_directory) 一時データ・ファイル (キャッシュ・ファイル) を格納するのに使用されるディレクトリー。理想的には、このディレクトリーは独立した高速のデバイスやコントローラに置かれるべきです。このディレクトリーへのアクセス速度が、パフォーマンスへの著しい影響があるからです。コマンドで区切って、複数の一時ディレクトリーを指定することもできます。複数のディレクトリーは、それぞれ別のディスクに配置する必要があります。最初のディレクトリーが最も頻繁に使用され、追加の各ディレクトリーは、ある種のデータ準備操作 (ソートなど) の実行中に並行処理が行われる場合に、一時的な作業ファイルを格納するのに使用されます。それぞれの実行が一時的な格納用に別々のディスクを交互に使用できるようにすると、パフォーマンスが向上する可能性があります。すべてのパス指定でフォワードスラッシュを使用してください。注：SPSS Modeler Server の起動時、このディレクトリーに一時ファイルが生成されます。このディレクトリーへの必要なアクセス権限があることを確認してください (例えば、一時ディレクトリーが共有ネットワーク・フォルダーである場合)。権限がない場合は、SPSS Modeler Server の始動が失敗します。

Python 実行パス。(python_exe_path) 実行可能ファイル名を含む Python 実行可能ファイルの絶対パス。Python がインストールされた場所に応じて、「プログラム・ファイル・パスのみにアクセスを制限」を「いいえ」に設定する必要がある場合があります。

パフォーマンス/最適化

ストリームの再書き込み。(stream_rewriting_enabled) サーバーがストリームを再書き込みすることで、ストリームを最適化できるようにします。例えば、サーバーは、データセットのサイズをできるだけ小さくするために、入力ノードの近くにデータ削減の操作をプッシュできる場合があります。このオプションの無効化は通常、最適化がエラーや予期しない結果の原因になる場合にだけ、推奨されます。この設定は、対応するクライアント最適化設定より優先されます。

並列処理。(max_parallelism) ストリームの実行中に SPSS Modeler で使用できる、並行ワーカー・スレッドの数を記述します。これを 0 または任意の負数に設定すると、IBM SPSS Modeler は、スレッド数をコンピューター上の使用可能なプロセッサ数と一致させます。このオプションのデフォルト値は -1 です。並列処理を無効にするには (複数のプロセッサを備えたマシンの場合)、このオプションを 1 に設定します。並列処理を制限付きで許可するには、マシン上のプロセッサ数よりも少ない数に設定します。ハイパースレッドまたはデュアルコア プロセッサは、2 つのプロセッサと見なされます。

バッファ・サイズ (バイト)。(io_buffer_size) サーバーからクライアントに転送されるデータ・ファイルが通過するバッファのサイズ (バイト数)。

キャッシュの圧縮。(cache_compression) サーバーの一時ディレクトリーにあるキャッシュやその他のファイルの圧縮を制御する、0 から 9 までの整数値。圧縮により使用されるディスク領域の量が削減され、このことは、領域が制限されているときには重要になります。また、キャッシュの作成と読み込みに必要なディスク・アクティビティの量が減るので、パフォーマンスが向上します。圧縮によってプロセッサ時間が増えますが、ほとんど常に、ディスク・アクセス時間の軽減がそれを埋め合わせます。シーケンシャルにアクセスされる特定のキャッシュだけが圧縮可能です。このオプションは、ネットワーク学習アルゴリズムに使用されるような、ランダム アクセスのキャッシュには適用されません。値を 0 にすると、圧縮が完全に無効になります。1 から上の値で圧縮の程度が示されますが、アクセス時間のコストが対応して増えます。デフォルト値は 1 です。これより高い値が必要になることは、ディスク・スペースが限られている場合を除いてほとんどありません。

メモリー使用量の乗数。(memory_usage) ソートやその他のメモリー内キャッシュに割り振られる物理メモリーの割合を制御します。デフォルトは、物理メモリーの約 10% に相当する 100 です。空きメモリーがある場合は、この値を高くするとソート・パフォーマンスが向上しますが、あまりに高くすると過度のページングの原因となるため、注意が必要です。

モデル作成メモリー制限率。(modelling_memory_limit_percentage) ニューラル・ネット、Kohonen、および K-means の各モデルの学習に割り振られる物理メモリーの割合を制御します。デフォルトは 25% です。空きメモリーがある場合は、この値を高くすると学習パフォーマンスが向上しますが、あまりに高くすると、データがディスクにあふれる場合に過度のページングが発生する原因となるため、注意が必要です。

モデル作成メモリーの指定変更を許可。(allow_modelling_memory_override) 特定のモデル作成ノードで「速度の最適化」オプションを有効化または無効化します。デフォルトは使用可能です。このオプションにより、モデル作成アルゴリズムは、利用できるメモリーの割合のオプションを無視して、利用可能なすべてのメモリーを要求します。サーバー・マシンのメモリー・リソースを共有する必要がある場合は、このオプションを無効にすることをお勧めします。

最大および最小サーバー・ポート。(max_server_port および min_server_port) インタラクティブ・モデルおよびストリームの実行に必要なクライアントとサーバー間の追加のソケット接続に使用できるポート番号の範囲を指定します。これらの場合は、ファイアウォールのあるシステム上のユーザーに問題を引き起こす可能性がある範囲を制限せずに、サーバーが他のポートをリスンすることが必要です。どちらもデフォルト値は -1 で、これは「制限なし」を意味します。例えば、サーバーがポート 8000 以上で listen するように設定するには、min_server_port を 8000、max_server_port を -1 に設定します。

メイン・サーバ ポートに追加のポートを開いてストリームを開くまたは実行する必要があります。また、並行ストリームを開くまたは実行する必要がある場合、それに応じてより多くのポートを開く必要があります。これは、ストリーム実行からフィードバックを取得するために必要です。

デフォルトでは、IBM SPSS Modeler では使用できる開いたポートを使用します。そのようなポートが検出されない場合 (例えばファイアウォールですべて閉じられている場合)、ストリーム実行時にエラーが表示されます。ポートの範囲を設定するには、IBM SPSS Modeler では、並行ストリームごとに 2 つの開いたポート (メインのサーバー・ポートに加えて)、ODBC 接続を持つ接続された各クライアントの追加ポートが必要です。より多くの並行ストリームを実行するには、ポート数の範囲を拡張し、ファイアウォールにより多くのポートを開きます。

注：これらのパラメーターを変更する場合、変更を有効にするためには SPSS Modeler Server を再起動する必要があります。

配列フェッチの最適化。(sql_row_array_size) SPSS Modeler Server が ODBC データ・ソースからデータをフェッチする方法を制御します。デフォルト値は 1 で、一度に 1 行をフェッチします。この値を増加した場合は、指定された数の行を配列にフェッチして、サーバーは情報を大きな単位で読み取ります。一部のオペレーティング・システム/データベースでは、こうすることで SELECT 文のパフォーマンスが改善します。

SQL

SQL 文字列の最大長。(max_sql_string_length) SQL を使用してデータベースからインポートされたストリングのうち、正常に渡されることが保証される最大文字数。OS に応じて、この値より長い文字列値は、警告なく切り捨てられます。有効な範囲は 1 文字から 65,535 文字です。このプロパティーはデータベース・エクスポート・ノードにも適用されます。

自動 SQL 生成。 (sql_generation_enabled) ストリームの自動 SQL 生成を可能にします。これにより、パフォーマンスが大幅に向上する場合があります。デフォルトは使用可能です。データベースが SPSS Modeler Server から送信されたクエリーをサポートできない場合にだけ、このオプションの無効化をお勧めします。この設定で、対応するクライアントの最適化設定が上書きされることに注意してください。さらに、スコアリングの目的の場合、SQL 生成は、この設定を無視してモデル作成の各ノードに対して個別に有効化する必要があることにも注意してください。

SQL 文字列のデフォルト長。 (default_sql_string_length)。データベース・キャッシュ テーブル内に作成される文字列の列のデフォルト幅を指定します。上流にデータ型情報がない場合、データベース・キャッシュ テーブル内の文字列フィールドはデフォルト幅の 255 で作成されます。データ内にこの値よりも幅広い値がある場合、上流のデータ型ノードをこれらの値でインスタンス化するか、このパラメーターをこれらの文字列値に対応した値に設定します。

SQL タイプのロギングのサポート。 (supported_sql_type_logging)。テクニカルサポートの指示がある場合の除き、変更しないでください。

データベース UDF の有効化。 (db_udf_enabled)。 Y (デフォルト) に設定されている場合、SQL 生成オプションが純粋な SPSS Modeler SQL の代わりにユーザー定義関数 (UDF) SQL を生成します。UDF SQL は通常、純粋な SQL より優れています。

SSL

SSL の有効化。 (ssl_enabled) SPSS Modeler と SPSS Modeler Server の間の接続で SSL 暗号化を有効にします。

証明書ファイル。 (ssl_certificate_file) サーバーの始動時にロードされる SSL 公開証明書ファイル。

秘密鍵ファイル。 (ssl_private_key_file) サーバーの始動時にロードされる SSL 秘密鍵ファイル。

秘密鍵のパスワード。 (ssl_private_key_password) 指定された証明書のパスワード・キー。このキーは、構成済みの証明書ファイル内に一覧にされた認証機関の少なくとも 1 つにより署名されている必要があります。

Coordinator of Processes の構成

ホスト。 (cop_host) Coordinator of Processes サービス のホスト名 および IP アドレス。デフォルトの「spsscops」は、管理者が DNS の IBM SPSS Collaboration and Deployment Services ホストのエイリアスとして追加できる仮の名前です。

Port numberポート番号。 (cop_port_number) Coordinator of Processes サービスのポート番号。デフォルトの 8080 は IBM SPSS Collaboration and Deployment Services デフォルトです。

コンテキスト・ルート。 (cop_context_root) Coordinator of Processes サービスの URL。

ログイン名。 (cop_user_name) Coordinator of Processes サービスに対する認証に使用するユーザー名。これは IBM SPSS Collaboration and Deployment Services ログイン名で、セキュリティー・プロバイダーの接頭辞が含まれます (例: ad/jsmith)。

パスワード。 (cop_password) Coordinator of Processes サービスに対する認証に使用するパスワード。

有効。 (cop_enabled) サーバーが Coordinator of Processes への登録を試行するかどうかを決定します。管理者は Coordinator of Processes を介して公開するサービスを選択するため、デフォルトは登録しないです。

サーバー名。 (cop_service_name) この SPSS Modeler Server インスタンスの名前。デフォルトはホスト名です。

説明。 (cop_service_description) このインスタンスの説明。

更新間隔 (分)。 (cop_update_interval) キープアライブ・メッセージ間の分数。デフォルトは 2 です。

重み。 (cop_service_weight) このインスタンスの重み。1 から 10 までの整数で指定します。重みが大きいほど、接続数は増えます。デフォルトは、1 です。

サービス・ホスト。 (cop_service_host) IBM SPSS Modeler Server ホストの完全修飾ホスト名。デフォルトのホスト名は自動的に作成されます。管理者は、マルチホームのホストを上書きすることができます。

options.cfg で表示されるオプション

administrators: 管理者アクセスを付与するユーザーのユーザー名を指定します。詳しくは、トピック 24 ページの『Modeler Administration Console でのアクセスの構成』を参照してください。

allow_config_custom_overrides: 技術サポート担当者の指示がある場合を除き、変更しないでください。

fips_encryption: FIPS に準拠した暗号化を使用可能にします。デフォルトは N です。

max_transfer_size: 変更しないでください。

shell: (UNIX サーバーのみ) UNIX シェルのデフォルト設定をオーバーライドします (例: shell, "/usr/bin/ksh")。デフォルトでは、IBM SPSS Modeler は IBM SPSS Modeler Server に接続しているユーザーのユーザー・プロファイルで定義されているシェルを使用します。

start_process_as_login_user: シングル・サインオンを使用し、IBM SPSS Modeler Server ホストと同じ認証方法を使用しない IBM SPSS Collaboration and Deployment Services サーバーに接続する場合、Y に設定します。また、プライベート・パスワード・データベースを使用して SPSS Modeler Server を実行している場合、Y に設定します。

time_ecode_execution_log: *message.log* ファイル (SPSS Modeler Server インストールの log フォルダー) にストリームの各ノードの実行時間が個別に書き込まれるようにする場合、Y に設定します。

use_bigint_for_count: カウントされるレコードの数が、通常の整数 (2³²-1) での上限を超えている場合、このオプションを Y に設定します。ストリームが DB2 または SQL Server に接続されている場合、このオプションを Y に設定すると、レコード件数が必要な場合には (例えば、レコード集計ノードによって生成された「Record_Count」フィールドで) COUNT_BIG() が使用されます。

SPSS Modeler Server の監視

Modeler Administration Console の監視用のウィンドウ枠には、Windows Task Manager と類似した、SPSS Modeler Server コンピューターで実行されているすべてのプロセスのスナップショットが表示されます。監視用ウィンドウ枠を活動状態にするには、「サーバー管理者」ウィンドウ枠の該当するサーバーの下にある、監視ノードをダブルクリックします。これで、サーバーからの現在のデータのスナップショットがウィンドウ枠に表示されます。データは、表示された間隔で更新されます (デフォルトは 1 分ごと)。データを手動で更新するには、「リフレッシュ」 ボタンをクリックします。このリストに SPSS Modeler Server プロセスだけを表示するには、「SPSS Modeler 以外のプロセスを除外」 ボタンをクリックします。

options.cfg ファイルの使用

options.cfg ファイルは、[サーバーのインストール先へのパス]/*config* ディレクトリーにあります。設定は、コンマで区切られた名前と値のペアで表現されます。この場合、**名前**はオプションの名前、**値**はオプションの**値**です。ポンド (ハッシュ) 記号 (#) で、コメントを示します。

注：この構成ファイルではなく、IBM SPSS Modeler Administration Console を使用して構成オプションを変更できますが、いくつかの例外があります。詳しくは、トピック 29 ページの『options.cfg で表示されるオプション』を参照してください。

IBM SPSS Modeler Administration Console を使用すると、サーバー・ポート以外のすべてのオプションは、サーバーを再起動しなくても変更できます。詳しくは、トピック 23 ページの『IBM SPSS Modeler Administration Consoleの使用』を参照してください。

デフォルト・ファイルに追加可能な構成オプション

デフォルトでは、データベース内キャッシングが IBM SPSS Modeler Server で有効化されています。次の行を *options.cfg* ファイルに追加して、この機能を無効にできます。

```
enable_database_caching, N
```

これにより、一時ファイルがデータベース内ではなくサーバー上に作成されます。

IBM SPSS Modeler Server の構成オプションを表示または変更する手順:

1. テキスト・エディターを使用して、*options.cfg* ファイルを開きます。
2. 変更対象のオプションを見つけます。すべてのオプションの一覧は、25 ページの『SPSS Modeler Server の構成』を参照してください。
3. 必要に応じて値を編集します。パラメーターの区切り記号として、すべてのパス名値には円記号でなく斜線 (/) を使用する必要があります。
4. ファイルを保存します。
5. 変更が効力を発揮するように、IBM SPSS Modeler Server を停止し、その後再起動します。詳しくは、トピック 21 ページの『IBM SPSS Modeler Server の開始および停止』を参照してください。

使用されていないデータベース接続の終了

デフォルトで、IBM SPSS Modeler は接続にアクセスされるとデータベースへの接続を少なくとも 1 つキャッシュします。データベース・アクセスが必要なストリームが実行されていない場合でも、データベース・セッションが開かれます。

データベース接続をキャッシュすると、ストリームが実行されるごとに IBM SPSS Modeler がデータベースに再接続する必要がなくなるため、実行時間を短縮できます。ただし、一部の環境では、アプリケーションができるだけ迅速にデータベース・リソースをリリースすることが重要です。使用されていないデータベースへの接続を維持する IBM SPSS Modeler セッションが多すぎる場合、データベース・リソースが消耗する場合があります。

カスタム・データベース構成ファイルの IBM SPSS Modeler オプション *cache_connection* をオフにして問題を回避できます。このオプションをオフにすることによって、IBM SPSS Modeler セッションで長期間接続を使用している場合に発生する可能性があるデータベース接続での障害 (タイムアウトなど) に対する影響を IBM SPSS Modeler が受けないようにすることもできます。

使用されていないデータベース接続を終了するには

1. [サーバーのインストール先へのパス]/config ディレクトリーを検索します。
2. 次のファイルを追加します (すでにある場合は、そのファイルを開きます)。

`odbc-custom-properties.cfg`

3. 次の行をファイルに追加します。

`cache_connection, N`

4. ファイルを保存してクローズします。
5. IBM SPSS Modeler Server を起動して、変更が有効になるようにします。

安全にデータを転送するための SSL の使用

Secure Sockets Layer (SSL) は、2 つのコンピューター間で転送するデータを暗号化するためのプロトコルです。SSL を使用することにより、コンピューター間の通信が安全なものとなります。SSL を使用して、ユーザー名/パスワードの認証およびサーバー/クライアント間で交換する内容を暗号化することができます。

SSL の動作方法

SSL は、サーバーの ID をパブリック キーにバインドするパブリック キー証明書のほか、サーバーのパブリック キーとプライベート・キーに依存します。

1. クライアントがサーバーに接続する場合、クライアントはパブリック キー証明書を使用してサーバーを認証します。
2. クライアントは乱数を生成し、サーバーのパブリック キーを使用して番号を暗号化し、暗号化されたメッセージをサーバーに送信します。
3. サーバーはプライベート・キーを使用して乱数を復号化します。
4. 乱数から、サーバーとクライアントは、後続する情報の暗号化および復号化に使用するセッション キーを作成します。

パブリック キー証明書は通常、認証期間によって署名されています。VeriSign や Thawte などの認証期間は、パブリック キー証明書に含まれるセキュリティー資格情報を発行、認証、および管理する組織です。認証期間は、サーバーの ID を確認します。認証期間は通常、証明書に課金しますが、自己署名証明書を作成することもできます。

SSL を使用したクライアント-サーバー間およびサーバー/サーバー間通信の保護

SSL を使用してクライアント/サーバー通信およびサーバー/サーバー通信を保護するための主なステップは、以下のとおりです。

1. サーバー・コンピューターに OpenSSL をインストールします。
2. SSL 証明書とキーを取得およびインストールします。
3. サーバー管理アプリケーション (IBM SPSS Collaboration and Deployment Services Deployment Manager) で、SSL の有効化と構成を行います。
4. 必要に応じて、クライアント・コンピューターに無限強度の暗号化をインストールします。
5. 証明書ファイルをクライアント・キーストアに追加します。
6. ユーザーにサーバー接続時に SSL を有効化するよう指示します。

注: サーバー製品がクライアントとして機能する場合があります。例えば、IBM SPSS Collaboration and Deployment Services Repository に接続する IBM SPSS Statistics Server などです。この場合、IBM SPSS Statistics Server は、クライアント となります。

OpenSSL のインストール

OpenSSL がサーバーでまだ使用可能でない場合は、OpenSSL をインストールする必要があります。

1. <http://www.openssl.org/> から OpenSSL をダウンロードします。サーバーのバージョンに適したバージョンの OpenSSL を使用してください。

表 1. OpenSSL のバージョン:

サーバー製品	互換性のある OpenSSL のバージョン
IBM SPSS Statistics 20 から 22	1.0.0
IBM SPSS Statistics 17 から 19 (Linux® on System z® 以外) IBM SPSS Modeler 16 (Linux® on System z® 以外)	0.9.8 およびサブバージョン (0.9.8a、0.9.8b、など)
IBM SPSS Statistics 19 (Linux® on System z®) IBM SPSS Modeler 16 (Linux® on System z®)	1.0.0

注 :IBM SPSS Modeler Server for UNIX が SSL に設定されている場合、次のような名前の 2 つのライブラリーとのリンク(.so はプラットフォーム固有の場合があります) として見ます。

- *libssl.so.0.9.8*
- *libcrypto.so.0.9.8*

SSL のサブバージョン (0.9.8a、0.9.8b、など) を使用する場合、IBM SPSS Modeler Server が期待する xxx.0.9.8 バージョンのライブラリーへのシンボリック・リンクを作成する必要があります。

2. 指示に従ってソフトウェアをインストールして構成します。OpenSSL をご自分で構築する場合、次のガイドラインに注意してください。

Windows : OpenSSL は DLL を使用して構築する必要があります (デフォルトではマルチスレッド)。

UNIX。 OpenSSL は、複数のスレッド (必ずしもデフォルトとは限りません) と共有ライブラリーをサポートする必要があります。

3. システムが SSL ライブラリーを置くことができるようにしてください。

Windows : OpenSSL がシステム・パスに含まれていることを確認します。

UNIX。 SSL ライブラリーを *modelersrv.sh* で定義されたロード ライブラリー・パスの末尾に追加します。例えば、Solaris の場合、次の行をサーバー起動スクリプトに追加します。

```
LD_LIBRARY_PATH_64=$LD_LIBRARY_PATH_64:/usr/local/openssl; export LD_LIBRARY_PATH_64
```

SSL 証明書とキーの取得およびインストール

1. SSL 証明書およびキー・ファイルを取得します。それには、2 つの方法があります。

- 公的認証機関 (Verisign、Thwaite など) から証明書およびキー・ファイルを購入します。公的認証機関は、使用するサーバーを認証するための証明書を署名します。
- 内部の自己署名認証局を使用してキーおよび証明書ファイルを生成します。OpenSSL は、この目的のために証明書管理ツールを提供します。または、自己署名 SSL 証明書の作成を行う方法については、インターネットを検索することができます。

2. 証明書およびキー・ファイルを、サーバー上のローカル・ディレクトリーにコピーします。パブリック・キーとプライベート・キーは、別個のディレクトリーに保存できます。単一のファイルに格納することもできます。プライベート・キーを含むディレクトリーは、通常の参照では表示できないようにアクセス権限が設定されていることを確認してください。

IBM SPSS Collaboration and Deployment Services Deployment Manager の SSL の有効化および設定

1. サーバー管理アプリケーション (IBM SPSS Collaboration and Deployment Services Deployment Manager) を起動し、サーバーに接続します。
2. 設定ページで、「**Secure Sockets Layer**」を「はい」に設定します。
3. 「**SSL パブリック キー・ファイル**」で、パブリック キー・ファイルへのフル・パスを指定します。
4. 「**SSL プライベート・キー・ファイル**」で、プライベート・キー・ファイルへのフル・パスを指定します。

注：パブリック キーおよびプライベート・キーを 1 つのファイルに保存する場合、「**SSL パブリック キー・ファイル**」および「**SSL プライベート・キー・ファイル**」で同じファイルを指定します。

5. メニューから次の項目を選択します。

「ファイル」 > 「保存」

6. サーバー・サービスまたはデーモンを再起動します。再起動すると、SSL パスワードを入力するよう指示があります。Windows では、「このパスワードを記憶する」を選択して、パスワードを安全に保存することはできます。このオプションを使用すると、サーバーを起動するごとにパスワードを入力する必要がなくなります。

IBM SPSS Modeler Server での SSL 設定

非公開の SSL キー・ファイルにアクセスするためにパスワードが必要な場合、パスワードの指定方法はオペレーティング・システムによって異なります。

Windows システム

Windows システムでは、次のいずれかの方法でパスワードを指定できます。

- `ssl_private_key_password` オプションを使用して、`options.cfg` ファイルでパスワードを指定します。
- プロンプト・ウィンドウを使用して指定します。`options.cfg` の `ssl_private_key_password` オプションを省略すると、プロンプト・ウィンドウが開始時に表示されます。

UNIX システム

UNIX システムでは、次のいずれかの方法でパスワードを指定できます。

- `ssl_private_key_password` オプションを使用して、`options.cfg` ファイルでパスワードを指定します。
- 起動時にパスワードを要求するスイッチを `modelersrv.sh` ファイルに追加します。

後者のオプションを使用する場合、`modelersrv.sh` ファイルで次の行を検索します。

```
if "$INSTALLEDPATH/$SCLEMDNAME" -server $ARGS; then
```

`-request_ssl_password` スイッチを次のように追加します。

```
if "$INSTALLEDPATH/$SCLEMDNAME" -request_ssl_password -server $ARGS; then
```

無限強度暗号のインストール

製品に付属している Java Runtime Environment では、米国から輸出可能な強度の暗号が有効になっていません。データのセキュリティを強化するために、無限強度暗号にアップグレードすることをお勧めします。

IBM J9

1. IBM.com Webサイトから、ご使用のバージョンの SDK 向け Java Cryptography Extension (JCE) Unlimited Strength Jurisdiction Policy ファイルをダウンロードしてください。
2. 圧縮ファイルにパッケージ化された、制限されていない管轄権ポリシー・ファイルを解凍します。圧縮ファイルには、*US_export_policy.jar* ファイルと *local_policy.jar* ファイルが含まれています。WebSphere Application Server のインストール・システムで、*\$JAVA_HOME/jre/lib/security* ディレクトリーへ進み、ご使用の *US_export_policy.jar* および *local_policy.jar* ファイルをバックアップします。
3. *US_export_policy.jar* ファイルと *local_policy.jar* ファイルをダウンロードした 2 つのファイルに置き換えます。
4. WebSphere Application Server 管理コンソールのセキュリティを有効にします。セル内のすべてのノード・エージェントが前もってアクティブであることを確認してください。詳細は、WebSphere のマニュアルを参照してください。「**セキュリティ > 管理、アプリケーション、インフラストラクチャを保護する**」 のリストから使用できる realm を選択し、「**現在の状態として設定**」 をクリックして、サーバー再起動時にセキュリティが有効になるようにします。
5. 管理コンソールをログオフします。
6. サーバーを停止します。
7. サーバーを再起動します。

Sun Java

1. Sun Java Webサイトから、ご使用のバージョンの SDK 向け Java Cryptography Extension (JCE) Unlimited Strength Jurisdiction Policy ファイルをダウンロードしてください。
2. ダウンロードしたファイルを抽出します。
3. 2 つの *.jar* ファイル、*local_policy.jar* と *US_export_policy.jar* を *<installation folder>/jre/lib/security* にコピーします。*<installation folder>* は製品をインストールしたフォルダーです。

ユーザーによる SSL の有効化

ユーザーがクライアント製品を使用してサーバーに接続する場合、サーバーに接続するために、ダイアログ・ボックスで SSL を有効化する必要があります。ユーザーが必ず該当するチェック・ボックスを選択するようにしてください。例えば、IBM SPSS Modeler で、「**ツール**」メニューの「**リポジトリオプション**」からアクセスする、「**リポジトリ: サーバー**」ダイアログ・ボックスの「**接続を確保 (SSL を使用)**」チェックボックスをオンにする必要があります。

第 5 章 パフォーマンスの概要

データを分析中の実際のパフォーマンスは、サーバーとデータベースの構成からストリーム内の個々のノードの順序まで、多くの要因に影響されます。通常、以下を実行すると、最高のパフォーマンスを得ることができます。

- データは DBMS に格納し、できるかぎり SQL 生成および最適化を使用する。
- 7 ページの『第 2 章 アーキテクチャーとハードウェアに関する推奨事項』に示された推奨事項を満たす、またはそれを超えるハードウェアを使用する。
- クライアントとサーバーのパフォーマンスと最適化の設定が適切に構成されていることを確認する。
SPSS Modeler が SPSS Modeler Server インストールに接続されている場合、サーバーのパフォーマンスを最適化の設定の方がクライアントのそれらの設定より優先されます。
- 最高のパフォーマンスを発揮できるようにストリームを設計する。

これらパフォーマンス要因のそれぞれの詳細は、次のセクションを参照してください。

サーバーのパフォーマンスと最適化の設定

一定の IBM SPSS Modeler Server 設定値は、パフォーマンスを最適化するように構成できます。これらの設定は、IBM SPSS Modeler Server に含まれている IBM SPSS Modeler Administration Console アプリケーションを使用して調整できます。詳しくは、トピック 23 ページの『IBM SPSS Modeler Administration Consoleの使用』を参照してください。

設定値は、IBM SPSS Modeler Administration Console 構成ウィンドウの「パフォーマンスと最適化」の見出しの下でグループ化されています。各設定値は、ほとんどのインストールで最適なパフォーマンスになるように、事前に構成されています。ただし、使用する特定のハードウェア、データ・セットのサイズ、およびストリームの内容に応じて、設定値を調整する必要があります。詳しくは、トピック 26 ページの『パフォーマンス/最適化』を参照してください。

クライアントのパフォーマンスと最適化の設定

クライアントのパフォーマンスと最適化の設定は、「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」タブから利用可能です。これらのオプションを表示するには、クライアントのメニューから、次を選択します。

「ツール」 > 「ストリーム・プロパティ」 > 「オプション」 > 「最適化」

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。

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 15 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 16 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、そのうちの 1 つまたは任意の組み合わせ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、 $\log(\text{salary})$ という式がある場合、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。これを使用して、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。
- **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に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。

データベースの使用率と最適化

データベース・サーバー。可能な場合、データ・マイニング専用のデータベース・インスタンスを作成し、実動サーバーが IBM SPSS Modeler のクエリーから影響を受けないようにします。IBM SPSS Modeler によって生成された SQL ステートメントには、多くのリソースが必要になることがあります。IBM SPSS Modeler Server マシン上の複数のタスクが、同じデータベースで SQL を実行する場合があるためです。

データベース内マイニング :多くのデータベース・ベンダーは、その製品の拡張版としてデータ・マイニングを提供しています。これらの拡張により、データ・マイニング・アクティビティ (モデル構築またはスコアリング) をデータベース・サーバー内、または別個の専用サーバー内で実行できます。IBM SPSS Modeler のデータベース内マイニング機能により、SQL 生成機能を補足および拡張し、ベンダー固有のデータベース拡張を運用することができます。このアプローチを採用することで、IBM SPSS Modeler とデータベース間のデータ転送に潜在する過度の負荷を避けることができる場合があります。データベース・キャッシングで、さらにパフォーマンス上の利点が増えます。詳細は、IBM SPSS Modeler インストール・ディスクの *Documentation* フォルダにある、*DatabaseMiningGuide.pdf* ファイルを参照してください。

SQL 最適化

最高のパフォーマンスを達成するためには、データベースのパフォーマンスとスケーラビリティを活用するために、生成される SQL の量を常に最大にする必要があります。SQL にコンパイルできないストリームの一部だけを IBM SPSS Modeler Server 内で実行するようにします。詳しくは、39 ページの『第 6 章 SQL 最適化』を参照してください。

ファイル ベースのデータのアップロード

データベースに格納されていないデータは、SQL 最適化の利益を受けることができません。分析するデータがまだデータベース内にはない場合は、データベース出力ノードを使用して、そのようなデータをアップロードできます。また、このノードを使用して、データの準備からの中間データ・セットや、展開の結果を格納することができます。

IBM SPSS Modeler には、多くの一般的なデータベース・システムの外部ローダーとのインターフェースがあります。いくつかの скриプトがソフトウェアに含まれており、ドキュメンテーションと共に IBM SPSS Modeler インストール・フォルダーの下の */scripts* サブディレクトリーにあります。

次の表に、バルク・ロードの潜在的なパフォーマンス上のメリットを示します。数字は、250,000 レコードと 21 フィールドを Oracle データベースへエクスポートする経過時間です。外部ローダーは Oracle の *sqlldr* ユーティリティーです。

表2. バルク・ロードのパフォーマンス上のメリット

エクスポート・オプション	時間 (秒数)
デフォルト (ODBC)	409
ODBC 経由のバルク・ロード	52
外部ローダー経由のバルク・ロード	33

第 6 章 SQL 最適化

IBM SPSS Modeler の最も強力な能力の 1 つは、データベース内で直接多くのデータ準備を行い、マイニング操作を実行する機能です。データベースにプッシュバックして実行可能な SQL コードを生成することにより、サンプリング、ソート、新しいフィールドの作成、およびある種のグラフ化などの多くの処理を、IBM SPSS Modeler または IBM SPSS Modeler Server コンピューターではなく、データベース内で実行できます。大規模なデータ・セットを処理している場合、これらのプッシュバックにより、いくつかある手段で、劇的にパフォーマンスを改善できます。

- DBMS から IBM SPSS Modeler に転送される結果のサイズを小さくする方法。大きな結果のセットが、ODBC ドライバ、ネットワーク I/O またはドライバーを介して読み込まれる場合、効率が低下する可能性があります。この理由のため、SQL 最適化からもっとも利益を受ける操作は、行と列の選択と集計で（条件抽出ノード、サンプリング・ノード、レコード集計ノード）。これらの操作は通常、転送されるデータ・セットのサイズを減らします。データは、将来パフォーマンスを改善するために、ストリーム内の重要な地点（例えば、レコード結合ノードや条件抽出ノードの後）で、データベース内の一時テーブルへキャッシュできます。
- データベースのパフォーマンスとスケーラビリティを活用する方法。効率は、DBMS が並列処理、より高性能なハードウェア、より洗練されたディスク ストレージの管理、およびインデックスの存在などをしばしば活用することで改善されます。

このような長所が与えられているので、IBM SPSS Modeler は各ストリームに生成される SQL の量を最大限にするように設計され、SQL に準拠できない操作だけを IBM SPSS Modeler Server に実行されるようにします。ただし、標準 SQL (SQL-92) で説明されている制限のため、ある種の操作がサポートされないことがあります。詳しくは、トピック 43 ページの『SQL 生成を最大限にするためのヒント』を参照してください。

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。例えば、文字列の比較および文字列の相関関係での大文字小文字の区別に対するデータベース構成によって、SQL プッシュバックを使用して実行する IBM SPSS Modeler ストリームでは、SQL プッシュバックを使用せずに実行したストリームとは異なる結果を生じる場合があります。データベース構成の詳細は、データベース管理者に問い合わせてください。IBM SPSS Modeler との互換性を最大化するには、データベース文字列では大文字小文字の区別をつける必要があります。

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」が表示されません。

詳しくは、トピック 15 ページの『IBM SPSS Modeler Server に接続中』を参照してください。

注：ストリームを Netezza データベースで実行すると、日付と時刻の詳細がデータベースから取得されず。例えば、データベースが異なる国やタイムゾーンにある場合などは、ローカルまたは IBM SPSS Modeler Server の日付および時刻と異なる場合があります。

データベースの要件

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>) にある製品互換性マトリクスを参照してください。

データベース・モデル作成を使用すると、更なるパフォーマンスの向上を獲得できます。

SQL 生成の動作原理

データベース入力ノードから導かれるストリーム内にある最初のものが、SQL 生成の主要な対象物です。SQL ヘコンパイルできないノードがあった場合、データがデータベースから抽出され、以後の処理は IBM SPSS Modeler Server で実行されます。

ストリームの準備中および実行に先立って、SQL の生成プロセスが、次のように行われます。

- サーバーは、それが安全であると証明できる場合は、ストリームを並べ替えて、下流ノードを「SQL ゾーン」に移動します。(この機能は、サーバー上で無効にできます。)
- 入力ノードからターミナル・ノードの方へ向かって、SQL 式が徐々に構築されます。このフェーズは、SQL に変換できないノードに出会うか、またはターミナル・ノード (テーブル・ノード、グラフ・ノードなど) が SQL に変換されると停止します。このフェーズの終わりに、変換の前後で各ノードの SQL が等価な場合、各ノードに SQL ステートメントでラベルが付けられます。
- 最も複雑な SQL 等価物のあるノードから入力ノードへ向かって作業し、SQL の有効性が検査されず。検証に成功した SQL が、実行用に選択されます。
- すべての操作が SQL に生成されたノードは、ストリーム作業領域で紫色に強調表示されます。この結果に基づいて、データベースの実行が完全に有利になるのに適切なストリームをさらに再編できます。詳しくは、トピック 43 ページの『SQL 生成を最大限にするためのヒント』を参照してください。

改善が行われる場所

SQL 最適化によって、多くのデータ操作のパフォーマンスが向上します。

- **キーによる結合**：結合操作で、データベース内の最適化が増加します。
- **集計関数**：レコード集計ノード、棒グラフ・ノード、および Web グラフ・ノードはすべて、結果を作成するために集計を使用します。要約されたデータは、元のデータに比べてかなり低い帯幅を使用します。
- **「選択」**。特定の基準に基づいてレコードを選択することで、レコードの数量が削減されます。
- **ソート**：レコードのソートは、データベース内の方が効率よく実行される、リソース集約型の活動です。
- **フィールドの導出**：新規フィールドは、データベース内の方が効率よく生成されます。
- **フィールドの投影**：IBM SPSS Modeler Server は、以降の処理に必要なフィールドだけをデータベースから抽出します。これにより、必要な帯域幅とメモリーを最小限に抑えることができます。同じことが、フラット・ファイルの余分なフィールドに対しても当てはまります。サーバーは余分なフィールドを読み込む必要がありますが、記憶域は割り当てられません。

- ・ **スコアリング**：SQL は、ディシジョン・ツリー、ルールセット、線形回帰、および因子生成の各モデルから、生成できます。

SQL 生成の例

次のストリームは、キーの操作で 3 つのデータベース・テーブルを結合してから、集計とソートを実行します。

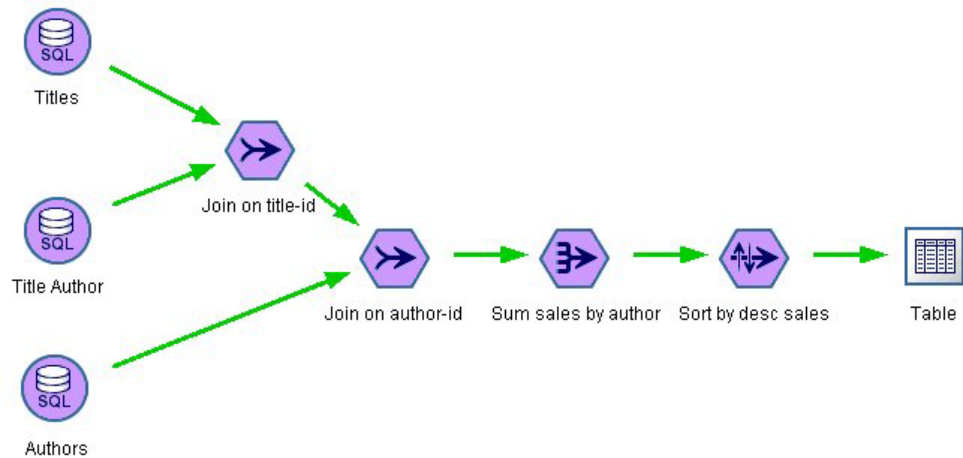


図 3. SQL プッシュバックを意味する紫色のノードを持つ最適化ストリーム (操作はデータベース内で実行されます)

生成された SQL

このストリーム用に生成された SQL は次のとおりです。

```

SELECT
  T2. au_lname AS C0,
  T2. au_fname AS C1,
  SUM({fn CONVERT(T0. ytd_sales ,SQL_BIGINT)}) AS C2
FROM
  dbo . titles T0,
  dbo . titleauthor T1,
  dbo . authors T2
WHERE
  (T0. title_id = T1. title_id )
  AND (T1. au_id = T2. au_id )
GROUP BY T2. au_lname ,T2. au_fname
ORDER BY 3 DESC
  
```

ストリームの実行

ストリームの末端にデータベース出力ノードがある場合、ストリーム全体をデータベース内で実行することが可能です。

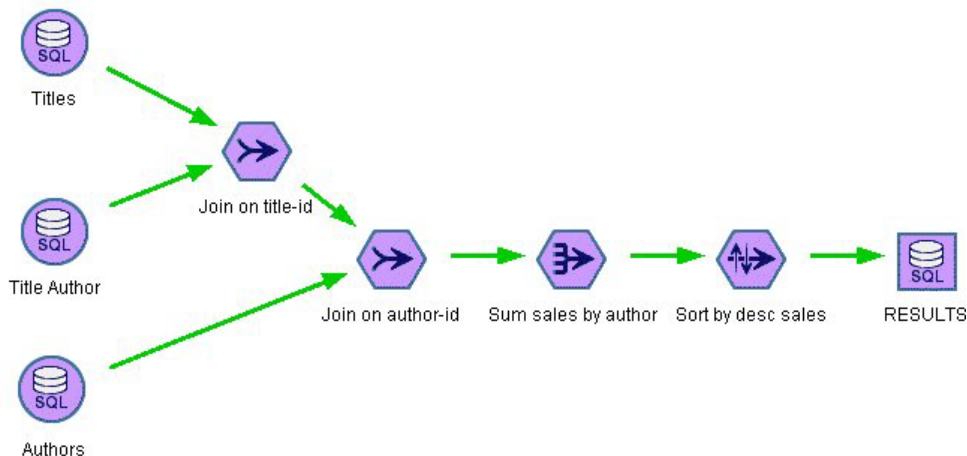


図4. データベース内で実行される全体のストリーム

SQL 最適化の設定

1. 使用するデータベースのために ODBC ドライバーをインストールして、データ・ソースを設定します。詳しくは、トピック 11 ページの『データへのアクセス』を参照してください。
2. データベースからデータを入力するために、入力ノードを使用するストリームを作成します。
3. クライアントと、適用可能な場合はサーバー上で、SQL 生成が有効になっていることを確認します。デフォルトでは、両方に対して有効にされています。

クライアントで SQL 最適化を有効にするには

1. 「ツール」メニューで「ストリームのプロパティ」 > 「オプション」を選択します。
2. 「最適化」タブをクリックします。「SQL 生成」を選択して、SQL 最適化を有効にします。オプションで、パフォーマンスを向上させるその他の設定も選択できます。詳しくは、トピック 35 ページの『クライアントのパフォーマンスと最適化の設定』を参照してください。

サーバーで SQL 最適化を有効にするには

サーバーの設定はクライアントで行われたどのような指定にも優先するので、サーバーの構成設定値の「ストリームのリライト」と「自動 SQL 生成」は、両方ともオンにする必要があります。IBM SPSS Modeler Server の設定を変更する方法の詳細は、26 ページの『パフォーマンス/最適化』を参照してください。

モデルのスコアリング時に最適化を有効にするには

スコアリングが目的の場合、サーバーまたはクライアント・レベルの設定にかかわらず、SQL 生成は各モデル作成ノードごとに個別に有効にする必要があります。常にデータベース内で効率的に評価されない極端に複雑な SQL 式が生成されるモデルもあります。SQL のサイズまたは複雑さにより、生成された SQL を実行しようとする、データベースにエラーが発生する場合があります。

試行とエラーを特定の回数繰り返すと、SQL 生成により指定したモデルのパフォーマンスが改善されているかどうかの確認が必要な場合があります。この作業は、生成されたモデルをストリームに追加した後に、「設定」タブで行います。

生成された SQL のプレビュー

データベース内で実行する前に、メッセージ・ログ内に生成された SQL をプレビューできます。これは、デバッグ時に便利です。また、生成された SQL をエクスポートして、後にデータベース内で編集したり実行したりすることができます。どのノードがデータベースへプッシュバックされるかも示されます。この情報は、パフォーマンス向上のためにストリームを再度順序付けるかどうかの判断に、役に立ちます。

1. 「ユーザー・オプション」ダイアログ・ボックスで「**ストリーム実行中に SQL をメッセージ・ログに表示する**」と「**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**」がオンになっていることを確認します。詳しくは、トピック 35 ページの『クライアントのパフォーマンスと最適化の設定』を参照してください。
2. ストリーム領域で、プレビューするノードまたはストリームを選択します。
3. ツールバーにある「**SQL のプレビュー**」 ボタンをクリックします。

SQL が生成された (そして、ストリームが実行されるときにデータベースへプッシュバックされる) すべてのノードが、ストリーム領域に紫色で表示されます。

4. 生成された SQL をプレビューするには、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」 > 「メッセージ...」

モデル・ナゲットの SQL の表示

一部のモデルについて、モデル・ナゲットの SQL を生成し、モデルのスコアリング段階をデータベースにプッシュバックします。この機能の主な使用はパフォーマンスの改善でなく、これらのナゲットを含むストリームがすべての SQL をプッシュバックできるようにするためです。詳しくは、トピック 44 ページの『SQL 生成をサポートしているノード』を参照してください。

SQL 生成をサポートするモデル・ナゲットの SQL を表示する手順は、次のとおりです。

1. モデル・ナゲットの「設定」タブを選択します。
2. 必要に応じて、「**欠損値のある (ない) サポートを生成**」 または 「**このモデルの SQL を生成**」 のいずれかを選択します。
3. モデル・ナゲットのメニューから、次を選択します。

「ファイル」 > 「SQL のエクスポート」

4. ファイルを保存します。
5. ファイルを開くと SQL が表示されます。

SQL 生成を最大限にするためのヒント

SQL 最適化で最高のパフォーマンスを発揮するには、次の項目に注意を払ってください。

ストリームの順序 : IBM SPSS Modeler のデータ・マイニング機能の方が標準的な SQL にサポートされる伝統的なデータ処理操作より豊かであるため、ノードの関数が意味論的に SQL 内の関数と同等でない場合、SQL 生成が停止する可能性があります。このようなことが発生すると、以後の下流のノードに対しても、SQL 生成が抑制されます。したがって、SQL を停止させる可能性のある操作はできるだけ下流に配置するようにノードを再配置すると、パフォーマンスを著しく向上させることができます。SQL オプティマイザは一定の再配置を自動的にできますが (ストリームのリライトが有効になっていることを確認)、それ以上の改善も可能です。この有力な候補は条件抽出ノードであり、しばしば前方へ移動されます。詳しくは、トピック 44 ページの『SQL 生成をサポートしているノード』を参照してください。

CLEM 式：ストリームを並べ替えることができない場合は、ノード・オプション (CLEM 式) を変更するか、操作の実行方法を変更することで、SQL 生成が阻害されないようにすることができます。レコード作成や条件抽出、および類似したノードは、すべての CLEM 式の演算子が SQL と機能的に同等なので、通常は SQL で表現できます。ほとんどの演算子は表現できますが、SQL 生成を阻害する演算子がいくつかあります (特にシーケンス関数 [“@ functions”])。ときには、生成されたクエリーがデータベースで処理するには複雑過ぎて、生成が停止することもあります。詳しくは、トピック 48 ページの『SQL 生成をサポートする CLEM 式と演算子』を参照してください。

複数の入力ノード：ストリームに複数のデータベース入力ノードがある場合、SQL 生成は、各入力枝葉 (ブランチ) に独立して適用されます。生成がある枝で停止すると、別の枝で続行できます。2 つの枝が併合される (さらに両方の枝は併合用に SQL で表現できる) 場合は、併合自体がしばしばデータベース結合に置き換えられて、生成は下流方向へと続きます。

データベースのアルゴリズム：モデルの推定は、必ずデータベースではなく IBM SPSS Modeler Server で行われますが、Microsoft、IBM、または Oracle のデータベースに組み込まれたアルゴリズムを使用する場合は例外です。

モデルのスコアリング：データベース内のスコアリングは、生成されたモデルを SQL へ変換することで、ある種のモデルに対してはサポートされます。しかし、モデルによっては、データベース内で必ずしも効果的に評価されるとは限らない極端に複雑な SQL 式を生成する場合があります。そのため、それぞれのモデル・ノードで個別に SQL 生成が有効にされる必要があります。あるモデル・ノードが SQL 生成を抑制していることが判明した場合は、そのノードのダイアログ・ボックスで「設定」タブを開き、「このモデルの SQL を生成」を選択します (モデルの種類に応じて、生成を制御する追加のオプションがあります)。このオプションが実行予定のアプリケーションにとって有益であるかを確認するために、テストしてください。詳しくは、トピック『SQL 生成をサポートしているノード』を参照してください。

モデルの SQL 生成が効率的に機能しているかどうかを確認するためにモデル作成ノードをテストする場合、まず IBM SPSS Modeler のすべてのストリームを保存することをお勧めします。複雑な場合もある生成 SQL をよりする場合、データベース・システムが中断する必要があります。Windows タスク マネージャーを使用して IBM SPSS Modeler を終了する必要があります。

データベース・キャッシング：ストリーム内の重要な点 (レコード結合ノードやレコード集計ノードの後など) でデータを保存するためにノード・キャッシュを使用している場合は、SQL 最適化とともにデータベース・キャッシングが有効になっていることを確認してください。ほとんどの場合、この設定によりデータが、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、キャッシュされます。詳しくは、トピック 42 ページの『SQL 最適化の設定』を参照してください。

ベンダー固有の SQL：生成された SQL のほとんどは標準的な SQL-92 ですが、標準的でないベンダー特有の機能が、実用的であれば活用されます。SQL 最適化の程度は、データベース ソースに応じて変化します。

SQL 生成をサポートしているノード

次の表に、SQL 生成をサポートしているデータ・マイニング操作を代行するノードを示します。データベース・モデル作成ノードを除き、ノードにこれらのテーブルが表示されない場合、SQL 生成をサポートしません。

実行する前に生成された SQL をプレビューできます。詳しくは、トピック 43 ページの『生成された SQL のプレビュー』を参照してください。

表 3. 入力

SQL 生成をサポートしているノード	注
データベース	このノードは、以後の分析に使用されるテーブルとビューを指定するのに使用されます。このノードで、SQL クエリーのエントリが有効にされます。結果セットで列名が重複することを回避します。詳しくは、トピック 51 ページの『SQL クエリーの記述』を参照してください。

表 4. レコード設定

SQL 生成をサポートしているノード	注
条件抽出	作成された式で SQL 生成がサポートされている場合のみ、生成をサポートします (次の式を参照)。フィールドにヌルがある場合、ネイティブの IBM SPSS Modeler で与えられるような、SQL 生成では破棄するための結果を与えません。
サンプル	単純なサンプリングは、データベースに応じたさまざまな度合いで SQL 生成をサポートします。表 5 を参照してください。
レコード集計	グループ集計の SQL 生成サポートは、データ・ストレージ・タイプによって異なります。46 ページの表 6 を参照してください。
RFM 集計	2 番目または 3 番目に最近のトランザクションの日付を保存する場合を除いて、または最新のトランザクションを含む場合のみ生成をサポートします。ただし、datetime_date(YEAR,MONTH,DAY) 関数がプッシュバックされる場合は、最新のトランザクションを含んでも機能しません。
ソート	
レコード結合	<p>順序によるレコード結合では、SQL 生成はサポートされません。</p> <p>完全外部結合または部分外部結合のキーによる結合は、データベース/ドライバがサポートしている場合にのみ対応しています。フィルター・ノードまたは入力ノードの「フィルター」タブを使用して、一致しない入力フィールドの名前を変更できます。</p> <p>条件による結合の SQL 生成をサポートしています。</p> <p>すべての結合のタイプで、入力が異なるデータベースで作成される場合は、SQL_SP_EXISTS はサポートされていません。</p>
レコード追加	入力がソートされていない場合のみ生成がサポートされます。
重複レコード	

表 5. 簡単なサンプリングのサンプル・ノードの SQL 生成のサポート :

モード	サンプル	最大サイズ	シード	DB2 for z/OS	DB2 for OS/400	DB2 for Windows/UNIX	Netezza (『注』参照)	Oracle	SQL Server	Teradata
含める	初めの n 件	なし		Y	Y	Y	Y	Y	Y	Y
	n 件ごと	off		Y	Y	Y	Y	Y		Y
		最大		Y	Y	Y	Y	Y		Y
	無作為 %	off	off			Y	Y	Y		Y

表 5. 簡単なサンプリングのサンプル・ノードの SQL 生成のサポート (続き):

モード	サンプル	最大サイズ	シード	DB2 for z/OS	DB2 for OS/400	DB2 for Windows/UNIX	Netezza (『注』参照)	Oracle	SQL Server	Teradata	
			on			Y		Y			
		最大	off			Y	Y	Y		Y	
			on			Y		Y			
破棄	初めの n 件	off					Y	Y			
		最大					Y	Y			
	n 件ごと	off		Y	Y	Y	Y	Y		Y	
		最大		Y	Y	Y	Y	Y		Y	
	無作為 %	off	off	off			Y	Y	Y		Y
			on				Y		Y		
		最大	off				Y	Y	Y		Y
			on				Y		Y		

注: Netezza データベースでは、SQL 生成の実行に使用されるサブクエリーでのキーワードの使用に制限が適用されます。回避策は、サンプル・ノードでキャッシングを有効にすることです。

表 6. レコード集計ノードの SQL 生成サポート:

ストレージ	合計	平均値	最小値	最大値	標準偏差	中央値	度数	分散	パーセンタイル
整数	Y	Y	Y	Y	Y		Y	Y	
実数	Y	Y	Y	Y	Y		Y	Y	
日付			Y	Y			Y		
時間			Y	Y			Y		
タイムスタンプ			Y	Y			Y		
文字列			Y	Y			Y		

表 7. フィールド設定

SQL 生成をサポートしているノード	注
データ型	データ型ノードがインスタンス化されており、ABORT または WARN データ型検査が指定されている場合、生成をサポートします。
フィルター	
フィールド作成	作成された式で SQL 生成がサポートされている場合、SQL 生成をサポートします (次の式を参照)。
アンサンプル	連続型対象の SQL 生成をサポートしています。その他の対象の場合は、「最高確信度勝ち取り」アンサンプル方法が使用されている場合のみ、生成をサポートします。
置換	作成された式で SQL 生成がサポートされている場合、SQL 生成をサポートします (次の式を参照)。

表7. フィールド設定 (続き)

SQL 生成をサポートしているノード	注
匿名化	連続型対象の SQL 生成、名義型、フラグ型対象の部分的 SQL 生成をサポートしています。
データ分類	
データ分割	「分位 (等カウント)」データ分割方法が使用されており、「可能ならピンの値から読み込む」オプションが選択されている場合、SQL 生成をサポートします。
RFM 分析	「可能ならピンの値から読み込む」オプションが選択されている場合 SQL 生成がサポートされますが、下流ノードはサポートしません。
データ区分	SQL 生成をサポートしてレコードをデータ区分に割り当てる :
フラグ設定	
再構成	

表8. グラフ作成(G)

SQL 生成をサポートしているノード	注
グラフボード	SQL 生成は、面、3-D 面、棒、3-D 棒、度数の棒、ヒート・マップ、円、3-D 円、度数の円のグラフ タイプにサポートされています。ヒストグラムの場合、SQL 生成はカテゴリー・データのみサポートされています。
分布	
Web	
評価	

表9. モデル作成

SQL 生成をサポートしているノード	注
自動分類	連続型対象の SQL 生成をサポートしています。その他の対象の場合は、「最高確信度勝ち取り」アンサンブル方法が使用されている場合にのみ、生成をサポートします。
自動数値	

一部のモデルについて、モデル・ナゲットの SQL を生成し、モデルのスコアリング段階をデータベースにプッシュ バックします。この機能の主な使用はパフォーマンスの改善でなく、これらのナゲットを含むストリームがすべての SQL をプッシュバックできるようにするためです。詳しくは、トピック 43 ページの『モデル・ナゲットの SQL の表示』を参照してください。

表10. モデル・ナゲット

SQL 生成をサポートしているモデル・ナゲット	注
C&R Tree	単一ツリー・オプションの SQL 生成はサポートしますが、ブースト、バギング、または大規模データセット・オプションはサポートしていません。
QUEST	
CHAID	
C5.0	

表 10. モデル・ナゲット (続き)

SQL 生成をサポートしているモデル・ナゲット	注
ディシジョン・リスト	
線型	標準モデル・オプションの SQL 生成はサポートしますが、ブースト、バギング、または大規模データセット・オプションはサポートしていません。
ニューラル・ネットワーク	標準モデル・オプションの SQL 生成はサポートしますが (多層パーセプトロンのみ)、ブースト、バギング、または大規模データセット・オプションはサポートしていません。
因子分析	
ロジスティック	多項式手続きの SQL 生成をサポートしていますが、2 項式ではサポートされていません。多項式の場合、確信度が選択されている場合、対象の種類がフラグ型でない限り生成はサポートされません。
生成されたルールセット	

表 11. 出力

SQL 生成をサポートしているノード	注
表	強調表示された式で SQL 生成がサポートされている場合、生成をサポートします (以下の式を参照)。
クロス集計	フィールド・オプションで「すべての数値」が選択されている場合を除いて、生成がサポートされています。
精度分析	選択されているオプションに応じて、生成をサポートします。
変換	
記述統計	「相関関係」オプションが使用されていない場合、生成をサポートします。
レポート	
グローバル	

表 12. エクスポート

SQL 生成をサポートしているノード	注
データベース	
Publisher	発行されたストリームは、生成された SQL を含みます。

SQL 生成をサポートする CLEM 式と演算子

次の表に、SQL の生成をサポートし、データ・マイニング中にしばしば使用される、数学演算と式を示します。この表に存在しない演算は、現在のリリースでは SQL 生成をサポートしません。

表 13. 演算子

SQL 生成をサポートしている操作	注
+	
-	
/	

表 13. 演算子 (続き)

SQL 生成をサポートしている操作	注
*	
><	文字列を結合するために使用されます。

表 14. 関係演算子

SQL 生成をサポートしている操作	注
=	
/=	「等しくない」ことを指定するために使用されます。
>	
>=	
<	
<=	

表 15. 関数

SQL 生成をサポートしている操作	注
abs	
allbutfirst	
allbutlast	
and	
arccos	
arcsin	
arctan	
arctanh	
cos	
div	
exp	
fracof	
hasstartstring	
hassubstring	
integer	
intof	
isaplhacode	
islowercode	
isnumbercode	
isstartstring	
issubstring	
isuppercode	
last	
length	
locchar	
log	

表 15. 関数 (続き)

SQL 生成をサポートしている操作	注
log10	
lowertoupper	
max	
member	
min	
negate	
not	
number	
または	
pi	
real	
rem	
round	
sign	
sin	
sqrt	
文字列	
strmember	
subscrs	
substring	
substring_between	
uppertolower	
to_string	

表 16. 特殊関数

SQL 生成をサポートしている操作	注
@NULL	
@GLOBAL_AVE	特殊グローバル関数は、グローバル・ノードにより計算されたグローバル値を取得するために使用されます。
@GLOBAL_SUM	
@GLOBAL_MAX	
@GLOBAL_MEAN	
@GLOBAL_MIN	
@GLOBALSDEV	

表 17. 集計関数

SQL 生成をサポートしている操作	注
Sum	
Mean	
Min	

表 17. 集計関数 (続き)

SQL 生成をサポートしている操作	注
Max	
Count	
SDev	

CLEM 式での SQL 関数の使用

@SQLFN 関数を使用すると、データベースの実行のみを目的として、指定された SQL 関数を CLEM 式に追加できます。この関数は、専有の SQL または他のベンダー固有のカスタマイズが必要な特別なケースで有用です。

この関数の実行は、IBM Corp. の制御下でない外部データベース・コンポーネントに依存するため、その使用は IBM SPSS Modeler の標準的なサポート契約の対象ではありません。ただし、特殊なケースでは、通常、サービス契約の一部として展開可能です。詳細は、<http://www.ibm.com/software/analytics/spss/services/> に問い合わせてください。

SQL クエリーの記述

データベース・ノードを使用するとき、重複する列名をデータ・セットに生成する SQL クエリーには、特に注意を払ってください。これらの重複する名前は、しばしば、下流のノードの SQL 最適化の妨げになります。

IBM SPSS Modeler は、データベース入力ノードで SQL クエリーを使用するストリームで SQL をプッシュバックするために、入れ子になった SELECT 文を使用します。つまりストリームは、下流のノードの最適化中に生成された SELECT 文の内側に、データベース入力ノードで指定されたクエリーを入れ子にします。そのため、クエリーの結果セットが重複する列名を含む場合、そのステートメントは RDBMS で入れ子にすることができません。入れ子の難しさは、結合されるテーブルのいくつかで同じ名前の列が選択されている場合のテーブル結合中に、しばしば発生します。例えば、入力ノード内の次のクエリーを考察します。

```
SELECT e.ID, e.LAST_NAME, d.*
FROM EMP e RIGHT OUTER JOIN
DEPT d ON e.ID = d.ID;
```

このクエリーは、この SELECT 文は、ID という名前の 2 つの列を持つデータ・セットを生成するため、これ以降の SQL 最適化ができなくなります。

完全な SQL 最適化を可能にするには、SQL クエリーをより明確に記述する必要があり、列名の重複が発生するような状況では、列名のエイリアスを指定する必要があります。次の文は、より明快なクエリーです。

```
SELECT e.ID AS ID1, e.LAST_NAME, d.*
FROM EMP e RIGHT OUTER JOIN
DEPT d ON e.ID = d.ID;
```

付録 A. UNIX プラットフォームでの Oracle の設定

SQL 最適化用の Oracle の設定

UNIX プラットフォームで IBM SPSS Modeler Server を実行し、Oracle データベースから読み込む場合は、次のヒントを検討して、生成された SQL がデータベース内で完全に最適化されることを確認してください。

適切なロケールの指定

Connect ODBC ドライバーと共に提供されているロケール以外のロケールで、IBM SPSS Modeler Server を実行する場合、SQL 最適化を最大限利用するためには、コンピューターを再設定する必要があります。Connect ODBC ドライバーには、*en_US* ロケール・ファイルだけが含まれています。その結果、IBM SPSS Modeler Server コンピューターが、異なるロケールで実行されている場合や、IBM SPSS Modeler Server を開始したシェルのロケールの定義が不完全な場合、生成された SQL を Oracle で完全に最適化できない場合があります。その理由は次の通りです。

- IBM SPSS Modeler Server は、データベースから返されたテキストをテキスト 文字列に変換するために使用するロケールに一致する ODBC ロケール・ファイルを使用します。その後、このテキスト文字列を、どのデータベースに実際に接続するかを決定するために使用します。
- システム \$LANG クエリーにより IBM SPSS Modeler Server に返されたロケールが、*en_US* でない場合、IBM SPSS Modeler は ODBC ドライバから受け取ったコードをテキストに変換できません。つまり、文字列 *Oracle* ではなく、未変換のコードが、データベース接続の開始時に IBM SPSS Modeler Server に返されます。これが、IBM SPSS Modeler が Oracle 用にストリームを最適化できない理由です。

ロケール指定を調べてリセットするには次のようにします。

1. UNIX シェルで、次を実行します。

```
#locale
```

このコマンドは、シェルのロケール情報を返します。以下に例を示します。

```
$ locale
LANG=en_US.ISO8859-15
LC_CTYPE="en_US.ISO8859-15"
LC_NUMERIC="en_US.ISO8859-15"
LC_TIME="en_US.ISO8859-15"
LC_COLLATE="en_US.ISO8859-15"
LC_MONETARY="en_US.ISO8859-15"
LC_MESSAGES="en_US.ISO8859-15"
LC_ALL=en_US.ISO8859-15
```

2. Connect ODBC/locale ディレクトリーに変更します (ここにあるディレクトリーは *en_US* 1 つだけです)。
3. シェル内に設定したロケールの名前を指定して、この *en_US* ディレクトリーにソフト・リンクを作成します。以下に例を示します。

```
#ln -s en_US en_US.ISO8859-15
```

fr_FR.ISO8859-1 などの英語以外のロケールでは、次のようにソフト・リンクを作成してください。

```
#ln -s en_US fr_FR.IS08859-1
```

4. リンクを作成したら、同じシェルから IBM SPSS Modeler Server を再起動します ((IBM SPSS Modeler Server は、起動したシェルからロケールの情報を受け取ります)。

注

UNIX コンピューターを、Oracle に SQL をプッシュバックするように最適化する場合、次のヒントを検討してください。

- 完全なロケールが指定されている必要があります。上の例では、`language_territory.code-page` からのリンクを作成する必要があります。既存の `en_US` ロケール・ディレクトリーでは不十分です。
- データベース内マイニングを完全に最適化するには、`LANG` および `LC_ALL` の両方を、IBM SPSS Modeler Server を開始するのに使用するシェルで定義する必要があります。`LANG` は、IBM SPSS Modeler Server を再起動する前に、他の環境変数を定義するときにシェル内で定義できます。例えば、次の定義を考えてみてください。

```
#LANG=en_US.IS08859-15; export LANG
```

- IBM SPSS Modeler Server を起動するたびに、シェル・ロケール変数が完全に指定されており、ODBC/locale ライブラリーへの適切なソフト・リンクが存在することを確認しなければなりません。

付録 B. UNIX 起動スクリプトの構成

はじめに

この付録では、IBM SPSS Modeler Server の UNIX バージョンに同梱されているスクリプトのいくつかについて、そのようなスクリプトの構成方法について説明します。次の目的で、スクリプトが使用されます。

- サーバー・コンピューターが再起動されたときに自動的に起動するように、IBM SPSS Modeler Server を構成する。
- IBM SPSS Modeler Server を手動で停止し、再起動する。
- IBM SPSS Modeler Server で作成されたファイルに対する権限を変更する。
- IBM SPSS Modeler Server とともに提供された ODBC 接続ドライバーと協調できるように、IBM SPSS Modeler Server を構成する。詳しくは、トピック 57 ページの『IBM SPSS Modeler Server と データ・アクセス・パック』を参照してください。
- Enterprise View ノードでデータを読み込む場合に IBM SPSS Collaboration and Deployment Services Enterprise View Driver を使用できるように、IBM SPSS Modeler Server を構成する。詳しくは、トピック 62 ページの『エンタープライズ・ビュー・ノードのドライバーを構成』を参照してください。

スクリプト

IBM SPSS Modeler Server では、次のようないくつかのスクリプトが使用されます。

- **modelersrv.sh.** IBM SPSS Modeler Server のインストール・ディレクトリーにある、IBM SPSS Modeler Server のための手動起動スクリプトです。このスクリプトで、サーバー・デーモン・プロセスが「手動で」起動されるとき、サーバーの環境を構成します。サーバーを手動で起動および停止するときに、これを実行します。手動による起動の構成を変更する必要がある場合は、このスクリプトを編集します。
- **auto.sh :** これは、起動時にサーバー・デーモン・プロセスを自動的に始動させるためにシステムを構成するスクリプトです。自動的な始動のためにシステムを構成するには、これを 1 回実行します。このスクリプトを編集する必要はありません。このスクリプトは、IBM SPSS Modeler Server のインストール・ディレクトリーにあります。
- **rc.modeler :** *auto.sh* を実行すると、このスクリプトが、使用中のサーバーのオペレーティング・システムに応じた場所に作成されます。このスクリプトで、サーバーが「自動的に」起動されるとき、環境を構成します。自動的な起動の構成を変更する必要がある場合は、このスクリプトを編集します。

表 18. オペレーティング・システムによって異なる *rc.modeler* の場所

オペレーティング・システム	場所
AIX	<i>/etc/rc.modeler</i>
HP-UX	<i>/sbin/init.d/rc.modeler</i>
Solaris	<i>/etc/init.d/rc.modeler</i>

IBM SPSS Modeler Server の自動的な起動と停止

IBM SPSS Modeler Server は、デーモン・プロセスとして起動される必要があります。インストール・プログラムには、IBM SPSS Modeler Server が自動的に停止および再起動するようにシステムを構成できるスクリプト (*auto.sh*) が含まれています。

自動的な起動とシャットダウンのためにシステムを構成する手順

1. root としてログオンします。
2. IBM SPSS Modeler Server のインストール・ディレクトリーに変更します。
3. スクリプトを実行します。UNIX プロンプトで、次のコマンドを入力します。

```
./auto.sh
```

自動的な起動のスクリプトである *rc.modeler* が、上の表に示した場所に作成されます。オペレーティング・システムでは、サーバーが再起動されたときに常に IBM SPSS Modeler Server デーモン・プロセスを始動するために、*rc.modeler* が使用されます。オペレーティング・システムではまた、システムがシャットダウンされるとときに常にデーモンを停止するために、*rc.modeler* が使用されます。

IBM SPSS Modeler Server の手動による起動と停止

modelersrv.sh スクリプトを実行することで、手動で IBM SPSS Modeler Server を起動したり停止したりできます。

IBM SPSS Modeler Server を手動で起動および停止する手順

1. IBM SPSS Modeler Server のインストール・ディレクトリーに変更します。
2. サーバーを起動するには、UNIX のコマンド・プロンプトで次のように入力します。

```
./modelersrv.sh start
```

3. サーバーを停止するには、UNIX コマンド・プロンプトで次のように入力します。

```
./modelersrv.sh stop
```

スクリプトの編集

手動と自動の両方の起動を行う場合は、*modelersrv.sh* と *rc.modeler* の両方を並行して変更します。手動による起動だけを行う場合は、*modelersrv.sh* を変更します。自動的な起動だけを行う場合は、*rc.modeler* を変更します。

スクリプトを編集する手順

1. IBM SPSS Modeler Server を停止します。(詳しくは、トピック『IBM SPSS Modeler Server の手動による起動と停止』を参照してください。)
2. 適切なスクリプトを検索します。(詳しくは、トピック 55 ページの『スクリプト』を参照してください。)
3. テキスト・エディター内でスクリプトを開き、変更を行い、ファイルを保存します。
4. サーバー・コンピューターを自動的に、または手動で再起動して、IBM SPSS Modeler Server を起動します。

ファイル作成時の権限の制御

IBM SPSS Modeler Server では、すべてのユーザーに対する読み込み、書き込み、および実行の権限付きで、一時ファイルが作成されます。このデフォルトは、*modelersrv.sh* と *rc.modeler* のどちらか、または両方の起動スクリプトの *UMASK* 設定を編集して、上書きできます。(詳細は、上記の 56 ページの『スクリプトの編集』を参照してください。) 使用するにはもっとも限定的な *UMASK* 設定として、077 をお勧めします。さらに制約が大きい設定をすると、IBM SPSS Modeler Server の権限問題の原因になる可能性があります。

IBM SPSS Modeler Server と データ・アクセス・パック

ODBC ドライバーを IBM SPSS Modeler Server とともに使用する場合は、IBM SPSS Modeler Server プロセスが始動するときに、ODBC 環境が *odbc.sh* によって構成されている必要があります。このデフォルトは、適切な IBM SPSS Modeler 起動スクリプトの *modelersrv.sh* と *rc.modeler* のどちらか、または両方を編集して、構成できます。(詳しくは、トピック 56 ページの『スクリプトの編集』を参照してください。)

詳細は Technical Support サイト <http://www.ibm.com/support> を参照してください。ODBC データ・ソースの権限の作成または設定についてわからないことがある場合は、データベース管理者に問い合わせてください。

IBM SPSS Modeler Server とともに起動するように ODBC を構成する手順

1. IBM SPSS Modeler Server ホストが稼動している場合は停止します。
2. IBM SPSS Modeler Server をインストールしたプラットフォームの関連する圧縮 tar アーカイブをダウンロードします。インストールした適切なバージョンの IBM SPSS Modeler Server のドライブをダウンロードしてください。ファイルを ODBC ドライバーをインストールする場所 (例: */usr/spss/odbc*) にコピーします。
3. アーカイブを回答します。
4. アーカイブから抽出した *setodbcpath.sh* スクリプトを実行します。
5. スクリプト *odbc.sh* を編集して、下記のようにこのスクリプトの一番下に ODBCINI の定義を追加します。

```
ODBCINI=/usr/spss/odbc/odbc.ini; export ODBCINI
```

ODBCINI は、定義する ODBC データソースのリストを取得するために IBM SPSS Modeler に読み取らせる *odbc.ini* ファイルのフルパス名を示す必要があります (デフォルトの *odbc.ini* はドライバーに付属しています)。

6. *odbc.sh* を保存します。
7. (64 ビットの IBM SPSS Modeler Server インストールのみ。その他のインストールは、次のステップから) *odbc.sh* の *LD_LIBRARY_PATH_64* を定義してエクスポートします。

```
if [ "$LD_LIBRARY_PATH_64" = "" ]; then
    LD_LIBRARY_PATH_64=<library_path>
else
    LD_LIBRARY_PATH_64=<library_path>:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
```

library_path はスクリプトの *LD_LIBRARY_PATH* 定義と同じです。LD_LIBRARY_PATH 定義はインストール・パスで初期化されています (例: */usr/spss/odbc/lib*)。最も簡単な方法は、*odbc.sh* ファイルの *LD_LIBRARY_PATH* の *if* 文と *export* 文をコピーしてファイルの最後に追加し、新しく追加した *if* 文と *export* 文の「LD_LIBRARY_PATH」文字列を「LD_LIBRARY_PATH_64」に置き換えます。

64 ビットの IBM SPSS Modeler Server のインストールの *odbc.sh* ファイルは、最終的に次のようになります。

```
if [ "$LD_LIBRARY_PATH" = "" ]; then
    LD_LIBRARY_PATH=/usr/spss/odbc/lib
else
    LD_LIBRARY_PATH=/usr/spss/odbc/lib:$LD_LIBRARY_PATH
fi
export LD_LIBRARY_PATH
if [ "$LD_LIBRARY_PATH_64" = "" ]; then
    LD_LIBRARY_PATH_64=/usr/spss/odbc/lib
else
    LD_LIBRARY_PATH_64=/usr/spss/odbc/lib:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
ODBCINI=/usr/spss/odbc/odbc.ini; export ODBCINI
```

if ループで LD_LIBRARY_PATH_64 を定義するほか、エクスポートを行ってください。

- 定義した *odbc.ini* ファイルを \$ODBCINI を使用して編集します。必要なデータ・ソース名を定義します (アクセスしているデータベースによって異なります)。
- odbc.ini* ファイルを保存します。
- これらのドライバーを使用するよう IBM SPSS Modeler Server を設定します。設定するには、*modelersrv.sh* を編集して、SCLEMDNAME を定義する行のすぐ下に次の行を追加します。
.*<odbc.sh_path>*

odbc.sh_path は、この手順の始めに編集した *odbc.sh* ファイルへのパスです。例えば、次のようになります。

```
./usr/spss/odbc/odbc.sh
```

注：ここではシンタックスが重要になります。最初のピリオドとファイルへのパスの間にスペースを入れてください。

- modelersrv.sh* を保存します。

接続をテストするには

- IBM SPSS Modeler Server を再起動します。
- クライアントから IBM SPSS Modeler Server に接続します。
- クライアント側で、データベース入力ノードを領域に追加します。
- ノードを開いて、設定手順で *odbc.ini* ファイルで定義したデータ・ソース名が表示されていることを確認します。

この時点で期待されているものが表示されない場合、また定義したデータ・ソースに接続しようとしたときにエラーが発生した場合、トラブルシューティングの手順に従ってください。詳しくは、トピック 59 ページの『ODBC 設定のトラブルシューティング』を参照してください。

IBM SPSS Modeler Solution Publisher Runtime とともに起動するように ODBC を構成する手順

IBM SPSS Modeler Server からデータベースに接続できる場合、IBM SPSS Modeler Solution Publisher Runtime の起動スクリプトから同じ *odbc.sh* スクリプトを参照して、同じサーバーに IBM SPSS Modeler Solution Publisher Runtime のインストールを設定することができます。

- IBM SPSS Modeler Solution Publisher Runtime で *modelerrun* スクリプトを編集して、スクリプトの最後の行のすぐ上に次の行を追加します。

. <odbc.sh_path>

odbc.sh_path は、この手順の始めに編集した *odbc.sh* ファイルへのパスです。例えば、次のようになります。

. /usr/spss/odbc/odbc.sh

注：ここではシンタックスが重要になります。最初のピリオドとファイルへのパスの間にスペースを入れてください。

2. *modelerrun* スクリプト・ファイルを保存します。

IBM SPSS Modeler Batch とともに起動するように ODBC を構成する手順

ODBC に IBM SPSS Modeler Batch スクリプトの設定は必要ありません。ストリームを実行するために IBM SPSS Modeler Batch から IBM SPSS Modeler Server に接続するためです。この項の始めに説明されているように、IBM SPSS Modeler Server ODBC 設定が行われ、適切に稼動していることを確認してください。

データ・ソース名を追加または編集するには

1. *odbc.ini* ファイルを編集して新しい名前または変更した名前が含まれるようにします。
2. この項の始めに説明されているように接続をテストします。

IBM SPSS Modeler Server と正しく接続している場合、新しいデータ・ソース名または変更されたデータ・ソース名も IBM SPSS Modeler Solution Publisher Runtime および IBM SPSS Modeler Batch に適切なものである必要があります。

ODBC 設定のトラブルシューティング

データ・ソースが表示されない、またはランダム テキストが表示される

データベース入力ノードを開いて利用できるデータ・ソースのリストが空であったり予期しないエントリがあったりした場合は、起動スクリプトの問題が原因の可能性があります。

1. \$ODBCINI が *modelersrv.sh* 内で、スクリプト自体で明示的に、または *modelersrv.sh* で参照される *odbc.sh* スクリプトで定義されていることを確認してください。
2. 後者の場合は、ODBCINI が ODBC データ・ソースの定義に使用する *odbc.ini* ファイルへのフル・パスを示すようにしてください。
3. ODBCINI のパスの指定が正しい場合、*modelersrv.sh* 内の変数を複製して、IBM SPSS Modeler Server 環境で使用されている \$ODBCINI の値を確認します。確認するには、ODBCINI を定義する場所の後の *modelersrv.sh* に次の行を追加します。

```
echo $ODBCINI
```
4. *modelersrv.sh* を保存して実行します。IBM SPSS Modeler Server 環境で設定される \$ODBCINI の値が検証のために *stdout* に書き込まれます。
5. 値が *stdout* に返されず、*modelersrv.sh* から参照している *odbc.sh* スクリプトで \$ODBCINI を定義している場合、参照しているシンタックスが正しいことを確認します。シンタックスは次のようになります。

. <odbc.sh_path>

odbc.sh_path は、この手順の始めに編集した *odbc.sh* ファイルへのパスです。例えば、次のようになります。

```
. /usr/spss/odbc/odbc.sh
```

注：ここではシンタックスが重要になります。最初のピリオドとファイルへのパスの間にスペースを入れてください。

modelersrv.sh の実行時に正しい値を *stdout* に複製すると、IBM SPSS Modeler Server を再起動してクライアントから接続するときデータベース入力ノードでデータ・ソース名を表示できる必要があります。

「データベース接続」ダイアログ・ボックスで「接続」をクリックした際の IBM SPSS Modeler クライアントの停止

この動作は、ライブラリー・パスが ODBC ライブラリーへのパスを含むよう正しく設定されていない場合に発生することがあります。ライブラリー・パスは \$LD_LIBRARY_PATH (64 ビット・バージョンでは \$LD_LIBRARY_PATH_64) によって使用されます。

IBM SPSS Modeler Server デーモン環境でライブラリー・パスの値を表示するには、ODBC ライブラリー・パスがライブラリー・パスの後に追加されている行の後で、*modelersrv.sh* の適切な環境変数の値を複製し、スクリプトを実行します。ライブラリー・パスの値は、次にスクリプトを実行した場合にターミナルに複製されます。

modelersrv.sh から *odbc.sh* を参照して IBM SPSS Modeler Server ODBC 環境を設定する場合、*odbc.sh* スクリプトを示す行の後の行からライブラリー・パスの値を複製します。値を複製するには、次の行をスクリプトに追加して、スクリプト・ファイルを保存し、実行します。

```
echo $<library_path_variable>
```

<library_path_variable> は、サーバーのオペレーティング・システムの該当するライブラリー・パス変数です。

ライブラリー・パスの戻り値には、ODBC の *lib* サブディレクトリーへのパスを含む必要があります。このパスを含まない場合、この場所をファイルに追加します。

64 ビットバージョンの IBM SPSS Modeler Server を稼動している場合、\$LD_LIBRARY_PATH_64 は \$LD_LIBRARY_PATH に優先されます (設定されている場合)。64 ビット プラットフォームのいずれかにこの問題が発生した場合、*modelersrv.sh* から \$LD_LIBRARY_PATH のほか LD_LIBRARY_PATH_64 を複製し、必要に応じて ODBC インストールの *lib* サブディレクトリーが含まれるよう \$LD_LIBRARY_PATH_64 を設定して定義をエクスポートします。

データ・ソース名が見つからず、デフォルトのドライバーが指定されていない

「データベース接続」ダイアログ・ボックスで「接続」をクリックした際にこのエラーが発生した場合、*odbc.ini* ファイルの定義が正しくないことを示します。ファイル上部の [ODBC Data Sources] セクション内で定義したデータ・ソース名 (DSN) が DSN を定義する *odbc.ini* の下部で大かっこに囲んで指定した文字列と一致していることを確認します。これらが異なる場合、IBM SPSS Modeler 内で DSN を使用して接続しようとするこのエラーが発生します。次に「正しくない」指定の例を示します。

```
[ODBC Data Sources]
Oracle=Oracle Wire Protocol
```

```
....
```

```
....
```

```
[Oracle Driver]
```

```
Driver=/usr/ODBC/lib/XEora22.so
Description=SPSS 5.2 Oracle Wire Protocol
AlternateServers=
....
```

正確に一致するよう、太字の 2 つの文字列のいずれかを変更する必要があります。変更することによってエラーが解決します。

指定したドライバーがロードできない

このエラーは、*odbc.ini* ファイルが正しく定義されていないことを示します。ドライバ節のドライバ パラメーターが次のように正しく設定されていない可能性があります。

```
[ODBC Data Sources]
Oracle=Oracle Wire Protocol
```

```
....
....
```

```
[Oracle]
Driver=/nosuchpath/ODBC/lib/XEora22.so
Description=SPSS 5.2 Oracle Wire Protocol
AlternateServers=
```

1. ドライバ パラメーターで指定された共有オブジェクトがあることを確認します。
2. 正しくない場合、共有オブジェクトへのパスを修正します。
3. ドライバ パラメーターが次の形式で指定されている場合

```
Driver=ODBCHOME/lib/XEora22.so
```

ODBC 関連スクリプトを初期化していないことを示します。ドライバでインストールした *setodbcpath.sh* スクリプトを実行します。詳しくは、トピック 57 ページの『IBM SPSS Modeler Server と データ・アクセス・パック』を参照してください。このスクリプトを実行すると、文字列「ODBCHOME」に ODBC インストールへのパスが代入されます。これにより、問題が解決されます。

他の原因がドライバーのライブラリーの問題となる場合があります。ドライバーがロードされないことを確認するために、ODBC とともに提供された *ivtestlib* を使用します。Connect64 の場合、*ddtestlib* ツールを使用します。起動スクリプトのライブラリー・パスの変数を設定して問題を修正します。

例えば、32-bit のインストールで Oracle ドライバーがロードされない場合は、次の手順に従ってください。

1. *ivtestlib* を使用して、ドライバーがロードされないことを確認します。例えば、UNIX のコマンド・プロンプトで、次のように入力します。

```
sh
cd ODBC DIR
. odbc.sh
./bin/ivtestlib MFor815
```

上記の ODBC DIR を、Connect for ODBC のインストール・ディレクトリーへのパスに置き換えます。

2. メッセージを読んで、エラーがあるかどうかを調べます。例えば、次のようなメッセージが表示されます。

Load of MFor815.so failed:ld.so.1:bin/ivtestlib:fatal:libclntsh.so:open failed:このようなファイルまたはディレクトリーがありません

このメッセージは、Oracle クライアント・ライブラリーの *libclntsh.so* がないか、または、このライブラリー・パス (例えば、Solaris の LD_LIBRARY_PATH) にないことを示しています。

- ライブラリーが存在することを確認します。存在しない場合は、Oracle クライアントを再インストールします。ライブラリーがある場合は、UNIX コマンド・プロンプトから次の一連のコマンドを入力します。

```
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/bigdisk/oracle/product/8.1.6/lib
export LD_LIBRARY_PATH
./bin/ivtestlib Mfor815
```

上記の */bigdisk/oracle/product/8.1.6/lib* は、*libclntsh.so* へのパスで置き換えます。また、LD_LIBRARY_PATH は、使用中のオペレーティング・システムのライブラリー・パス変数です。

IBM SPSS Modeler 64 ビットを Linux または Solaris で実行している場合、ライブラリー・パス変数には接尾辞 *_64* が含まれています。そのため、前述の例の最初の 2 行は次のようになります。

```
LD_LIBRARY_PATH_64=$LD_LIBRARY_PATH_64:/bigdisk/oracle/product/8.1.6/lib
export LD_LIBRARY_PATH_64
```

- メッセージを読んで、ドライバーがこれでロードされたことを確認します。例えば、次のようなメッセージが表示されます。

Load of MFor815.so successful, qehandle is 0xFF3A1BE4

このメッセージは、Oracle クライアント・ライブラリーがロードできたことを示しています。

- IBM SPSS Modeler 起動スクリプト内のライブラリー・パスを訂正します。
- 編集した起動スクリプト (*modelersrv.sh* または *rc.modeler*) を使用して、IBM SPSS Modeler Server を再起動します。

ライブラリー・パス

ライブラリー・パスの変数名は、使用するオペレーティング・システムによって異なります。次の表は、システムを構成またはトラブルシューティングしているときに、適切な置き換えを行うためのガイドとして含めてあります。

表 19. オペレーティング・システム別のライブラリー・パス

オペレーティング・システム	ライブラリー・パス変数名
AIX	LIBPATH
HP-UX	SHLIB_PATH
Solaris または Linux 32-bit	LD_LIBRARY_PATH
Solaris または Linux 64-bit	LD_LIBRARY_PATH_64

エンタープライズ・ビュー・ノードのドライバーを構成

エンタープライズ・ビュー・ノードを使用すると、共有の IBM SPSS Collaboration and Deployment Services Repository で IBM SPSS Modeler セッションおよび IBM SPSS Collaboration and Deployment Services Enterprise View 間の接続を作成および維持管理することができます。

IBM SPSS Modeler Server でエンタープライズ・ビュー・ノードを使用するには、IBM SPSS Collaboration and Deployment Services Enterprise View Driver をインストールし構成する必要があります。PEV ドライ

パーのインストールの詳細は、各サイトの管理者に問い合わせてください。インストール後、*pev.sh* への参照を、*modelersrv.sh*および *rc.modeler*の両方またはいずれかの始動スクリプトに、以下のフォーマットで追加する必要があります。

```
./usr/odbc/pev.sh
```

付録 C. SPSS Modeler Server を UNIX 上で root 以外のプロセスとして設定および実行

はじめに

以下の説明は、UNIX システムで root 以外のプロセスとして IBM SPSS Modeler Server を実行することについての情報です。

root として実行： IBM SPSS Modeler Server のデフォルトのインストールでは、サーバー・デーモン・プロセスが root として実行されると想定されています。root として実行されることにより、IBM SPSS Modeler は各ユーザーのログインを確実に認証し、対応する UNIX ユーザー・アカウントで各ユーザー・セッションを始動できるようになります。これで、ユーザーは自分のファイルとディレクトリーだけへのアクセス権限を持っていることが確実にになります。

非 root として実行： 非 root プロセスとして IBM SPSS Modeler Server を実行することは、サーバー・デーモン・プロセスの、現実の、そして有効なユーザー ID を、自分が選択したアカウントへ設定させる、ということです。SPSS Modeler Server によって開始するすべてのユーザー・セッションでは同じ UNIX アカウントを使用します。つまり SPSS Modeler によって読み込まれたり書き込まれたりするファイル・データはすべての SPSS Modeler ユーザーによって共有されます。ユーザーが使用するそれぞれのデータベース・データに対して独立して自分自身を認証する必要があるため、データベースのデータへのアクセスは影響を受けません。root 権限がない場合、IBM SPSS Modeler は次のいずれかの方法で処理します。

- **非公開のパスワード・データベースがない：** この方法では、SPSS Modeler は、UNIX システムのユーザー認証に通常使用される既存の UNIX パスワード・データベース、NIS、LDAP サーバーを使用します。詳しくは、トピック『非公開のパスワード・データベースがない状態で root 以外のユーザーとして設定』を参照してください。
- **非公開のパスワード・データベースがある：** この方法では、SPSS Modeler は、UNIX 認証に通常使用される UNIX パスワード・データベース、NIS、LDAP サーバーとは異なる非公開のパスワード・データベースに対してユーザーを認証します。詳しくは、トピック 66 ページの『非公開のパスワード・データベースを使用して root 以外のユーザーとして設定』を参照してください。

非公開のパスワード・データベースがない状態で root 以外のユーザーとして設定

非公開のパスワード・データベースを使用せずに非 root アカウントで実行するように IBM SPSS Modeler Server を設定するには、次の手順に従ってください。

1. SPSS Modeler Server *options.cfg* ファイルを開いて編集します。
2. オプション *start_process_as_login_user* を Y に設定します。
3. *options.cfg* ファイルを保存して閉じます。

非 root として実行する場合、設定の更新が必要になることがあります。詳しくは、トピック 68 ページの『ユーザー認証エラーのトラブルシューティング』を参照してください。

注意: `start_process_as_login_user` 設定を有効にしてから、IBM SPSS Modeler Server を `root` として開始しないでください。これにより、サーバーに接続しているすべてのユーザーにとって、サーバー・プロセスが `root` として実行されます。これはセキュリティ上リスクがあります。

非公開のパスワード・データベースを使用して root 以外のユーザーとして設定

非公開のパスワード・データベースを使用してユーザー認証する場合、すべてのユーザー・セッションが同じ非 `root` ユーザー・アカウントで開始します。

非 `root` アカウントで実行するために IBM SPSS Modeler Server を構成するには、次の手順に従ってください。

1. すべてのユーザーを含むグループを作成します。このグループの名前はどのように付けることもできますが、この例では `modelerusers` とします。
2. IBM SPSS Modeler Server を実行するユーザー・アカウントを作成します。このアカウントは、IBM SPSS Modeler Server デモン・プロセスの専用です。この例では、このアカウントを `modelerserv` とします。

アカウントを作成するときは、次の点に注意してください。

- プライマリ グループは、以前作成した `<modelerusers>` にする必要があります。
- ホーム・ディレクトリーは、IBM SPSS Modeler のインストール・ディレクトリーであってもよいし、その他の便利なデフォルトのディレクトリーであってもかまいません (アカウントをアップグレードでも存続させる必要がある場合は、インストール・ディレクトリー以外の別のディレクトリーを使用することを考慮します)。

3. 次に、新しく作成されたアカウントを使用して IBM SPSS Modeler Server を起動する起動スクリプトを構成します。適切な起動スクリプトを特定し、テキスト・エディター内で開きます。詳しくは、トピック 55 ページの『スクリプト』を参照してください。

- a. 作成されたファイルに対して少なくともグループ読み取りアクセス権が持てるように、`umask` を変更します。

```
umask 027
```

4. 次の行を追加することでプライベート・パスワード・データベースに対する認証を指定するために、サーバー・オプション・ファイルの `config/options.cfg` を編集します。

```
authentication_methods, "pasw_modeler"
```

5. オプション `start_process_as_login_user` を `Y` に設定します。

6. 次に、`config/passwords.cfg` ファイル内に格納されるプライベート・パスワード・データベースを作成する必要があります。このパスワード・ファイルで、IBM SPSS Modeler へのログインが許可されるユーザー名とパスワードの組み合わせが定義されます。注：これらは IBM SPSS Modeler 専用であり、UNIX へログインするのに使用されるユーザー名とパスワードとは、何のつながりもありません。便宜上同じユーザー名を使用できますが、同じパスワードを使用することはできません。

パスワード・ファイルを作成するには、IBM SPSS Modeler Server のインストール先の `bin` ディレクトリー内にある、`pwutil` というパスワード ユーティリティー プログラムを使用する必要があります。このプログラムの概要は、次のとおりです。

```
pwutil [ username [ password ] ]
```


このプログラムでユーザー名とプレーンテキストのパスワードが指定され、ユーザー名と暗号化されたパスワードは、パスワード・ファイルに含めるのに適した形式で、標準出力に書き出されます。例えば、ユーザー *modeler* とパスワード「data mining」を定義するには、次のように入力します。

```
bin/pwutil modeler "data mining" > config/passwords.cfg
```

ほとんどの場合、1 つのユーザー名を定義するだけで十分です。そのときは、すべてのユーザーが同じ名前とパスワードでログインします。ただし、次の例のように、>> 演算子を使用してファイルに付加する追加ユーザーを作成できます。

```
bin/pwutil modeler "data miner2" >> config/passwords.cfg
```

注：1 つの > が使用された場合は、*passwords.cfg* の内容が以前のユーザーの設定を置き換えて、毎回上書きされます。すべてのユーザーが同じ UNIX ユーザー・アカウントを共有していることに注意してください。

注: SPSS Modeler Server の実行中に新規ユーザーをプライベート・パスワード・データベースに追加した場合は、新しく定義したユーザーが認識されるように、SPSS Modeler Server を再起動する必要があります。これを行うまでは、SPSS Modeler Server が最後に再起動されてから pwutil を使用して追加されたすべての新規ユーザーによるログインは失敗します。

7. IBM SPSS Modeler のインストール先ディレクトリーとその内容の所有権がユーザー <*modelerserv*> とグループ <*modelerusers*> のものになるように、再帰的に変更します。参照されるこれらの名前は、以前作成したものです。例:

```
chown -R -h modelerserv:modelerusers .
```

8. 他からの干渉なしに作業データを格納できるように、IBM SPSS Modeler ユーザーのデータ・ディレクトリー内にサブディレクトリーを作成することを考慮します。これらのディレクトリーは <*modelerusers*> グループによりグループ所有され、読み込み、書き込み、および検索の権限を持っている必要があります。例えば、ユーザー *bob* 用の作業ディレクトリーを作成するには、次のように入力します。

```
mkdir data/bob
chown bob:modelerusers data/bob
chmod ug=rwx,o= data/bob
```

さらに、グループ ID 設定ビットをそのディレクトリーに設定して、そのディレクトリーにコピーされたすべてのデータ・ファイルが自動的に <*modelerusers*> にグループ所有されるようにすることができます。

```
chmod g+s data/bob
```

SPSS Modeler Server を非 root ユーザーとして実行する

非 root アカウントで SPSS Modeler Server を実行するためには、次の手順に従ってください。

1. 以前作成した非 root のユーザーアカウントでログインします。
2. 構成ファイルのオプション *start_process_as_login_user* を有効にした状態で実行している場合は、SPSS Modeler Server の開始、停止、および状況の確認を実行できます。詳しくは、トピック 21 ページの『UNIX 上で起動、停止、またはステータスをチェックする手順』を参照してください。

エンド・ユーザーは、クライアント・ソフトウェアからログインして、SPSS Modeler Server に接続します。サーバー・コンピューターの IP アドレスまたはホスト名など、接続に必要な情報をエンド・ユーザーに提供する必要があります。

IBM SPSS Modeler Server を AIX の非 root ユーザーとして実行する

AIX の非 root ユーザーとして IBM SPSS Modeler Server を起動する場合、エラー `lsattr not found` がコンソールに出力される場合があります。これは、`/usr/sbin` ディレクトリーに `lsattr` があるにもかかわらず、新規ユーザー (非 root として実行) が IBM SPSS Modeler Server にログインすると、`/usr/bin` と `/bin` のみを含むパスにリセットされるために発生します。

このエラーが発生すると、解決策として `/usr/bin` の下に次のリンクが作成されます。

```
su
cd /usr/bin
ln -s /usr/sbin/lsattr lsattr
```

リンクを作成した後、IBM SPSS Modeler Server を再起動します。

ユーザー認証エラーのトラブルシューティング

認証を行うためにオペレーティング・システムを設定する方法によって、非 root 設定で実行する場合に SPSS Modeler Server にログオンしようとしてエラーが発生する場合があります。例えば、オペレーティング・システムを構成 (`/etc/nsswitch.conf` ファイルなどを使用) して NIS または LDAP ではなくローカルのシャドー・パスワード・ファイルを確認する場合にエラーが発生することがあります。これは、SPSS Modeler Server に安全なユーザー・アカウント情報を保存する、`/etc/shadow` ファイルなど (AIX の `/etc/security/passwd` など)、認証に使用されるファイルへの読み取りアクセス権限が必要なためです。しかし、オペレーティング・システム・ファイルの権限は、`/etc/shadow` が root ユーザーによってのみアクセス可能なように設定されています。こうした条件のもと、非 root プロセスではユーザー・パスワードを検証する `/etc/shadow` を読み取ることができず、認証エラーとなります。

この問題を解決するにはいくつかの方法があります。

- システム管理者に、NIS や LDAP を認証に使用するようオペレーティング・システムを設定するよう依頼します。
- SPSS Modeler Server の実行に使用されたローカル・ユーザー・アカウントがファイルにアクセスできるよう、`/etc/shadow` ファイルへの読み取りアクセスを付与することなどによって、保護されたファイルに対するファイルアクセス権限を変更します。実動環境では不適切かもしれませんが、テスト環境に一時的に適用して、認証のエラーがオペレーティング・システムにリンクしているかどうかを検証することもできます。
- `/etc/shadow` ファイルのアクセス制御リスト (ACL) を指定します (本セクションの例を参照)。
- SPSS Modeler Server を root として実行し、サーバー・プロセスで `/etc/shadow` ファイルを読み取れるようにします。警告: この場合は、前述したセキュリティー上の問題を防ぐため、SPSS Modeler Server の `options.cfg` ファイルでオプション `start_process_as_login_user`, N が指定されていることを確認してください。

アクセス制御リスト (ACL) の例

Solaris では、`setfacl` コマンドを使用してアクセス制御リストを作成することができます (管理者に確認してこのコマンドがシステムで利用可能となるようにします)。次の例では、ユーザー `modserv` の ACL を作成する方法を示しています。

```
# vi /tmp/shadow.acl
user::r--
user:modserv:r--
group::--
mask:r--
```

```
other:--  
  
# setfacl -f /tmp/shadow.ac1 /etc/shadow  
  
# getfacl /etc/shadow  
  
# file: /etc/shadow  
# owner: root  
# group: sys  
user::r--  
user:modserv:r--      #effective:r--  
group::---            #effective:---  
mask:r--  
other:---
```

付録 D. Windows での非公開パスワード・ファイルを使用した SPSS Modeler Server の構成および実行

はじめに

以下の手順では、Windows システムで非公開パスワード・ファイルを使用して IBM SPSS Modeler Server を実行する方法について説明します。この方法を使用すると、IBM SPSS Modeler は、Windows におけるシステム認証とは異なり、プライベート・パスワード・データベースに対してユーザーを認証します。

プライベート・パスワード・データベースを使用した設定

プライベート・パスワード・データベースを使用してユーザー認証する場合、すべてのユーザー・セッションが同じユーザー・アカウントで開始されます。

この方法で IBM SPSS Modeler Server を設定するには、次の手順に従ってください。

1. IBM SPSS Modeler Server を実行するユーザー・アカウントを作成します。このアカウントは、IBM SPSS Modeler Server デーモン・プロセスの専用です。この例では、このアカウントを *modelerserv* とします。
2. サーバー・オプション・ファイル *config/options.cfg* を編集して *start_process_as_login_user* オプションを Y に設定し、次の行を追加してプライベート・パスワード・データベースに対する認証を指定します。

```
authentication_methods, "pasw_modeler"
```

3. 次に、*config/passwords.cfg* ファイル内に格納されるプライベート・パスワード・データベースを作成する必要があります。このパスワード・ファイルで、IBM SPSS Modeler へのログインが許可されるユーザー名とパスワードの組み合わせが定義されます。注：これらは IBM SPSS Modeler 専用であり、Windows へのログインに使用されるユーザー名およびパスワードとは、何の関連もありません。便宜上同じユーザー名を使用できますが、同じパスワードを使用することはできません。

パスワード・ファイルを作成するには、IBM SPSS Modeler Server のインストール先の *bin* ディレクトリー内にある、*pwutil* というパスワード ユーティリティー プログラムを使用する必要があります。このプログラムの概要は、次のとおりです。

```
pwutil [ username [ password ] ]
```

このプログラムでユーザー名とプレーンテキストのパスワードが指定され、ユーザー名と暗号化されたパスワードは、パスワード・ファイルに含めるのに適した形式で、標準出力に書き出されます。例えば、ユーザー *modeler* とパスワード「data mining」を定義するには、次のように入力します。

```
bin/pwutil modeler "data mining" > config/passwords.cfg
```

ほとんどの場合、1 つのユーザー名を定義するだけで十分です。そのときは、すべてのユーザーが同じ名前とパスワードでログインします。ただし、次の例のように、*>>* 演算子を使用してファイルに付加する追加ユーザーを作成できます。

```
bin/pwutil modeler "data miner2" >> config/passwords.cfg
```

注：1 つの > が使用された場合は、*passwords.cfg* の内容が以前のユーザーの設定を置き換えて、毎回上書きされます。すべてのユーザーが同じ UNIX ユーザー・アカウントを共有していることに注意してください。

注: SPSS Modeler Server の実行中に新規ユーザーをプライベート・パスワード・データベースに追加した場合は、新しく定義したユーザーが認識されるように、SPSS Modeler Server を再起動する必要があります。これを行うまでは、SPSS Modeler Server が最後に再起動されてから *pwutil* を使用して追加されたすべての新規ユーザーによるログインは失敗します。

4. ステップ 1 で作成したユーザーに、サーバー・オプション・ファイル *config/options.cfg* および `%ALLUSERSPROFILE%\IBM\SPSS` ディレクトリーに対する完全な制御権を割り当てます
5. システム・サービスで、*IBM SPSS Modeler Server* サービスを停止し、「ログオン」を「ローカル システム アカウント」から、ステップ 1 で作成したユーザー・アカウントに変更して、サーバーを再起動します。

付録 E. サーバー・クラスターでの負荷バランシング

IBM SPSS Collaboration and Deployment Services により、Coordinator of Processes というプラグインを使用してネットワーク上のサービスを管理できるようになりました。Coordinator of Processes は、クライアント - サーバー間の通信および処理を最適化するように設計された、サーバー管理機能を備えています。

IBM SPSS Statistics Server や IBM SPSS Modeler Server など、管理対象のサービスは、起動時に Coordinator of Processes に登録し、更新されたステータス メッセージを定期的送信します。サービスは、必要な構成ファイルを IBM SPSS Collaboration and Deployment Services Repository に格納し、初期化時に取得することもできます。

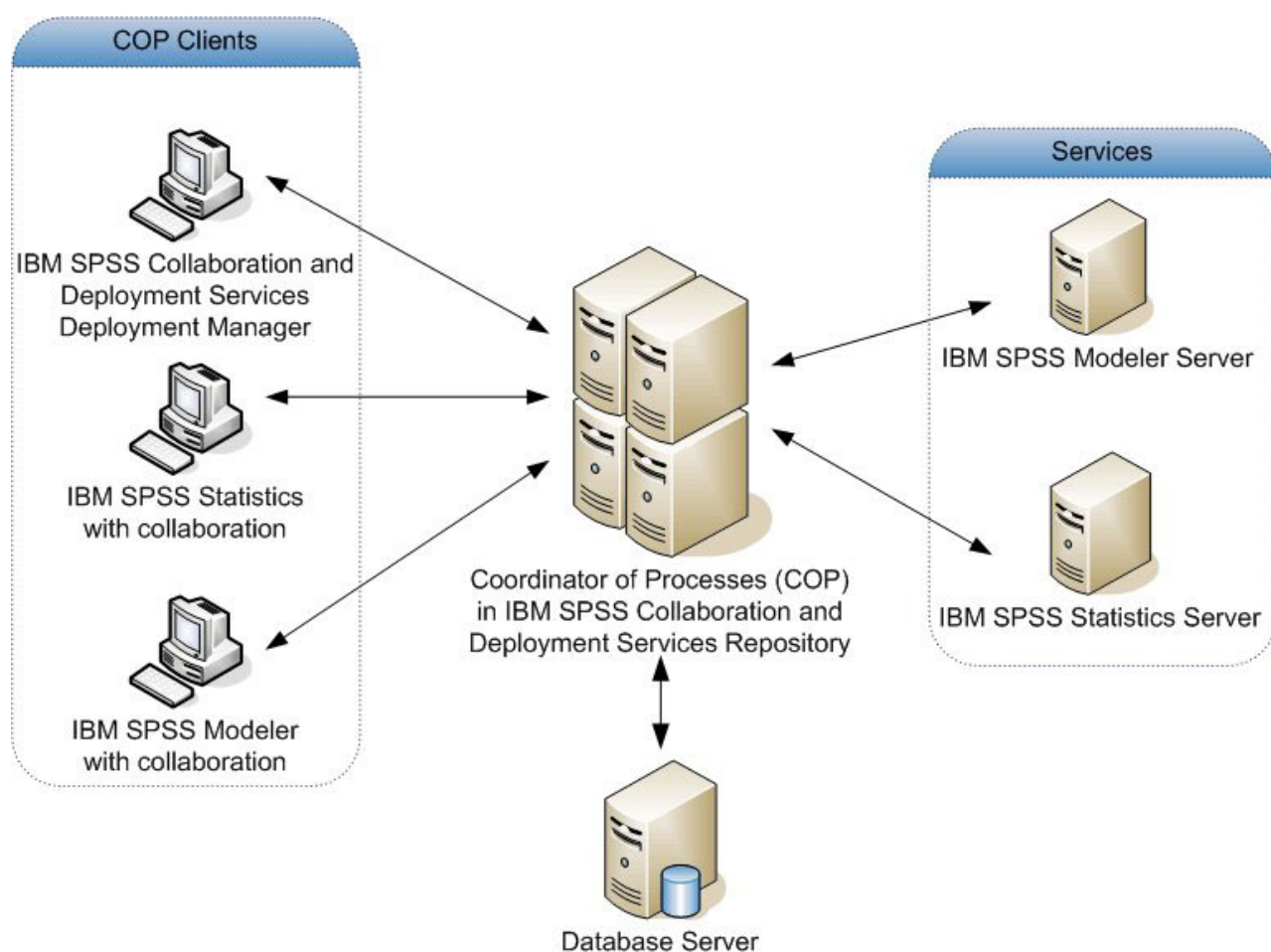


図 5. Coordinator of Processes アーキテクチャー

IBM SPSS Modeler ストリームをサーバー上で実行すると、パフォーマンスが向上する場合があります。サーバーの選択肢が 1 台か 2 台の場合もあれば、サーバーの選択肢が多数になる場合もあります。これは、所有者、アクセス権、サーバー・データ、テスト・サーバーと実動サーバーなど、サーバー間に顕著な違いがある場合です。さらに、Coordinator of Processes がネットワーク上にある場合は、サーバー・クラスターが提供されることもあります。

サーバー・クラスターとは、構成やリソースを相互に交換可能なサーバーのグループのことです。Coordinator of Processes は、サーバーの重み、ユーザーの優先度、現在の処理負荷など、いくつかの条件に従って負荷のバランスを取るアルゴリズムを使用して、処理要求への応答に最も適したサーバーを特定します。詳細は、「*Coordinator of Processes Service Developer's Guide*」を参照してください (IBM Corp. に請求して入手できます)。

IBM SPSS Modeler でサーバーまたはサーバー・クラスターに接続するたびに、サーバーを手動で入力するか、Coordinator of Processes を使用してサーバーまたはクラスターを検索することができます。詳しくは、トピック 15 ページの『IBM SPSS Modeler Server に接続中』を参照してください。

特記事項

本情報は全世界で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。

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

以下の保証は、国または地域の法律に沿わない場合は、適用されません。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、お客様に対してなん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 Software Group

ATTN: Licensing

200 W. Madison St.

Chicago, IL; 60606

U.S.A.

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。

この文書に含まれるいかなるパフォーマンス・データも、管理環境下で決定されたものです。そのため、他の操作環境で得られた結果は、異なる可能性があります。一部の測定が、開発レベルのシステムで行われた可能性があります。その測定値が、一般に利用可能なシステムのものと同じである保証はありません。さらに、一部の測定値が、推定値である可能性があります。実際の結果は、異なる可能性があります。お客様は、お客様の特定の環境に適したデータを確かめる必要があります。

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をお願いします。

IBM の将来の方向または意向に関する記述については、予告なしに変更または撤回される場合があります、単に目標を示しているものです。

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、名称や住所が類似する企業が実在しているとしても、それは偶然にすぎません。

この情報をソフトコピーでご覧になっている場合は、写真やカラーの図表は表示されない場合があります。

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。

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。

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。

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]

- アーキテクチャー
 - コンポーネント 7
- アプリケーションの例 3
- 一時ディレクトリー
 - IBM SPSS Modeler Server の 25
- 一時ファイル 10
 - 権限 (IBM SPSS Modeler Server) 20
- 演算子
 - SQL 生成 48
- オペレーティング・システム
 - 64 ビット 9

[カ行]

- 管理
 - IBM SPSS Modeler Server の 23
- 管理者アクセス
 - IBM SPSS Modeler Server の 24
- キャッシュの圧縮 26
- キャッシング、データベース内 30
- 結果
 - クライアントとサーバーの違い 20
 - 丸め 20
 - レコードの並び順 20
- 構成オプション
 - 一時ディレクトリー 25
 - 概要 25
 - 自動 SQL 生成 27
 - ストリームの再書き込み 26
 - 接続とセッション 25
 - データ・ファイル・アクセス 25
 - パフォーマンスと最適化 26
 - 並列処理 26
 - ポート番号 25
 - メモリー管理 26
 - ログインの試行 25
 - Coordinator of Processes 28
 - COP 28
 - IBM SPSS Modeler Server の 23
 - SQL 文字列の長さ 27
 - SSL データ暗号化 28

[サ行]

- サーバー・ポート設定
 - options.cfg ファイル 26
- 最適化
 - SQL 生成 39, 40, 42
- 自動的なサーバーの起動
 - UNIX 上で構成 56
- 状況
 - UNIX 上の IBM SPSS Modeler Server の 21
 - Windows 上の IBM SPSS Modeler Server の 21
- シングル・サインオン 15
- ストリーム実行のエラー 26
- ストリームの再書き込み 43
- セキュリティ
 - ファイルの作成 20
 - SSL 31
 - UNIX 上でファイルの作成を構成 57
- 接続
 - サーバー・クラスター (server cluster) 19
 - IBM SPSS Modeler Server への 15, 18, 19
- 接続の COP の検索 19
- ゾンビ・プロセス、IBM SPSS Modeler Server 22

[タ行]

- データベース
 - アクセス 11
- データベース接続
 - 閉じる 30
- データベース内キャッシング 30
- データベース内マイニング 37
- データベース・キャッシング
 - options.cfg からの制御 30
 - SQL 生成 43
- データベース・サーバー 37
- データ・アクセス・バック
 - および UNIX ライブラリー・パス 62
 - ODBC、UNIX 上で構成 57
 - UNIX 上での ODBC のトラブルシューティング 59
 - UNIX の構成 57
- データ・ファイル
 - インポートとエクスポート 13
 - IBM SPSS Statistics 13
- ディスク・スペース 10

- ディスク・スペース (続き)
 - 計算 10
- ドメイン名 (Windows)
 - IBM SPSS Modeler Server 15

[ナ行]

- 認証 19
- ノード
 - SQL 生成をサポートする 44
- ノード・キャッシング
 - データベースへの書き込み 43
 - SQL 生成 43

[ハ行]

- ハードウェアの要件
 - IBM SPSS Modeler Server の 9
- ハード・ドライブ 10
- パス 12
- パスワード
 - IBM SPSS Modeler Server 15
- パフォーマンス
 - IBM SPSS Modeler Server の 35
- ファイアウォールの設定
 - options.cfg ファイル 26
- ファイル権限
 - IBM SPSS Modeler Server 上の 20
 - UNIX 上で構成 57
- ファイル名
 - UNIX 12
 - Windows 12
- 複数ストリームの実行 26
- プッシュバック 39, 40, 42, 44
 - CLEM 式 48
- プロセス、非応答 22
- プロセッサ 9
 - 複数 26
- 文書 3
- 並列処理
 - 制御 26
- ポート設定
 - options.cfg ファイル 26
- ポート番号
 - IBM SPSS Modeler Server 15, 18, 25
- ホスト名 (host name)
 - IBM SPSS Modeler Server 15, 18

[マ行]

- 紫色のノード
 - SQL 最適化 41
- メッセージ
 - 生成された SQL の表示 43
- メモリー 11
- メモリー管理
 - 管理オプション 26
- モデル
 - メモリー管理 26
- モデル・ナゲット
 - SQL の表示 43

[ヤ行]

- ユーザー ID
 - IBM SPSS Modeler Server 15
- ユーザーの認証 19
- ユーザー・アカウント
 - IBM SPSS Modeler Server 19

[ラ行]

- 例
 - アプリケーション ガイド 3
 - 概要 5
- ログ・ファイル
 - 生成された SQL の表示 43

[数字]

- 64-bit のオペレーティング・システム 9

A

- allow_modelling_memory_override
 - options.cfg ファイル 26
- auto.sh (UNIX)
 - 場所 55

C

- cache_compression
 - options.cfg ファイル 26
- cache_connection オプション 30
- chemsrv.sh (UNIX)
 - 場所 55
- CLEM 式
 - SQL 生成 48
- Coordinator of Processes 19
 - サーバー・クラスター 73
 - ロード・バランシング (load balancing) 73

- Coordinator of Processes の構成
 - IBM SPSS Modeler Server の 28
- COP 19
 - サーバー・クラスター 73
 - ロード・バランシング (load balancing) 73
- COP の構成
 - IBM SPSS Modeler Server の 28
- cop_enabled
 - options.cfg ファイル 28
- cop_host
 - options.cfg ファイル 28
- cop_password
 - options.cfg ファイル 28
- cop_port_number
 - options.cfg ファイル 28
- cop_service_description
 - options.cfg ファイル 28
- cop_service_host
 - options.cfg ファイル 28
- cop_service_name
 - options.cfg ファイル 28
- cop_service_weight
 - options.cfg ファイル 28
- cop_update_interval
 - options.cfg ファイル 28
- cop_user_name
 - options.cfg ファイル 28

D

- data access 11
- data_files_restricted
 - options.cfg ファイル 25
- data_file_path
 - options.cfg ファイル 25
- DB2
 - SQL 最適化 39, 40

E

- encryption
 - FIPS 29
 - SSL 31

F

- FIPS 暗号化 29

I

- IBM SPSS Analytic Server
 - 構成オプション 25
- IBM SPSS Modeler 1
 - 文書 3

- IBM SPSS Modeler Administration
 - Console 23
 - 管理者アクセス 24
- IBM SPSS Modeler Server 2
 - 一時ディレクトリー 25
 - エンド・ユーザー向けの情報 19
 - 管理オプション 23
 - 管理者アクセス 24
 - クライアントと比べた結果の違い 20
 - 構成オプション 25
 - サーバー・プロセス 29
 - 使用率の監視 29
 - シングル・サインオン 16, 17
 - ドメイン名 (Windows) 15
 - の管理 23
 - パスワード 15
 - 非応答プロセス 22
 - ファイルの作成 20
 - ポート番号 15, 18, 25
 - ホスト名 (host name) 15, 18
 - ユーザー ID 15
 - ユーザーの認証 19
 - ユーザー・アカウント 19
 - Coordinator of Processes の構成 28
 - COP の構成 28

- IBM SPSS Modeler Server 接続の追加 18, 19

- IBM SPSS Modeler Server の開始
 - UNIX 上で 21
 - Windows の場合 21

- IBM SPSS Modeler Server の停止
 - UNIX 上で 21
 - Windows の場合 21

- IBM SPSS Modeler Server へのログイン 15

- IBM SPSS Statistics データ・アクセス・テクノロジー 11

- IBM SPSS Statistics データ・ファイルインポートとエクスポート 13

- io_buffer_size
 - options.cfg ファイル 26

M

- max_file_size
 - options.cfg ファイル 25
- max_login_attempts
 - options.cfg ファイル 25
- max_parallelism
 - options.cfg ファイル 26
- max_sessions
 - options.cfg ファイル 25
- max_sql_string_length
 - options.cfg ファイル 27
- memory_usage
 - options.cfg ファイル 26

Microsoft SQL Server
SQL 最適化 39, 40
modelling_memory_limit_percentage
options.cfg ファイル 26

O

ODBC
UNIX 上で構成 57
ODBC データ・ソース
および UNIX 57
ODBC と UNIX スクリプト 57
options.cfg ファイル 30
Oracle
SQL 最適化 39, 40, 53

P

port_number
options.cfg ファイル 25
preview
SQL 生成 43
program_files_restricted
options.cfg ファイル 25
program_file_path
options.cfg ファイル 25

R

RAM 11
rc.modeler (UNIX)
場所 55

S

Secure Sockets Layer 31
server
サーバーの COP の検索 19
シングル・サインオン 16, 17
接続の追加 18
ログイン 15
SQL
クエリー 51
生成のプレビュー 43
重複する列名 51
モデル・ナゲットの表示 43
Oracle を最適化する 53
SQL Server
SQL 最適化 39, 40
SQL 生成 39, 40, 44
使用可能化 42
ストリームの再書き込み 43
ヒント 43
レビュー 43
モデル・ナゲットの表示 43

SQL 生成 (続き)
ログ記録 43
CLEM 式 43, 48
IBM SPSS Modeler Server での有効化
27
SQL プッシュバック。「SQL 生成」も参
照 39
sql_generation_enabled
options.cfg ファイル 27
SSL 31
概要 31
通信の保護 31
SSL データ暗号化
IBM SPSS Modeler Server での有効化
28
ssl_certificate_file
options.cfg ファイル 28
ssl_enabled
options.cfg ファイル 28
ssl_private_key_file
options.cfg ファイル 28
ssl_private_key_password
options.cfg ファイル 28
stream_rewriting_enabled
options.cfg ファイル 26

T

temp_directory
options.cfg ファイル 25

U

UNC ファイル名 12
UNIX
シングル・サインオン 17
ファイル権限の構成 57
ユーザーの認証 19
ライブラリー・パス 62
UNIX カーネルの制限 21
UNIX スクリプト
編集 56
auto.sh 55
modelersrv.sh 55
rc.modeler 55
UNIX のカーネルの制限 21

[特殊文字]

@SQLFN 関数 51



Printed in Japan