

IBM SPSS Modeler Server 14.2  
管理およびパフォーマンス ガ  
イド



注： サポートされている情報および製品をご利用いただく前に、「注意事項」（ p. ）の一般情報をお読みください。

本マニュアルには、SPSS Inc., an IBM Company が所有する情報が含まれています。これらの情報は使用許諾契約書に基づいて提供され、著作権法によって保護されています。本文書に記載されている情報には、製品の保証は含まれていません。また本マニュアルに記載されている文は製品の保証を規定しないものとします。

IBM または SPSS に情報を送信すると、あなたに対する義務を負うことなく、適切とする方法でその情報を使用または配布する非独占的権利と IBM および SPSS 付与するものとします。

© Copyright IBM Corporation 1994, 2011..

---

# はじめに

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 を使っ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イスを使用すると、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、さまざまなモデル作成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、意思決定者がビジネス パフォーマンスを向上させるために信頼する完全で、一貫した正確な情報を提供します。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、および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。豊富な業界のソリューション、実績ある実例、専門サービスと組み合わせ、さまざまな規模の組織が、高い生産性を実現、意思決定を自信を持って自動化し、より良い決定をもたらします。

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 ソフトウェアを日々の業務に取り入れることによって、組織は業務目標を達成し、大きな競争的優位を獲得することができるよう、意思決定を方向付け、自動化することができるようになります。お問い合わせは、<http://www.ibm.com/spss> を参照してください。

## テクニカル 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 サポートの詳細は、IBM Corp. Web ページ <http://www.ibm.com/support> を参照してください。ご本人、組織、サポートの同意を確認できるものをご用意ください。

---

# 内容

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>IBM SPSS Modeler について</b>                                      | <b>1</b>  |
|          | IBM SPSS Modeler Server . . . . .                                 | 1         |
|          | IBM SPSS Modeler のオプション . . . . .                                 | 2         |
|          | IBM SPSS Text Analytics . . . . .                                 | 2         |
|          | IBM SPSS Modeler ドキュメント . . . . .                                 | 3         |
|          | アプリケーションの例 . . . . .  | 4         |
|          | Demos フォルダ . . . . .  | 5         |
| <b>2</b> | <b>アーキテクチャとハードウェアに関する推奨事項</b>                                     | <b>6</b>  |
|          | IBM SPSS Modeler のアーキテクチャ . . . . .                               | 6         |
|          | アーキテクチャの説明 . . . . .  | 6         |
|          | 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 . . . . .   | 12        |
|          | データ ファイルの参照 . . . . .   | 14        |
|          | IBM SPSS Statistics データ ファイルのインポート . . . . .                      | 14        |
|          | インストール手順 . . . . .  | 15        |
| <b>3</b> | <b>IBM SPSS Modeler サポート</b>                                      | <b>16</b> |
|          | IBM SPSS Modeler Server への接続 . . . . .                            | 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 . . . . . | 20        |
|          | データとファイル システム . . . . .   | 21        |
|          | ユーザーの認証 . . . . .   | 21        |
|          | ファイルの作成 . . . . .   | 22        |
|          | 結果の違い . . . . .   | 22        |

## 4 IBM SPSS Modeler Server 管理 24

|   |    |
|---|----|
| IBM SPSS Modeler Server の起動と停止                | 24 |
| Windows 上で起動、停止、またはステータスをチェックする手順             | 24 |
| UNIX 上で起動、停止、またはステータスをチェックする手順                | 25 |
| 非応答性サーバー プロセス (UNIX システム) の処理                 | 26 |
| 管理  | 26 |
| IBM SPSS Modeler Administration Console の使用方法 | 27 |
| Modeler Administration Console の起動            | 27 |
| Modeler Administration Console でのアクセスの構成      | 28 |
| SPSS Modeler Server 接続                        | 29 |
| SPSS Modeler Server の構成                       | 30 |
| SPSS Modeler Server の監視                       | 38 |
| options.cfg ファイルの使用                           | 39 |
| 使用されていないデータベース接続の終了                           | 40 |
| IBM SPSS Statistics ライセンス ロケーション              | 41 |
| 安全にデータを転送するための SSL の使用                        | 41 |
| SSL の動作方法                                     | 42 |
| 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       | 42 |
| URL プレフィックスの設定                                | 47 |
| SSL による LDAP の保護                              | 47 |
| サーバー ログ                                       | 48 |

## 5 パフォーマンスの概要 51

|                       |    |
|-----------------------|----|
| サーバーのパフォーマンスと最適化の設定   | 51 |
| クライアントのパフォーマンスと最適化の設定 | 52 |
| データベースの使用率と最適化        | 55 |
| SQL 最適化               | 55 |

## 6 SQL 最適化 57

|                   |    |
|-------------------|----|
| SQL 生成の動作原理       | 58 |
| SQL 生成の例          | 59 |
| SQL 最適化の設定        | 61 |
| 生成された SQL のプレビュー  | 62 |
| モデル ナゲットの SQL の表示 | 63 |

|                          |    |
|--------------------------|----|
| SQL 生成を最大限にするためのヒント      | 64 |
| SQL 生成をサポートしているノード       | 65 |
| SQL 生成をサポートする CLEM 式と演算子 | 69 |
| CLEM 式での SQL 関数の使用       | 72 |
| SQL クエリーの記述              | 72 |

## 付録

### A UNIX プラットフォームでの Oracle の設定 74

|                      |    |
|----------------------|----|
| SQL 最適化用の Oracle の設定 | 74 |
|----------------------|----|

### B UNIX 起動スクリプトの構成 76

|  |    |
|--|----|
| はじめに                                   | 76 |
| スクリプト                                  | 76 |
| IBM SPSS Modeler Server の自動的な起動と停止     | 77 |
| IBM SPSS Modeler Server の手動による起動と停止    | 77 |
| スクリプトの編集                               | 78 |
| ファイル作成時の権限の制御                          | 78 |
| IBM SPSS Modeler Server と データ アクセス パック | 78 |
| Connect for ODBC 構成のトラブルシューティング        | 82 |
| ライブラリ パス                               | 86 |
| エンタープライズ ビュー ノードのドライバを構成               | 86 |

### C UNIX 上での非ルート プロセスとしての実行 87

|   |    |
|---|----|
| はじめに  | 87 |
| 非 root プロセスとしての IBM SPSS Modeler Server の構成 | 87 |

|   |                     |    |
|---|---------------------|----|
| D | サーバー クラスタでの負荷バランシング | 91 |
| E | 注意事項                | 93 |
|   | 索引                  | 97 |



# 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。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。

SPSS Modeler ではさまざまなモデル作成方法を提供しています。[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。詳細は、<http://www.ibm.com/software/analytics/spss/products/modeler/> 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。ここに挙げた以外にも、ほかの製品やアップデートも利用できる可能性があります。詳細は、<http://www.ibm.com/software/analytics/spss/products/modeler/> を参照してください。

**SPSS Modeler:** SPSS Modeler はこの製品のすべての機能を搭載したバージョンであり、ユーザーのデスクトップ コンピュータにインストールし、そのコンピュータで実行します。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 Server:** SPSS Modeler Server は、1 つまたは複数の IBM® SPSS® Modeler のインストールと同時に分散分析モードで継続的に実行し、大規模なデータセットを使用する際にパフォーマンスが大幅に向上しますが、それは、データをクライアント コンピュータへダウンロードする

ことなく、メモリー主体の操作をサーバー上で実行できるからです。また、SPSS Modeler Server は SQL 最適化のサポート、データベース内モデル作成機能を提供し、パフォーマンスおよび自動化にさらなるメリットをもたらします。分析を実行するには、少なくとも 1 つの SPSS Modeler をインストールしておく必要があります。

## IBM SPSS Modeler のオプション

次のコンポーネントおよび機能を個別に購入し、ライセンス供与を受け SPSS Modeler と合わせて使用できます。追加の製品や更新が利用できる可能性があることに注意してください。 詳細は、<http://www.ibm.com/software/analytics/spss/products/modeler/> を参照してください。

- SPSS Modeler Server へのアクセスにより、大規模なデータセット上のスケーラビリティおよびパフォーマンスを向上させ、SQL 最適化のサポート、およびインデータベース モデリング能力を提供します。
- SPSS Modeler Solution Publisher は、SPSS Modeler 環境の外側でのリアルタイムまたは自動スコアリングで使用します。 詳細は、2 章 [IBM SPSS Modeler Solution Publisher in IBM SPSS Modeler 14.2 Solution Publisher](#) を参照してください。
- アダプタを使用して IBM SPSS Collaboration and Deployment Services またはシンククライアント アプリケーションの IBM SPSS Modeler Advantage に展開します。 詳細は、9 章 [IBM SPSS Collaboration and Deployment Services Repository オブジェクトの保存と展開 in IBM SPSS Modeler 14.2 ユーザー ガイド](#) を参照してください。

## IBM SPSS Text Analytics

IBM® SPSS® Text Analytics は、SPSS Modeler の完全に統合されたアドオンで、高度な言語テクノロジーと自然言語処理（NLP: Natural Language Processing）を使用して、さまざまな無構造テキスト データを高速で処理し、重要な概念を抽出および整理してカテゴリに分類します。抽出されたコンセプトとカテゴリを、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。

- テキスト マイニング ノードは、テキスト リンクとクラスタの高度な洞察を実行できるインタラクティブ ワークベンチと同様にコンセプトおよびカテゴリ モデル作成を提供、独自のカテゴリを作成し。言語リソースのテンプレートを調整します。

- ブログやその他の Web ベースのソースなど、多くのインポート形式がサポートされています。
- CRM (Customer Relationship Management、顧客関係管理) やゲノム研究のような特定の分野用のカスタム テンプレート、ライブラリ、辞書も付属しています。

注 : このコンポーネントを利用するには、別途、ライセンスが必要です。詳細は、<http://www.ibm.com/software/analytics/spss/products/modeler/>を参照してください。

## IBM SPSS Modeler ドキュメント

オンライン ヘルプ形式の完全なドキュメントは、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は、各製品 DVD の ¥Documentation フォルダにもあります。

- **IBM SPSS Modeler ユーザー ガイド:** SPSS Modeler の使用方法への全体的な入門で、データ ストリームの構築方法、欠損地の処理方法、CLEM 式の処理方法、プロジェクトおよびレポートの処理方法、IBM SPSS Collaboration and Deployment Services、予測アプリケーション製品、または IBM SPSS Modeler Advantage へ展開するストリームのパッケージ化方法が含まれています。
- **IBM SPSS Modeler 入力ノード、プロセス ノード、出力ノード:** さまざまな形式のデータを読み込み、処理し、出力するために使用するすべてのノードの説明があります。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の説明。IBM® SPSS® Modeler には、マシン学習、人工知能、および統計に基づいたさまざまなモデル作成手法が用意されています。詳細は、[3 章 モデル作成ノードの概要 in IBM SPSS Modeler 14.2 Modeling Nodes](#) を参照してください。
- **IBM SPSS Modeler アルゴリズム ガイド:** 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。
- **IBM SPSS Modeler アプリケーション ガイド:** 本ガイドの例では、特定のモデル作成手法および技術に関する簡単で、目的に沿った説明を行います。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。詳細は、[アプリケーションの例 in IBM SPSS Modeler 14.2 ユーザー ガイド](#) 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:** スクリプトの実行によるシステムのオートメーションの情報で、ノードおよびストリームを操作するために使用することができるプロパティが含まれています。
- **IBM SPSS Modeler 展開ガイド:**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、 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:** ユーザーのデータベースを最大限に活用して、パフォーマンスを改善する方法と、サードパーティー製のアルゴリズムを使用して分析可能な範囲を拡大する方法についての情報があります。
- **IBM SPSS Modeler Server 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 管理コンソール ユーザー ガイド:**SPSS Modeler Server を監視して設定するためのコンソール ユーザー インターフェイ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 Solution Publisherガイド:** SPSS Modeler Solution Publisher はアドオン コンポーネントです。組織はこれを使用すると、標準的な SPSS Modeler 環境の外部へストリームを公開できます。
- **IBM SPSS Modeler CRISP-DM Guide.** CRISP-DM 手法を使用した SPSS Modeler によるデータ マイニングの段階的なガイドです。

## アプリケーションの例

SPSS Modeler のデータ マイニング ツールは、多様なビジネスおよび組織の問題解決を支援しますが、アプリケーションの例では、特定のモデル作成手法および技術に関する簡単で、目的に沿った説明を行います。ここで使用されるデータセットは、データ マイニング作業によって管理された巨大なデータ 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 ディレクトリの Demos フォルダにインストールされています。詳細は、[Demos フォルダ in IBM SPSS Modeler 14.2 ユーザー ガイド](#) を参照してください。

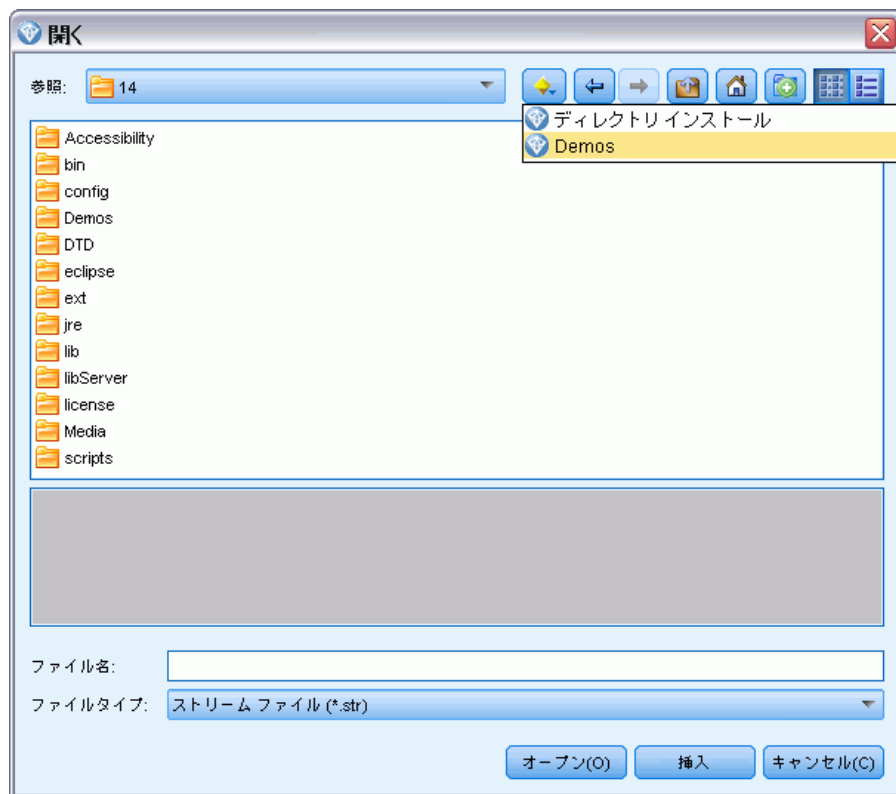
**データベース モデル作成の例:** 例は、『IBM SPSS Modeler データベース内マイニング ガイド』を参照してください。

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

## Demos フォルダ

アプリケーションの例で使用されるデータ ファイルとサンプル ストリームは、製品のインストール ディレクトリの Demos フォルダにインストールされています。このフォルダには、Windows [スタート] メニューの IBM SPSS Modeler 14.2 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の一覧から [Demos] を選択してアクセスすることもできます。

図 1-1  
最近使用されたディレクトリの一覧から Demos フォルダを選択



# アーキテクチャとハードウェアに関する推奨事項

## IBM SPSS Modeler のアーキテクチャ

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 のアーキテクチャを説明します。最適なパフォーマンスを達成するために 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 ドライバのセットアップ場所も説明します。

## アーキテクチャの説明

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

- SPSS Modeler Server は、インストール先の決定で柔軟性があり、Windows に加えて、UNIX でも実行できます。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 コンピュータであることに専念できます。
- 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、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、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® Collaboration and Deployment Services Deployment Manager (SPSS Modeler Server に付属) や IBM® SPSS® Collaboration and Deployment Services および IBM® SPSS® Collaboration and

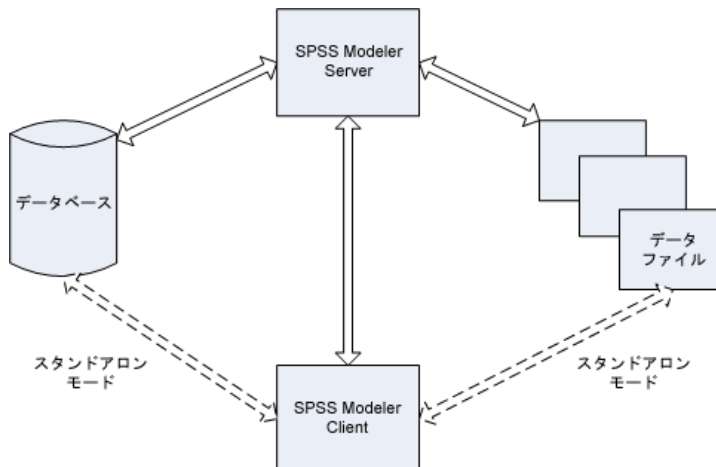
Deployment Services Repository (別売) のような管理ツールを使用すると、データマイニングのプロセスを監視でき、適切なコンピューティングリソースを確実に利用できます。これで、一定のデータマイニングタスクを自動化し、データモデルへのアクセスを管理し、組織全体で結果を共有することができます。

IBM® SPSS® Modeler の分散型アーキテクチャのコンポーネントは、次に示されています。“IBM SPSS Modeler Server のアーキテクチャ” の図 ( p.7 )。

- **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イスが提供され、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 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 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。SPSS Modeler は、Windows オペレーティングシステムでのみ動作します。
- **SPSS Modeler Server**: 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にインストールされたサーバーソフトウェア。SPSS Modeler Server はサービス (Windows) またはデーモンプロセス (UNIX) として実行されます。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、活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) であってもよいし、ほかの運用システムへの影響を減らすために、ローカルまたは部門別のサーバー (Windows 上の SQL Server など) 上のデータマートであってもかまいません。

## IBM SPSS Modeler Server のアーキテクチャ

図 2-1  
IBM SPSS Modeler Server のアーキテクチャ



分散型アーキテクチャでは、ほとんどの処理がサーバー コンピュータで発生します。エンド ユーザーがストリームを実行すると、SPSS Modeler により、ストリームの記述がサーバーへ送られます。サーバーでどの操作が SQL で実行可能かが判断され、適切なクエリーが作成されます。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。

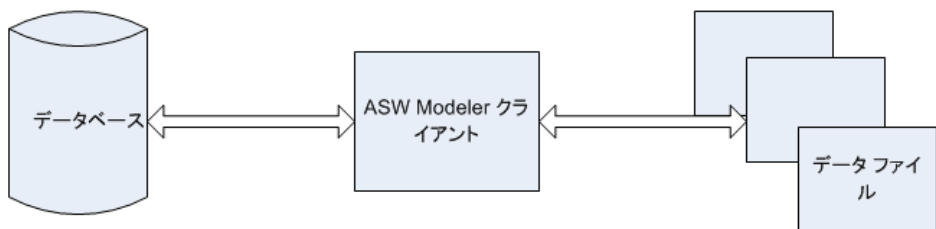
必要な場合には、SPSS Modeler Server により、すべての SPSS Modeler 操作がデータベースの外側で実行されます。操作のためのデータを保持するために、RAM とディスク メモリーとの使用率のバランスが自動的にとられます。このプロセスで、SPSS Modeler Server はフラット ファイルとの完全な互換性がとられます。

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。詳細は、D 付録 p.91 サーバー クラスタでの負荷バランシングを参照してください。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に接続できます。詳細は、3 章 IBM SPSS Modeler Server への接続 in IBM SPSS Modeler 14.2 ユーザー ガイド を参照してください。

## スタンドアロン クライアント

SPSS Modeler は、次に示すように、デスクトップ内蔵型の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。下の “IBM SPSS Modeler スタンドアロン” の図 . 詳細は、3 章 p.16 IBM SPSS Modeler サポート を参照してください。

図 2-2  
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。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** :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 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** :SPSS Modeler Server をデータベースと同じコンピュータ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に分けることです。2 つのコンピュータ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、プロセッサ数が増えるほど、パフォーマンスがよくなります。

- 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、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（たとえば、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 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あります。

- レコード集計ノード
- 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。たとえば、データ ファイルが 2GB で SQL 生成が使用されない場合、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`（レコード当たりのセル数）は、名義型フィールドがある場合に、非常に大きくなります。

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 ドライバが含まれています。また、これらのドライバは、本リリースに付属する IBM SPSS Data Access Pack のインストール ディスクから入手できます。ODBC データ ソースに関する質問がある場合は、データベース管理者に問い合わせてください。

IBM® SPSS® Modeler のデータベースのサポートは 3 つの段階に分類され、それぞれデータベース ベンダーに応じて SQL プッシュバックおよび最適化のサポートについて異なるレベルを示しています。各レベルのサポートは、さまざまなシステム設定を使用して実装され、サービス契約の一環としてカスタマイズすることができます。

データベース サポートの 3 つの段階は次のようになります。

テーブル 2-1  
データベース サポートの段階

| サポートの段階 | 説明   |
|---------|--|
| 段階 1    | すべての SQL プッシュバックを使用でき、データベース固有の SQL 最適化を行います。        |
| 段階 2    | 多くの SQL プッシュバックを使用でき、データベース固有でない SQL 最適化を行います。       |
| 段階 3    | SQL プッシュバックを使用できず、最適化も行いません。データベースの読み込み、および書き込みのみです。 |

### サポートされている ODBC ドライバ

SPSS Modeler 14.2 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。

## ドライバをインストールする場所

ODBC ドライバは、処理が行われる各コンピュータ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、ドライバをローカル コンピュータにインストールする必要があります。
- 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 ドライバは SPSS Modeler Server がインストールされたコンピュータにインストールする必要があります。
- SPSS Modeler と 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 ドライバは両方のコンピュータ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 SPSS Modeler を実行する場合、ODBC ドライバは 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。
- 公開されたストリームを別のコンピュータで実行するために IBM® SPSS® Modeler Solution Publisher ランタイムを使用する場合、そのコンピュータに ODBC ドライバをインストールして設定する必要があります。

注：UNIX で SPSS Modeler Server を使用して Teradata データベースを使用する場合、Teradata ODBC ドライバとともにインストールされた ODBC ドライバマネージャを使用する必要があります。この変更を SPSS Modeler Server に行うには、コメントで示される `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 ドライバに付属する ODBC ドライバ マネージャの場所（Teradata ODBC ドライバのデフォルト 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、SPSS Modeler Server を再起動する必要があります。Teradata アクセスに対応する SPSS Modeler Server プラットフォーム、サポートされている Teradata ODBC のバージョンの詳細は、サポート サイト <http://www.ibm.com/support> を参照してください。

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、SPSS Modeler 内の [ディレクトリの設定] 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が影響を受けるわけではありません。

**UNIX** :UNIX 上で IBM® SPSS® Statistics データ アクセス テクノロジーとともに動作する SPSS Modeler Server の構成方法は、「[UNIX 起動スクリプトの構成](#)」（ p. 76 ）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にデータを格納する場合は、サーバー コンピュータ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（たとえば 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（たとえば \\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、完全ファイル指定とスラッシュ（斜線）を使用します（たとえば、/public/data/ServerData/Sales 1998.csv）。SPSS Modeler Server とともに使用するデータの UNIX ディレクトリとファイル名には、円記号（¥）の使用を避けてください。テキスト ファイルに UNIX または DOS フォーマットを使用するかどうかは（両方とも自動的に処理される）、問題ではありません。

## 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 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。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 SPSS Statistics ノードでは、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、SPSS Statistics クライアントが 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。

# IBM SPSS Modeler 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
- 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
- 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

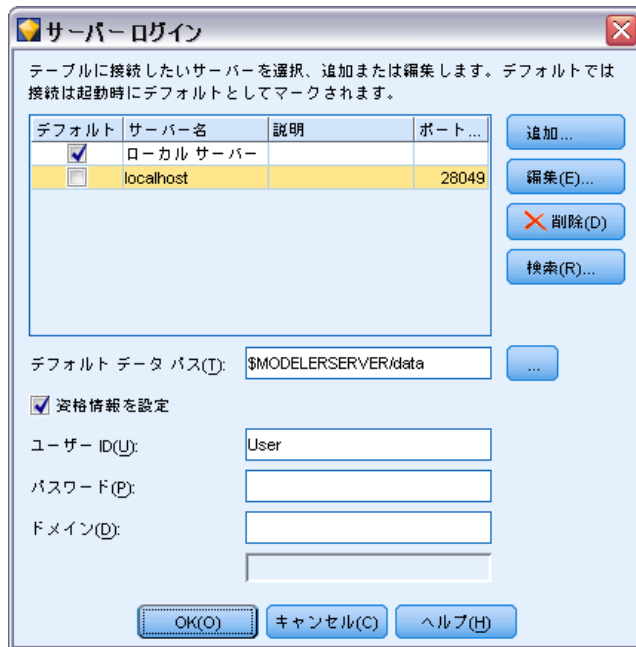
## IBM SPSS Modeler Server への接続

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 SPSS Modeler Server またはサーバー クラスタ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、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 を使用する場合は、[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。詳細は、D 付録 p.91 **サーバー クラスタでの負荷バランシング** を参照してください。



図 3-1  
[サーバー ログイン] ダイアログ ボックス



### サーバーに接続するには

- ▶ [ツール] メニューの [サーバーへのログイン] をクリックします。[サーバーへのログイン] ダイアログ ボックスが開きます。または、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。
- ▶ 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。  
詳細は、[IBM SPSS Modeler Server 接続の追加および編集 in IBM SPSS Modeler 14.2 ユーザー ガイド](#) 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 クラスタにアクセスします。詳細は、[IBM SPSS Collaboration and Deployment Services のサーバーの検索 in IBM SPSS Modeler 14.2 ユーザー ガイド](#) 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、定義されたサーバー接続のセットが含まれています。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 コンピュータ上のデータへのパスを指定します。[...] 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を解除すると、**シングル サインオン**機能を有効にし、ローカル コンピュータのユーザー名とパスワードの詳細を入力してサーバーにログインします。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（たとえば、管理者アカウントにログインした場合）、資格情報を入力するための次のフィールドが表示されます。

**ユーザー ID：** サーバーにログインするユーザー名を入力します。

**パスワード：** 指定したユーザー名に対応するパスワードを入力します。

**ドメイン：** サーバーにログインする際に使用するドメイン名を指定します。サーバー コンピュータが クライアント コンピュータとは異なる Windows ドメインにある場合にのみ、ドメイン名が必要です。

- ▶ [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

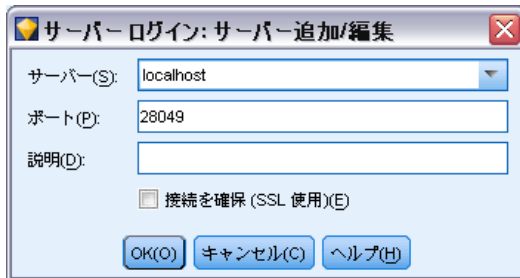
- ▶ [ツール] メニューの [サーバーへのログイン] をクリックします。[サーバーへのログイン] ダイアログ ボックスが開きます。または、SPSS Modeler ウィンドウの接続ステータス領域をダブル クリックします。
- ▶ ダイアログ ボックスで、[ローカル サーバー] を選択し、[OK] をクリックします。

## IBM SPSS Modeler Server 接続の追加および編集

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

注： IBM® SPSS® Collaboration and Deployment Services から追加されたサーバー接続は、名前、ポート、およびそのほかの詳細が IBM SPSS Collaboration and Deployment Services で定義されているため、編集することができません。

図 3-2  
[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



### サーバー接続を追加するには

- ▶ [ツール] メニューの [サーバーへのログイン] をクリックします。[サーバーへのログイン] ダイアログ ボックスが開きます。
- ▶ 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。
- ▶ 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 コンピュータは、英数字の名前（たとえば、myserver）、または、サーバー コンピュータに割り当てられた IP アドレス（たとえば、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。
  - **接続を確保 (SSL 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

- ▶ [ツール] メニューの [サーバーへのログイン] をクリックします。[サーバーへのログイン] ダイアログ ボックスが開きます。
- ▶ 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。

- ▶ 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 クラスタを選択できます。サーバー クラスタは、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。詳細は、D 付録 p.91 [サーバー クラスタでの負荷バランシング](#) を参照してください。

[サーバーへのログイン] ダイアログ ボックスでサーバーを手動で追加できますが、使用できるサーバーを検索して、正しいサーバー名やポート番号を知らなくてもサーバーに接続することができます。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。詳細は、[IBM SPSS Modeler Server 接続の追加および編集 in IBM SPSS Modeler 14.2 ユーザー ガイド](#) を参照してください。

図 3-3  
[サーバーの検索] ダイアログ ボックス



### サーバーおよびクラスタを検索するには

- ▶ [ツール] メニューの [サーバーへのログイン] をクリックします。[サーバーへのログイン] ダイアログ ボックスが開きます。
- ▶ 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。 [詳細は、9 章 IBM SPSS Collaboration and Deployment Services Repository への接続 in IBM SPSS Modeler 14.2 ユーザー ガイド を参照してください。](#)

- ▶ リストからサーバーまたはサーバー クラスタを選択します。
- ▶ [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**: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 ディレクトリです)。

## ユーザーの認証

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 コンピュータのオペレーティング システムを使用します。ユーザーが 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 SPSS Modeler Server にログオンすることはできません。

**UNIX**:デフォルトでは、SPSS Modeler Server は UNIX 上でルートとして実行されると想定されます。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 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ます、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）。詳細は、C 付録 p.87 非 root プロセスとしての IBM SPSS Modeler Server の構成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、SPSS Modeler Server は PAM を使用して認証を行います。必要な場合は、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。詳細は、2 章 p.10 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**:SPSS Modeler Server の UNIX バージョンでは、一時ファイルへの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。詳細は、B 付録 p.78 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 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。

**丸めの相違:** 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、SPSS Modeler Server とは別のフォーマットが使用されます。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。

# IBM SPSS Modeler Server 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® Collaboration and Deployment Services Deployment Manager 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 CD に入っている 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 の起動と停止

IBM® SPSS® Modeler Server は、Windows 上のサービスとして、または UNIX および。

スケジューリング上の注意 :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。

- ▶ **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。

- ▶ [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 ディレクトリで `modelersrv.sh` スクリプトを実行して、SPSS Modeler Server を起動または停止します。

- ▶ SPSS Modeler Server のインストール ディレクトリに変更します。たとえば、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、SPSS Modeler Server のインストール先ディレクトリです。

- ▶ 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
```

- ▶ 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
```

- ▶ 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 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。

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。詳細は、[B 付録 p.77 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、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、SPSS Modeler Server 環境内では制限なしに設定する必要があります。このためには、次の手順を実行します。

- ▶ 次のコマンドを `modelersrv.sh` に追加します。

```
ulimit -d unlimited
```

```
ulimit -m unlimited
```

```
ulimit -f unlimited
```

- ▶ 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 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 プロセスの集積の影響を受けやすいシステムの場合、次の一連のコマンドを使用して、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。

```
cdmodeler_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 SPSS Modeler プロセスにより（一時ディレクトリの）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、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、SPSS Modeler 一時ディレクトリ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、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® Collaboration and Deployment Services Deployment Manager アプリケーションを使用します。詳細は、[p.27 IBM SPSS Modeler Administration Console の使用方法](#) 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 ディレクトリにある options.cfg テキスト ファイルを使用します。詳細は、[p.39 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 Deployment Manager 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 SPSS Modeler Server のファイル システムにアクセスする必要がありますが、Deployment Manager 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。その上、Deployment Manager 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。さらに、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、SPSS Modeler Server を再起動しなくても変更できます。

Deployment Manager 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 の使用方法

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 インターフェイ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、モニタリングできる共有のグラフィカル インターフェイスを提供します。

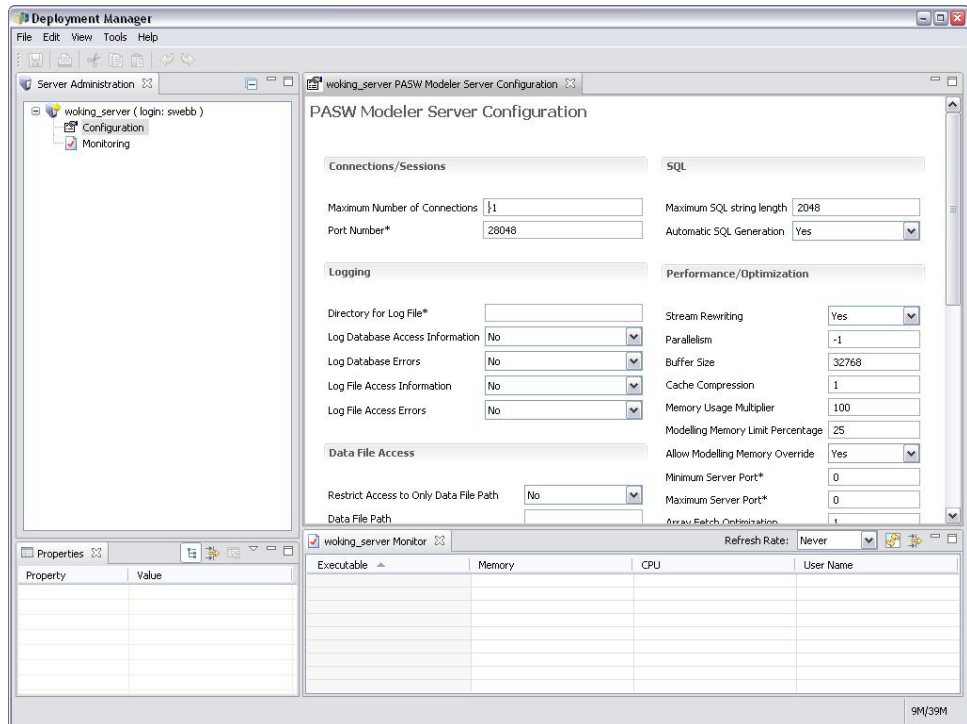
### Modeler Administration Console の起動

- ▶ Windows の [スタート] メニューから、[すべてのプログラム] → [IBM SPSS Collaboration and Deployment Services] → [Deployment Manager] → [IBM SPSS Modeler 管理コンソール <nn.n>] を選択します。

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。

図 4-1

Modeler Administration Console インターフェイス



## Modeler Administration Console でのアクセスの構成

Modeler Administration Console を介した SPSS Modeler Server への管理者アクセスは、/config の下の SPSS Modeler Server インストールディレクトリにある options.cfg ファイルの **administrators** の行で制御されています。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、エントリは二重引用符で囲む必要があります。エントリでは、大文字小文字を区別します。
- 複数のユーザー ID はカンマで区切ります。
- 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 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、ログインの資格証明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

- ▶ 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。
- ▶ [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。
- ▶ 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、たとえば「Production Server」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。
- ▶ 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。[完了] ボタンをクリックします。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

- ▶ [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。

- ▶ [ログイン] ダイアログ ボックスで、資格証明を入力します。(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。詳細は、[p.28 Modeler Administration Console](#) での[アクセスの構成](#)を参照してください。

ログインに失敗して、メッセージ [サーバー '<server>' への接続に失敗しました] が表示された場合、ユーザー ID とパスワードが正しいことを確認し、SPSS Modeler Server サービスが稼動していることを確認します。たとえば、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、ログインを再実行します。

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] ウィンドウ枠には、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。各オプションのかっこ内に、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

図 4-2  
接続/セッションの設定

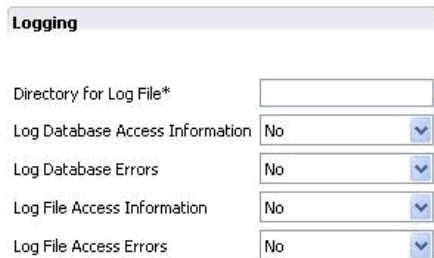
| Connections/Sessions          |       |
|-------------------------------|-------|
| Maximum Number of Connections | -1    |
| Port Number*                  | 28047 |

**接続の最大数：** (max\_sessions) 一度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。

**ポート番号 :**(port\_number) SPSS Modeler Server が待機するポート番号。別のアプリケーションがすでにデフォルト (28047) を使用している場合は、これを変更します。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。

## ログ

図 4-3  
ログ記録の設定



| Logging                         |                      |
|---------------------------------|----------------------|
| Directory for Log File*         | <input type="text"/> |
| Log Database Access Information | No                   |
| Log Database Errors             | No                   |
| Log File Access Information     | No                   |
| Log File Access Errors          | No                   |

**ログ ファイルのディレクトリ :** (log\_directory) messages.log ファイルの場所。空白の文字列は、[サーバーのインストール先ディレクトリ]/log という意味です。スラッシュ記号 (斜線) を使用する必要があります。このログ ファイルの内容の詳細は、「[サーバー ログ](#)」 ( p. 48 ) を参照してください。

**データベース アクセス情報をログに記録する :** (log\_database\_access) データベース アクセス情報をログ ファイルへ追加します。

**データベース エラーをログに記録する :** (log\_database\_errors) データベース エラーをログ ファイルへ追加します。

**ファイル アクセス情報をログに記録する :** (log\_file\_access) ファイル アクセス情報をログ ファイルへ追加します。

**ファイル アクセス エラーをログに記録する :** (log\_file\_errors) ファイル アクセス情報をログ ファイルへ追加します。

## データ ファイル アクセス

図 4-4  
データ ファイル アクセスの設定

| Data File Access                           |     |
|--|-----|
| Restrict Access to Only Data File Path     | No  |
| Data File Path                             |     |
| Restrict Access to Only Program Files Path | Yes |
| Program Files Path                         |     |
| Maximum File Size (MB)                     | -1  |
| Temporary Directory                        |     |
| Python Exe Path                            |     |

**データ ファイル パスのみにアクセスを制限：** (`data_files_restricted`) `yes` に設定すると、このオプションで、標準のデータ ディレクトリと下の [データ ファイル パス] に表示されたディレクトリのデータ ファイルにだけ、アクセスが制限されます。

**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ディレクトリのリスト。[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かぎり、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。たとえば [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`) [はい] に設定すると、このオプションで、標準の bin ディレクトリと下の [プログラム ファイル パス] に表示されたディレクトリのプログラム ファイルにだけ、アクセスが制限されます。

**プログラム ファイル パス：** (`program_file_path`) クライアントがここからプログラムを実行できる、追加ディレクトリのリスト。[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かぎり、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を指定します。

**最大ファイル サイズ (MB)：** (`max_file_size`) ストリーム実行時に作成された一時データ 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 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：** (`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。理想的には、このディレクトリ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。カンマで区切って、複数の一時ディレクトリを指定することもできます。複数のディレクトリは、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が最も頻繁に使用され、追加の各ディレクトリは、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、このディレクトリに一時ファイルが生成されます。このディレクトリに対する必要なアクセス権限があることを確認してください（たとえば、一時ディレクトリが共有ネットワーク フォルダである場合）。権限がない場合、SPSS Modeler Server の起動に失敗します。

**Python 実行パス：** (`python_exe_path`) 実行ファイル名を含む Python 実行ファイルへの完全パス。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ります。

## パフォーマンス/最適化

図 4-5  
パフォーマンス/最適化の設定

| Performance/Optimization          |       |
|-----------------------------------|-------|
| Stream Rewriting                  | Yes   |
| Parallelism                       | -1    |
| Buffer Size                       | 32768 |
| Cache Compression                 | 1     |
| Memory Usage Multiplier           | 100   |
| Modelling Memory Limit Percentage | 25    |
| Allow Modelling Memory Override   | Yes   |
| Minimum Server Port               | 0     |
| Maximum Server Port               | 0     |
| Array Fetch Optimization          | 1     |

**ストリームのリライト：** (`stream_rewriting_enabled`) サーバーがストリームをリライトすることで、ストリームを最適化できるようにします。たとえば、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。この設定により、対応するクライアントの最適化設定が上書きされます。

**並行処理：** (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。0 または負の数に設定すると、IBM® SPSS® Modeler はスレッド数をコンピュータ上の使用できるプロセッサの数に一致させます。このオプションのデフォルト値は -1 です。並行処理を無効にするには（複数のプロセッサを持つコンピュータの場合）、このオプションを 1 に設定します。制限付き並行処理を有効にするには、コンピュータ上のプロセッサの数より小さい値に設定します。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へ転送されたデータ ファイルは、このバイト数のバッファを通過します。

**キャッシュの圧縮：** (cache\_compression) キャッシュと、サーバーの一時ディレクトリ内のほかのファイルの圧縮を制御する 0 から 9 の範囲内の整数。圧縮により使用され、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。値が 0 の場合は、圧縮機能がすべて無効化されます。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。これより上の値は、ディスク容量が非常に高額な場合を除いて、稀にしか必要ありません。

**メモリー使用率の乗数：** (memory\_usage) ソートやその他のメモリー内キャッシュに割り当て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。未使用のメモリーが利用できる場合にソート パフォーマンスを向上させるためにこの値を増やしますが、この値が増えると過度のページングの原因となるので、注意が必要です。

**モデル作成のメモリー制限パーセント：** (modelling\_memory\_limit\_percentage) ニューラル、Kohonen、および K-Means の各モデルの学習に割り当て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。未使用のメモリーが利用できる場合にソート パフォーマンスを向上させるためにこの値を増やしますが、データがディスク上にこぼれだす場合にこの値が増えると過度のページングの原因となるので、注意が必要です。

**モデル作成メモリーの上書きの許可：** (allow\_modelling\_memory\_override) 一定のモデル作成ノード内の [速度優先で最適化] オプションを有効または無効にします。デフォルトは、有効化です。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 コンピュータ上で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できます。

**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実行のために必要なクライアントとサーバーの間の追加ソケット接続のため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。両方のデフォルト値は、「制限なし」の意味の -1 です。このために、たとえば、サーバーをポート 8000 を大気するように設定するには、min\_server\_port を 8000 に設定し、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、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（たと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、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コントロールします。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

図 4-6  
SQL の設定

The image shows a dialog box titled "SQL". It contains two settings:

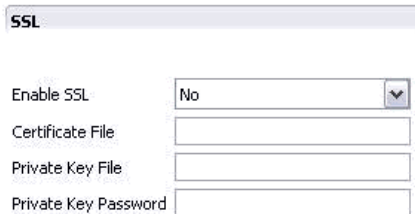
- "Maximum SQL string length" is set to "2048".
- "Automatic SQL Generation" is set to "Yes" (indicated by a downward arrow in a dropdown menu).

**SQL 文字列の最大の長さ：** (max\_sql\_string\_length) データベースから SQL でインポートされる文字列の最大の長さ。これより長い文字列の値は、警告なしに右側で切り捨てられます。有効な範囲は 1 から 65,535 文字です。

**自動 SQL 生成：**(`sql_generation_enabled`) 以後のパフォーマンス向上の可能性のあるストリームに対し、自動的な SQL 生成を許可します。デフォルトは、有効化です。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、このオプションの有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。

## SSL

図 4-7  
SSL 設定



| SSL                  |    |
|----------------------|----|
| Enable SSL           | No |
| Certificate File     |    |
| Private Key File     |    |
| Private Key Password |    |

**SSL の有効化：**(`ssl_enabled`) SPSS Modeler と SPSS Modeler Server の接続のために SSL 暗号化を有効にします。

**証明書ファイル：**(`ssl_certificate_file`) サーバーの起動時にロードされる公開 SSL 証明書ファイル。

**プライベートキー ファイル：**(`ssl_private_key_file`) サーバーの起動時にロードされるプライベート SSL キー ファイル。

**プライベートキー 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

図 4-8  
Coordinator of Processes の構成設定

| Co-ordinator Of Processes Configuration |         |
|---|---------|
| Host*                                   | spsscop |
| Port Number*                            | 8080    |
| Login Name*                             |         |
| Password*                               |         |
| Enabled*                                | No ▼    |
| Server Name*                            |         |
| Description*                            |         |
| Update Interval (min)*                  | 2       |
| Weight                                  | 1       |
| Service Host*                           |         |

Items marked (\*) require a server restart to take effect.

**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「spsscop」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。

**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。

**ログイン名:** (`cop_user_name`) Coordinator of Processes サービスへの認証用のユーザー名。これは IBM SPSS Collaboration and Deployment Services ログイン名で、セキュリティ プロバイダの接頭辞が含まれます (例: ad/jsmith)。

**パスワード:** (`cop_password`) Coordinator of Processes サービスへの認証用の暗号化されたパスワード。

パスワード ファイルを作成するには、SPSS Modeler Server のインストール先の bin ディレクトリ内にある、pwwutil というパスワード ユーティリティ プログラムを使用する必要があります。

**有効:** (`cop_enabled`) サーバーを Coordinator of Processes で登録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です。デフォルトはホスト名です。

**説明:** (`cop_service_description`) インスタンスの説明。

**更新間隔 (最小):** (`cop_update_interval`) メッセージが有効な時間 (分) です。デフォルトは 2 です。

**重み:** (`cop_service_weight`) インスタンスの重みで、1 から 10 の整数で指定されます。重みの数値が高いほど、接続が多くなります。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 のみで表示されるオプション

**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、このパラメータをこれらの文字列値に対応した値に設定します。

**max\_transfer\_size:** 内部システムの使用のみ。変更しないでください。

**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、**true** に設定します。このオプションを **true** に設定すると、子プロセスで SPSS Modeler Server ログインの詳細を使用できます。

**administrators:** 管理者アクセス権を付与するユーザーのユーザー名を指定します。詳細は、[p.28 Modeler Administration Console](#) での [アクセスの構成](#) を参照してください。

**supported\_sql\_type\_logging:** デバッグ用のみ。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。

**allow\_config\_custom\_overrides:** 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。

## 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。

図 4-9

IBM SPSS Modeler Administration Console の監視用のウィンドウ枠

| Executable       | Memory   | CPU | User Name |
|------------------|----------|-----|-----------|
| FrameworkService | 13,036 K | 00  | SYSTEM    |
| LocalSch         | 1,640 K  | 00  | SYSTEM    |
| Mcshield         | 23,364 K | 00  | SYSTEM    |
| SoftMon          | 1,632 K  | 00  | SYSTEM    |
| System           | 176 K    | 00  | SYSTEM    |
| UpdaterUI        | 2,156 K  | 00  | cstrahs   |
| VsTskMgr         | 464 K    | 00  | SYSTEM    |
| clemserv         | 10,452 K | 00  | SYSTEM    |
| clemserv         | 22,312 K | 00  | SYSTEM    |

## 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 ディレクトリにあります。設定は、カ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： 次を例外として、すべての構成オプションは、この構成ファイルでなく、IBM® SPSS® Collaboration and Deployment Services Deployment Manager を使用して変更できます。

- administrators
- allow\_custom\_config\_overrides
- max\_transfer\_size
- start\_process\_as\_login\_user
- supported\_sql\_type\_logging

Deployment Manager 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。 [詳細は、p. 27 IBM SPSS Modeler Administration Console の使用方法 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

- ▶ テキスト エディタを使用して、`options.cfg` ファイルを開きます。
- ▶ 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、「[SPSS Modeler Server の構成](#)」（p. 30）を参照してください。
- ▶ 必要に応じて値を編集します。パラメータ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。
- ▶ ファイルを保存します。
- ▶ 変更が効力を発揮するように、SPSS Modeler Server を停止し、その後再起動します。詳細は、[p. 24 IBM SPSS Modeler Server の起動と停止](#)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 SPSS Modeler がデータベースに再接続する必要なくなるため、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 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 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、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 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

- ▶ [サーバー インストール パス]/`config` ディレクトリを検索します。



- ▶ 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
```

- ▶ 次の行をファイルに追加します。

```
cache_connection, N
```

- ▶ ファイルを保存して閉じます。
- ▶ IBM® SPSS® Modeler Server を起動して、変更が有効になるようにします。

## IBM SPSS Statistics ライセンス ロケーション

IBM® SPSS® Modeler が Statistics 変換ノード、Statistics モデル ノードおよび Statistics 出力ノードを使用できるようにするには、ストリームが実行されているコンピュータに IBM® SPSS® Statistics をインストールしライセンスされている必要があります。サーバーに対して実行するとき、ライセンスの構成はそのサーバー上で行う必要があります。

### Windows

- ▶ コマンド プロンプトから、IBM® SPSS® Modeler Server の bin ディレクトリに変更して、以下を実行します。

```
statisticsutility -location =<path to IBM SPSS Statistics Server license file>
```

この場合、<path to SPSS Statistics Server license file> は、ライセンスが与えられた SPSS Statistics サーバーのインストール ディレクトリを示します。

### UNIX

- ▶ コマンド プロンプトから、SPSS Modeler Server の bin ディレクトリに変更して、以下を実行します。

```
./statisticsutility -location =<path to IBM SPSS Statistics Server license file>
```

この場合、<path to SPSS Statistics Server license file> は、ライセンスが与えられた SPSS Statistics サーバーのインストール ディレクトリを示します。

## 安全にデータを転送するための SSL の使用

Security Sockets Layer (SSL) は、2 つのコンピュータ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。

- ▶ 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。
- ▶ 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。
- ▶ 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。
- ▶ 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 によってクライアント-サーバー間およびサーバー-サーバー間の通信を保護する主な手順は次のとおりです。

- ▶ サーバー コンピュータに OpenSSL をインストールします。
- ▶ SSL 証明書とキーを取得およびインストールします。
- ▶ サーバー管理アプリケーションで SSL を有効化し、設定します (IBM® SPSS® Collaboration and Deployment Services Deployment Manager)。
- ▶ 必要に応じて、クライアント コンピュータに無限強度の暗号をインストールします。
- ▶ 自己署名証明書を使用する場合は、クライアント コンピュータの証明書をコピーします。
- ▶ 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。
- ▶ エンド 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。たとえば、IBM® SPSS® Collaboration and Deployment Services Repository に接続する IBM® SPSS® Statistics Server などです。この場合、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。

- ▶ <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。

テーブル 4-1  
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 17-19 (System z® の Linux® 以外)<br>IBM® SPSS® Modeler 14.2 (System z® の Linux® 以外) | 0.9.8 およびサブバージョン (0.9.8a、0.9.8b、など) |
| SPSS Statistics 19 (System z® の Linux®)<br>SPSS Modeler 14.2 (System z® の Linux®)                      | 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、SPSS Modeler Server が期待する xxx.0.9.8 バージョンのライブラリへのシンボリック リンクを作成する必要があります。

- ▶ 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) と共有ライブラリをサポートする必要があります。

- ▶ システムが SSL ライブラリ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 ライブラリを `modelersrv.sh` で定義されたロード ライブラリパスの終わりに追加します。たとえば、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

- ▶ 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に署名します。
  - 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 には、この目的のための証明書管理ツールがあります。または、自己署名 SSL 証明書作成に関する指示をインターネットで検索できます。
- ▶ 証明書およびキー ファイルを、サーバー上のローカル ディレクトリにコピーします。公開キーとプライベート キーを 1 つのファイルに保存することができます。アクセス権限を使用して、このディレクトリの不正な参照を拒否することができます。
- ▶ アプリケーション サーバーに SSL 証明書およびキーをインストールします。キーおよび証明書を特定のアプリケーション サーバーと同時に操作する方法の詳細は、発売元のベンダーのマニュアルを参照してください。証明書とキーを Java キーストアに追加する必要があります。

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

- ▶ 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。
- ▶ 設定ページで、[Secure Sockets Layer] を [はい] に設定します。
- ▶ [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。
- ▶ [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。

- ▶ メニューから次の項目を選択します。  
File > [保存(S)]
- ▶ 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せん。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 の SSL の設定 (UNIX システムのみ)

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、UNIX システムで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` パスワードを使用して、`options.cfg` ファイルで指定する。 [詳細は、 p. 36 SSL を参照してください。](#)
- 起動時にパスワードを要求するよう、スイッチを `modelersrv.sh` ファイルに追加する。

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す。データのセキュリティを強化するために、弊社では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。

- ▶ <http://java.sun.com/javase/downloads/index.jsp> から Java Cryptography Extension (JCE) Unlimited Strength Jurisdiction Policy Files 5.0 をダウンロードします。
- ▶ ダウンロードしたファイルを解凍します。
- ▶ 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です。

## 証明書のクライアント キーストアへの追加 (リポジトリまたは IBM SPSS Modeler Serverへの接続)

注： 認証機関に署名された証明書を使用している場合は、この手順を省略してください。

SSL を使用して IBM® SPSS® Collaboration and Deployment Services リポジトリまたは IBM® SPSS® Modeler Serverに接続し、自己署名証明書を使用する場合、証明書をクライアントの Java キーストアに追加する必要があります。クライアント コンピュータで、次の手順を実行します。

サーバー コンピュータがクライアントとして機能する場合がありますので、注意してください。たとえば、IBM® SPSS® Collaboration and Deployment Services Repository に接続する IBM® SPSS® Statistics Server などです。この場合、SPSS Statistics Server はクライアントとなるため、IBM SPSS Collaboration and Deployment Services Repository Server の証明書を SPSS Statistics Server にコピーする必要があります。

- ▶ 証明書 (.cer) ファイルを SSL サーバーからクライアント コンピュータの <install\_dir>%jre%lib%security ディレクトリにコピーします。
- ▶ 同じディレクトリで、cacerts ファイルを右クリックして [プロパティ] を選択します。
- ▶ [読み取り専用] チェック ボックスをオフにします。
- ▶ コマンド プロンプトを開いて、ディレクトリを次の場所に変更します。<product install dir> は、製品をインストールしたディレクトリです。

```
<product install dir>/jre/bin
```

- ▶ 次のコマンドを入力します。

```
keytool -import -alias <alias name> -file <path to cert> -keystore <path to key store>
```

この場合、<alias name> は証明書の任意のエイリアス、<path to cert> は証明書へのフル パス、<path to key store> は、Java キーストアへのフル パスで、<product install dir>/lib/security/jssecacerts または <product install dir>/lib/security/cacerts となります。

- ▶ プロンプトが表示されたら、キーストアのパスワードを入力します。デフォルトは changeit です。
- ▶ 証明書の信頼についてプロンプトが表示されたら、yes と入力します。

## エンド ユーザーによる SSL の有効化

エンド 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。たとえば IBM® SPSS® Modeler では、[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 [セキュア接続の確認] チェック ボックスをオンにする必要があります。

## URL プレフィックスの設定

IBM® SPSS® Collaboration and Deployment Services Repository が SSL アクセスに設定されている場合、URL プレフィックス設定の値を、次のように変更する必要があります。

1. ブラウザベースのコンソールを使用してリポジトリにログインします。
2. URL プレフィックスの設定オプションを開きます。  
布置 > 設定 > URL プレフィックス
3. プレフィックスの値を **http** ではなく **https** に設定し、ポート番号を SSL ポート番号に設定します。次に例を示します。

```
[default]
http://<hostname>:<port>
[SSL-enabled]
https://<hostname>:<SSLport>
```

## SSL による LDAP の保護

Lightweight Directory Access Protocol (LDAP) は、任意のレベルの情報を含むネットワーク ディレクトリとデータベース間で情報を交換するためのインターネット技術タスク フォース (IETF) 基準です。追加のセキュリティが必要なシステムに対し、Microsoft の Active Directory などの LDAP プロバイダは、Web またはアプリケーション サーバーが SSL で LDAP をサポートしている場合、Secure Socket Layer (SSL) で操作することができます。LDAP と連動させて SSL を使用すると、ログイン パスワード、アプリケーション情報、そのほかの重要データの侵入、改ざん、漏洩されません。

次の例では、セキュリティ プロバイダとして Microsoft の Active Directory を使用した LDAP の有効化方法を説明しています。手順についての詳細または、特定のリリースのセキュリティ プロバイダを扱う詳細については、発売元のベンダーのマニュアルを参照してください。

1. Active Directory および企業の認証機関が導入され、機能していることを確認します。


2. 認証機関を通じて証明書を生成し、証明書を IBM® SPSS® Collaboration and Deployment Services Deployment Manager インストールの証明書ストアにインポートします。これにより、IBM® SPSS® Collaboration and Deployment Services Repository と Active Directory サーバー間の LDAPS 接続を確立します。

安全な Active Directory 接続向けに Deployment Manager を設定するために、リポジトリへの接続が終了していることを確認します。

3. IBM® SPSS® Collaboration and Deployment Services Deployment Manager を起動します。
4. [ツール] メニューの [サーバー管理] を選択します。
5. 以前に定義された管理対象のサーバーにログインします。
6. サーバーの [構成] アイコンをダブルクリックして、階層を展開します。
7. [セキュリティプロバイダ] アイコンをダブルクリックして、階層を展開します。
8. Active Directory セキュリティ プロバイダをダブルクリックします。
9. セキュリティ証明書をインストールした Active Directory のインスタンスに構成値を入力します。
10. [SSL を使用] チェック ボックスを選択します。
11. [ドメイン ユーザー] フィールドに名前を入力します。Active Directory を使用した今後のログインは、SSL を使用して認証されます。

特定のアプリケーション サーバーに LDAP のインストール、構成、実行の詳細は、発売元のベンダーのマニュアルを参照してください。

## サーバー ログ

IBM® SPSS® Modeler Server では、messages.log というログ ファイルに、重要なアクションの記録が維持されます。デフォルト  は、このファイルは [サーバーのインストール先へのパス]/log ディレクトリにありますが、この場所は設定可能です。SPSS Modeler Server が起動されると、起動時間と待機しているポート番号を示すために、ログ ファイルが更新されます。この時点から、クライアントがサーバーに接続しようとしたときは常に、その結果を示すためにログ ファイルが更新されます。

デフォルトでは、サーバーにより次のイベントがログに書き出されます。

- サーバーの起動
- サーバーの停止
- 接続の受け入れ



- 接続の拒否
- セッションの起動
- セッションの停止
- システム エラー
- ログインの成功
- ログインの失敗

データベースへのアクセス イベントについての情報を記録する場合は、SPSS Modeler Server の構成で、[データベース アクセス情報をログに記録する] と [データベース エラーをログに記録する] を [はい] に変更します。次のデータベース アクセス イベントがログに記載されます。

- データベースへのログインの成功
- データベースへのログインの失敗
- データベースからのログアウト

ファイルへのアクセス イベントについての情報を記録する場合は、SPSS Modeler Server の構成で、[ファイル アクセス情報をログに記録する] と [ファイル アクセス エラーをログに記録する] を [はい] に変更します。次のファイル アクセス イベントがログに記載されます。

- ファイルを開く/作成する
- ファイルを閉じる

次は、ログ ファイルの例です（行番号を付けてあります）。

1. 2006/06/14 08:41:23 [167]: IBM SPSS Modeler Server Version 0.1.0 started
2. 2006/06/14 08:41:23 [167]:Listening on port 28028
3. 2006/06/14 08:49:28 [167]:Accepted connection from localhost:1827
4. 2006/06/14 08:49:28 [167]:Session 105 (localhost:1827) started
5. 2006/06/14 08:49:29 [167-105]:Login failed for user:SPSS\awilson
6. 2006/06/14 09:27:27 [167]:Session 105 ended
7. 2006/06/14 09:53:39 [167]:Accepted connection from localhost:1837
8. 2006/06/14 09:53:39 [167]:Session 63 (localhost:1837) started
9. 2006/06/14 09:53:40 [167-63]:Login succeeded for user:SPSS\awilson
10. 2006/06/14 10:27:31 [167]:Session 63 ended

ログ ファイル内のすべての項目には、タイム スタンプが含まれています。1 行目と 2 行目は、SPSS Modeler Server が起動されたときにログ ファイルに追加されました。3 行目は、ポート番号 1827 で localhost コンピュータのユーザーがサーバーに接続しようとしたときに追加されました。4 行目は、サーバー セッションがこのユーザーにより開始されたことを示しています。5 行目は、このユーザーがログインに失敗したことを示しています。しばらく後に、このユーザーはセッションを終了しました。30 分後に、別の接続が localhost から試行され、セッション

ンが開始されました。今回のログインは成功でした。さらに 30 分後、セッションは終了して接続が閉じられました。

# 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 最適化を使用する。
- 2 章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。
- 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。これらの設定のいくつかはクライアントでも利用できますが、IBM® SPSS® Modeler をスタンドアロン モードで使用している場合にだけ適用可能です。IBM® SPSS® Modeler がサーバーに接続されていると、サーバーのパフォーマンスと最適化の設定が自動的に使用されます。
- 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最適化する。  
詳細は、13 章 [ストリームおよびノードのパフォーマンスの考察 in IBM SPSS Modeler 14.2 ユーザー ガイド](#) を参照してください。

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、SPSS Modeler Server に含まれている IBM® SPSS® Collaboration and Deployment Services Deployment Manager アプリケーションを使用して調整できます。詳細は、4 章 [p.27 IBM SPSS Modeler Administration Console の使用方法](#) を参照してください。

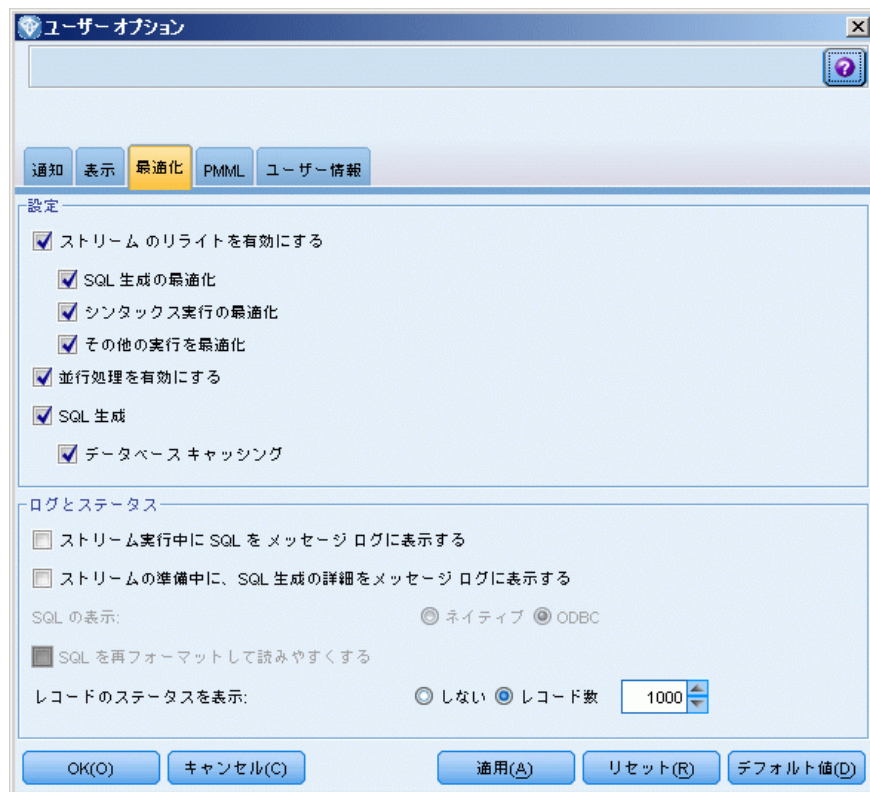
設定値は、Deployment Manager 構成ウィンドウの [パフォーマンスと最適化] の見出しの下でグループ化されています。各設定値は、ほとんどのインストールで最適なパフォーマンスになるように、事前に構成されています。ただし、使用する特定のハードウェア、データ セットのサイズ、およびストリームの内容に応じて、設定値を調整する必要があります。詳細は、4 章 [p.33 パフォーマンス/最適化](#) を参照してください。

## クライアントのパフォーマンスと最適化の設定

[ユーザー オプション] ダイアログ ボックスの [最適化] タブを使用すると、ストリームのパフォーマンスを最適化できるようになります。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、クライアント内で上書きされます。このタブは、SPSS Modeler Server 接続が IBM® SPSS® Modeler コンピュータで有効化されている場合にのみ使用できます。

注： データベース モデル作成および SQL 最適化では、SPSS Modeler Server 接続が SPSS Modeler コンピュータ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、SPSS Modeler から SQL を直接プッシュバック、SPSS Modeler Server にアクセスできます。現在のライセンス ステータスを検証するには、SPSS Modeler の [ヘルプ] メニューから [バージョン情報] を選択します。詳細は、3 章 IBM SPSS Modeler Server への接続 in IBM SPSS Modeler 14.2 ユーザー ガイド を参照してください。

図 5-1  
[ユーザー オプション] ダイアログ ボックス、[最適化] タブ



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 14.2 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。

**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、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 2 種類あり、どちらかまたは両方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 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 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。詳細は、6 章 p.57 SQL 最適化 を参照してください。
- **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。複数のノードを使用、または C5.0、結合（キーごと）、ソート、ビン（ランクおよびタイル手法）、レコード集計（1 つまたは複数のキーフィールドを使用）を使用すると、平行処理のメリットが得られます。

**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。たとえば、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、たと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 最適化が有効でない場合、キャッシュは、代わりにファイル システムへ書き込まれます。詳細は、5 章 [ノードのキャッシュ オプション in IBM SPSS Modeler 14.2 ユーザー ガイド](#) 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、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。

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

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

**SQL の表示：**ログ中に表示される SQL を、固有の SQL 関数、または、SPSS Modeler に生成されたとおりの {fn FUNC(...)} 形式の標準 ODBC 関数のどちらかの形式で含めるのかを指定します。前者は、実装されていない可能性がある ODBC ドライバの機能に依存しています。たとえば、このコントロールは、SQL Server に対して何の影響も及ぼしません。

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

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

このタブのデフォルト設定に戻す場合は、[デフォルト値] を選択します。

## データベースの使用率と最適化

**データベース サーバー：**可能であれば、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、運用中のサーバーが IBM® SPSS® Modeler のクエリーに影響を受けないようにします。SPSS Modeler で生成された SQL ステートメントは、同じデータベース内で SQL を実行できる IBM® SPSS® Modeler Server コンピュータ上で、複数のタスクを要求している可能性があります。

**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。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、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 の量を常に最大になるよう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。詳細は、[6 章](#) 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。詳細は、[7 章 データベース エクスポートの詳細オプション in IBM SPSS Modeler 14.2 入力ノード、プロセス ノード、出力ノード](#) を参照してください。

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イスがあります。いくつかのスクリ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 SPSS Modeler インストール フォルダの下の /scripts サブディレクトリにあります。

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` ユーティリティです。

テーブル 5-1  
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      |



# 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 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す（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（たとえば、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、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 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。詳細は、[p. 64 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、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。たとえば、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 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。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、SPSS Modeler Server 接続が SPSS Modeler コンピュータ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、SPSS Modeler から SQL を直接プッシュバック、SPSS Modeler Server にアクセスできます。現在のライセンス ステータスを検証するには、SPSS Modeler の [ヘルプ] メニューから [バージョン情報] を選択します。詳細は、3 章 IBM SPSS Modeler Server への接続 in IBM SPSS Modeler 14.2 ユーザー ガイド を参照してください。

### データベースの要件

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 モデリングを使用すると、更なるパフォーマンスの向上を獲得できます。詳細は、2 章 データベース モデル作成の概要 in IBM SPSS Modeler 14.2 データベース内マイニング ガイド を参照してください。

## 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 へ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。

- サーバーにより、下流のノードを安全であると確認できる「SQL ゾーン」へ移動するために、ストリームを並べ替えます。(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ます。検証に成功した SQL が、実行用を選択されます。
- 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。 [詳細は、 p.64 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

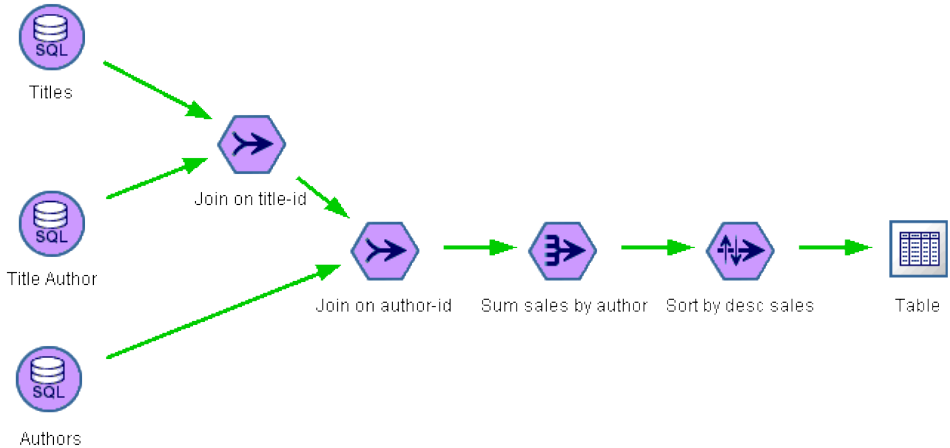
SQL 最適化によって、多” のデータ操作のパフォーマンスが向上します。

- **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。
- **選択部分**：ある種の基準に基づいてレコードを選択すると、レコードの品質が落ちます。
- **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**:SPSS Modeler Server では、以後の処理が必要なフィールドだけをデータベースから抽出します。このことで、帯幅とメモリーの要件が最小になります。同じことがフラット ファイルの余分なフィールドにも当てはまります。サーバーは余分なフィールドを読む必要がありますが、それらにストレージは割り当てません。
- **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 6-1  
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。

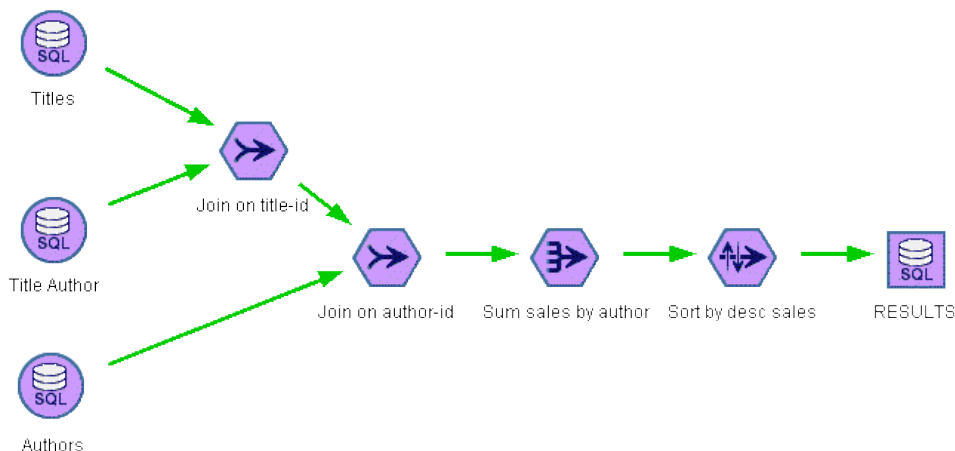
```

SELECT
  T2.au_fname AS C0,
  T2.au_lname 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 6-2  
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

- ▶ 使用するデータベースのために ODBC ドライバをインストールして、データソースを設定します。詳細は、2 章 p.12 データへのアクセスを参照してください。
- ▶ 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。
- ▶ 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

- ▶ [ツール] メニューの [ユーザー オプション] を選択します。
- ▶ [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。詳細は、5 章 p.52 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、4 章 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。詳細は、3 章 [ストリーム内でのモデル ナゲットの使用](#) in IBM SPSS Modeler 14.2 Modeling Nodes を参照してください。

## 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。

- ▶ [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。詳細は、5 章 [p. 52 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。
- ▶ 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。
- ▶ 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。

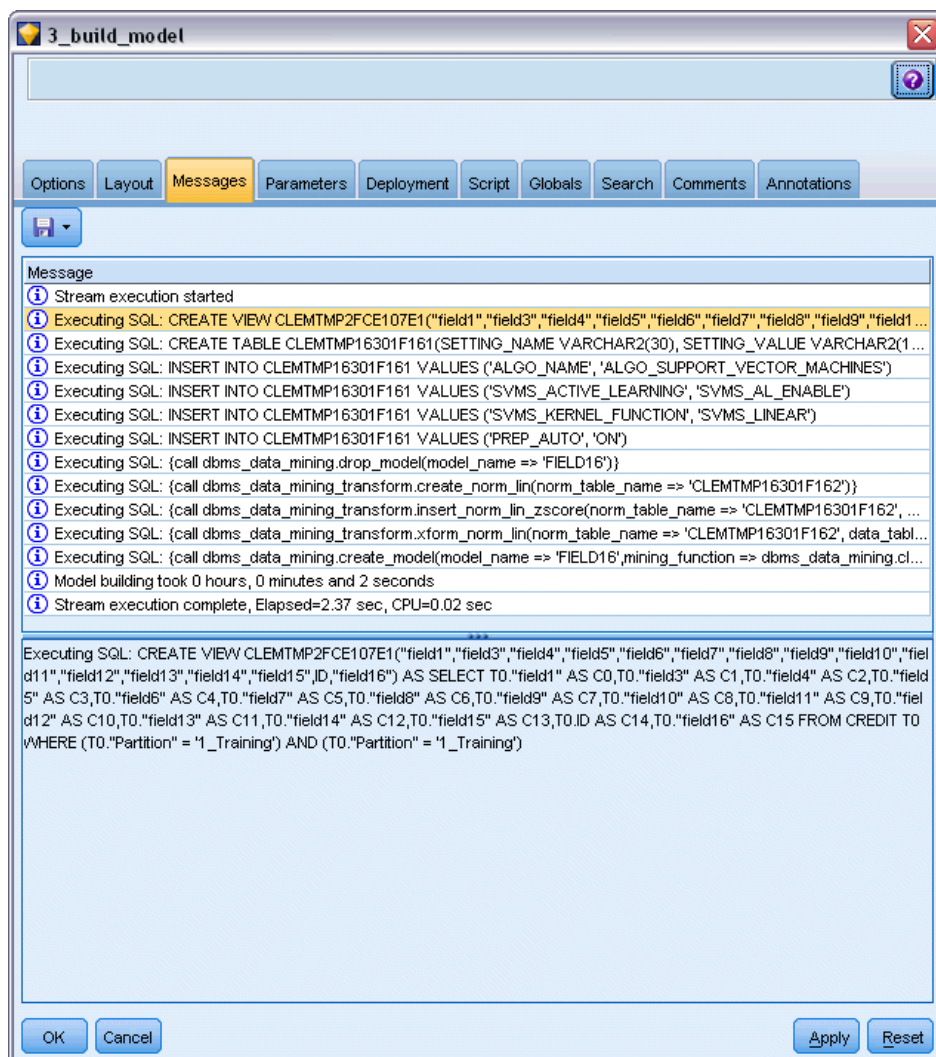
SQL が生成された（そして、ストリームが実行されるときにデータベースへプッシュバックされる）すべてのノードが、ストリーム領域に紫色で表示されます。

図 6-3  
[SQL のプレビュー] ボタン



- ▶ 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(M)...]

図 6-4  
メッセージ ログに表示される生成された SQL



## 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。詳細は、[p. 65 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。

- ▶ モデル ナゲットの [設定] タブを選択します。
- ▶ 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。
- ▶ モデル ナゲットのメニューから、次を選択します。  
File > SQL のエクスポート
- ▶ ファイルを保存します。
- ▶ 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。 [詳細は、p. 65 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** : ストリーム内で再順序付けができない場合は、ノードのオプションまたは CLEM 式を変更することができます。または、操作が実行される方法を変更して、以後 SQL 生成が抑制されないようにします。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変換できますが、SQL 生成を抑制する多数の演算子もあります（特に、シーケンス関数の “@ 関数”）。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。 [詳細は、p. 69 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（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 アルゴリズム：** モデルの推定は、データベースでなく、常に IBM® SPSS® Modeler Server で行われます。ただし、Microsoft、IBM、または Oracle のデータベース固有のアルゴリズムを使用している場合は別です。詳細は、[2 章 データベース モデル作成の概要 in IBM SPSS Modeler 14.2 データベース内マイニング ガイド](#) を参照してください。

**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せん。ただし、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。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る場合があります)。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。詳細は、[p.65 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 マネージャを使用して 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地点 (たとえば、レコード結合ノードやレコード集計ノードの次) 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。詳細は、[p.61 SQL 最適化の設定](#) を参照してください。詳細は、[ノードのキャッシュ オプション](#) 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。詳細は、[p.62 生成された SQL のプレビュー](#) を参照してください。

## 入力

| SQL 生成をサポートしているノード | メモ   |
|--------------------|--|
| 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。詳細は、 <a href="#">p. 72 SQL クエリーの記述</a> を参照してください。 |

## レコード設定

| SQL 生成をサポートしているノード | メモ   |
|--------------------|--|
| Select             | 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、次の表に示すように SQL 生成をサポートしています。複雑なサンプリングでは、SQL 生成をサポートしていません。   |
| グループ集計             |  |
| RFM レコード集計         | 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、 <code>datetime_date(YEAR,MONTH,DAY)</code> 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。  |
| Sort               |  |
| 結合                 | 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。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 ノードまたは入力ノードの [フィルタ] 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。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、 <code>SQL_SP_EXISTS</code> はサポートされていません。 |
| 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             |  |

単純なサンプリングを使用する場合のサンプル ノードの SQL 生成のサポートを次の表に示します。

| モード     | サンプル  | 最大サイズ | シード | DB2 OS/Z | DB2 OS/400 | DB2 Windows/UNIX | Netezza | Oracle | SQL Server | Teradata |
|---------|-------|-------|-----|----------|------------|------------------|---------|--------|------------|----------|
| Include | First | 利用不可  |     | Y        | Y          | Y                | Y       | Y      | Y          | Y        |
|         | n 件ごと | off   |     | Y        | Y          | Y                | Y       | Y      |            | Y        |
|         |       | max   |     | Y        | Y          | Y                | Y       | Y      |            | Y        |
|         | 無作為 % | off   | off |          |            | Y                | Y       | Y      |            | Y        |
|         |       |       | on  |          |            | Y                |         | Y      |            |          |
|         |       | max   | off |          |            | Y                | Y       | Y      |            | Y        |
|         |       |       | on  |          |            | Y                |         | Y      |            |          |
| 破棄      | First | off   |     |          |            |                  | Y       | Y      |            |          |
|         |       | max   |     |          |            |                  | Y       | Y      |            |          |
|         | n 件ごと | off   |     | Y        | Y          | Y                | Y       | Y      |            | Y        |
|         |       | max   |     | Y        | Y          | Y                | Y       | Y      |            | Y        |
|         | 無作為 % | off   | off |          |            | Y                | Y       | Y      |            | Y        |
|         |       |       | on  |          |            | Y                |         | Y      |            |          |
|         |       | max   | off |          |            | Y                | Y       | Y      |            | Y        |
|         |       |       | on  |          |            | Y                |         | Y      |            |          |

## フィールド設定

| SQL 生成をサポートしているノード | メモ  |
|--------------------|---|
| データ型               | データ型ノードがインスタンス化されており、 <b>ABORT</b> または <b>WARN</b> 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               |   |
| 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)。                             |
| アンサンブル             | 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。 |
| 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)。                             |
| 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。     |

| SQL 生成をサポートしているノード | メモ  |
|--------------------|---|
| 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。 |
| Partition          | 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                |   |

### グラフ作成

| SQL 生成をサポートしているノード | メモ   |
|--------------------|--|
| 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 データのみサポートされています。 |
| 分布                 |  |
| Web                |  |
| 評価                 |  |

### 生成されたモデル ノード

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。 [詳細は、p.63 モデル ナゲットの SQL の表示 を参照してください。](#)

| SQL 生成をサポートしているモデル ナゲット | メモ  |
|-------------------------|---|
| 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                    |   |
| ディシジョン リスト              |   |
| Linear                  | 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                    |   |

|                         |  |
|-------------------------|--|
| SQL 生成をサポートしているモデル ナゲット | メモ   |
| ロジスティック                 | 多項式プロシージャの 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             |  |

### 出力

|                    |   |
|--------------------|---|
| SQL 生成をサポートしているノード | メモ  |
| テーブル               | 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。 |
| 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                 |   |
| 統計値                | [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              |   |

### エクスポート

|                    |                              |
|--------------------|------------------------------|
| SQL 生成をサポートしているノード | メモ                           |
| 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。

### 演算子

|                   |    |
|-------------------|----|
| SQL 生成をサポートしている操作 | メモ |
| +                 |    |
| -                 |    |
| /                 |    |

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

### 関係演算子

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

### 関数

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

| SQL 生成をサポートしている操作 | メモ |
|-------------------|----|
| length            |    |
| locchar           |    |
| log               |    |
| log10             |    |
| lowertoupper      |    |
| max               |    |
| member            |    |
| min               |    |
| negate            |    |
| not               |    |
| number            |    |
| or                |    |
| pi                |    |
| real              |    |
| rem               |    |
| round             |    |
| sign              |    |
| sin               |    |
| sqrt              |    |
| string            |    |
| strmember         |    |
| subscrs           |    |
| substring         |    |
| substring_between |    |
| uppertolower      |    |
| to_string         |    |

### 特殊関数

| SQL 生成をサポートしている操作 | メモ  |
|-------------------|---|
| @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       |   |

## 集計関数

|                   |    |
|-------------------|----|
| SQL 生成をサポートしている操作 | メモ |
| Sum               |    |
| Mean              |    |
| Min               |    |
| Max               |    |
| Count             |    |
| SDev              |    |

## CLEM 式での SQL 関数の使用

@SQLFN関数を使用して、データベースを実行する目的のみで、指定の SQL 関数を CLEM 式内に追加できます。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. が制御できない外部データベース コンポーネントに依存するため、この関数の使用は標準の IBM® SPSS® Modeler のサポート同意契約にはカバーされていません。ただし、特別なケースの場合、サービス契約の一環として展開することができます。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には、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。たとえば、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ます。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;
```

# 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 ドライバと共に提供されているロケール以外のロケールで、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、コンピュータを再設定する必要があります。Connect ODBC ドライバには、en\_US ロケール ファイルだけが含まれています。その結果、SPSS Modeler Server コンピュータが、異なるロケールで実行されている場合や、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。

- 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。それから、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 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 SPSS Modeler Server に返されます。これが、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。

- ▶ 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。例：

```
$ 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
```

- ▶ Connect ODBC/locale ディレクトリに変更します（ここにあるディレクトリは en\_US 1 つだけです）。
- ▶ 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.ISO8859-1
```

- ▶ リンクを作成したら、同じシェルから SPSS Modeler Server を再起動します（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）。

## 注意

UNIX コンピュータを、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、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、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。たとえば、次の定義を考えてみてください。

```
#LANG=en_US.ISO8859-15; export LANG
```

- 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。

# 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と、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 コンピュータが再起動されたときに自動的に起動するように、SPSS Modeler Server を構成する。
- SPSS Modeler Server を手動で停止し、再起動する。
- SPSS Modeler Server で作成されたファイルに対する権限を変更する。
- SPSS Modeler Server とともに提供された ODBC 接続ドライバと協調できるように、SPSS Modeler Server を構成する。 [詳細は、 p. 78 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、SPSS Modeler Server を構成する。 [詳細は、 p. 86 エンタープライズ ビュー ノードのドライバを構成 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**: SPSS Modeler Server のインストール ディレクトリにある、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」起動されるときの、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、SPSS Modeler Server のインストール ディレクトリにあります。
- **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」起動されるときの環境を構

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。

テーブル B-1  
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 | 位置                      |
|---------------|-------------------------|
| AIX           | /etc/rc.modeler         |
| HP-UX         | /sbin/init.d/rc.modeler |
| Solaris       | /etc/init.d/rc.modeler  |

## 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、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

- ▶ root としてログオンします。
- ▶ SPSS Modeler Server のインストール ディレクトリに変更します。
- ▶ 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 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

- ▶ SPSS Modeler Server のインストール ディレクトリに変更します。
- ▶ 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
```

- ▶ 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

- ▶ IBM® SPSS® Modeler Server を停止します。（詳細は、[p. 77 IBM SPSS Modeler Server の手動による起動と停止](#) を参照してください。）
- ▶ 適切なスクリプトを検索します。（詳細は、[p. 76 スクリプト](#) を参照してください。）
- ▶ テキスト エディタ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。
- ▶ サーバー コンピュータを自動的に、または手動で再起動して、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、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

Connect for ODBC ドライバを IBM® SPSS® Modeler Server とともに使用する場合は、SPSS Modeler Server プロセスが始動するときに、Connect for 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。（詳細は、p. 78 スクリプトの編集 を参照してください。）

詳細は、当社 Web サイト (<http://www.spss.com/drivers/>) を参照してください。ODBC データ ソースに関する質問がある場合は、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 Connect for ODBC を構成する手順

- ▶ Web サイト <http://www.spss.com/drivers/clientCLEM.htm> から、SPSS Modeler Server をインストールするプラットフォームに関連する圧縮 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 SPSS Modeler Server のドライバをダウンロードしてください。ファイルを ODBC ドライバをインストールする場所（例： /usr/spss/odbc）にコピーします。
- ▶ アーカイブを回答します。
- ▶ アーカイブから抽出した setodbcpath.sh スクリプトを実行します。
- ▶ 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 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 はドライバに付属しています）。

- ▶ odbc.sh を保存します。
- ▶ (64 ビットの 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 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。
- ▶ これらのドライバを使用するよう 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。たとえば、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

- ▶ SPSS Modeler Server を再起動します。
- ▶ クライアントから 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。詳細は、[p. 82 Connect for ODBC 構成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 Connect for ODBC を構成する手順

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 SPSS Modeler Solution Publisher Runtime のインストールを設定することができます。

- ▶ 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。たとえば、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。

- ▶ `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 Connect for 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 SPSS Modeler Batch から 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、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

- ▶ `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。
- ▶ 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。

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 SPSS Modeler Solution Publisher Runtime および SPSS Modeler Batch に適切なものである必要があります。

## Connect for ODBC 構成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あります。

- ▶ \$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。
- ▶ 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。
- ▶ 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
```

- ▶ modelersrv.sh を保存して実行します。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。
- ▶ 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。たとえば、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、SPSS Modeler Server を再起動してクライアントから接続するときに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 パスが ODBC ライブラリ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 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。

SPSS Modeler Server デーモン環境でライブラリ パスの値を表示するには、ODBC ライブラリ パスがライブラリ 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 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 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 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 パス変数です。

ライブラリ パスの戻り値には、ODBC の lib サブディレクトリ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 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=
```

- ▶ ドライバ パラメータ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。
- ▶ 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。
- ▶ ドライバ パラメータ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。詳細は、[p. 78 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のライブラリの問題となる場合があります。ドライバがロードされないことを確認するために、Connect for 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パスの変数を設定して問題を修正します。

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

- ▶ **ivtestlib** を使用して、ドライバがロードされないことを確認します。たとえば、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
. odbctest.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へのパスに置き換えます。

- ▶ メッセージを読んで、エラーがあるかどうかを調べます。たとえば、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の **libclntsh.so** がないか、または、このライブラリパス（たとえば、Solaris の **LD\_LIBRARY\_PATH**）にないことを示しています。

- ▶ ライブラリ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パス変数です。

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、ドライバがこれでロードされたことを確認します。たとえば、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 ライブラリがロードできたことを示しています。

- ▶ SPSS Modeler 起動スクリプト内のライブラリ 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、SPSS Modeler Server を再起動します。

## ライブラリ パス

ライブラリ 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。

テーブル B-2  
オペレーティング システム別のライブラリ パス

| オペレーティング システム            | ライブラリ 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。詳細は、[9 章 IBM SPSS Collaboration and Deployment Services Repository オブジェクトの保存と展開 in IBM SPSS Modeler 14.2 ユーザー ガイド](#) を参照してください。

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 ドライバ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 の参照を、modelersrv.sh または rc.modeler のいずれかまたは両方の起動スクリプト に、次の形式で追加する必要があります。

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

# UNIX 上での非ルート プロセスとしての実行

## 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** :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。ただし、次の手順に従うと、root 以外のアカウントでデーモン プロセスを実行することが可能になります。

**非 root として実行** : root 権限がなくても、SPSS Modeler では、UNIX パスワード データベースとは区別されるプライベート パスワード データベースに対してユーザーを認証し、同じ UNIX ユーザー アカウントで、すべてのユーザー セッションを起動します。つまり、SPSS Modeler によって読み書きされたすべてのファイル データは、すべての SPSS Modeler ユーザーに共有されます。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。

## 非 root プロセスとしての IBM SPSS Modeler Server の構成

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。サーバーを起動したり停止したりするためと一般的な管理を行うには、引き続き root または root に対する su としてログインする必要があります。

非 root アカウントで実行するために 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. 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、SPSS Modeler Server デーモン プロセスの専用です。つまり、サーバーを起動したり停止したりするためと、通常の IBM® SPSS® Modeler 管理を実行するためには、このアカウントを使用できません。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。

- アカウントにログイン機能は必要ありません。
- 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 ディレクトリは、SPSS Modeler のインストール ディレクトリであってもよいし、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 ディレクトリ以外の別のディレクトリ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 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 エディタ内で開きます。詳細は、[B 付録 p. 76 スクリプト](#) を参照してください。

- ▶ 適切なスクリプト内で、SPSS Modeler Server を実行するために作成されたユーザー アカウントの名前を参照する、**MODELERUSER** 変数を設定する行を変更します。次に例を示します。

```
MODELERUSER=modelerserv
```

- ▶ 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. 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、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 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、SPSS Modeler Server のインストール先の bin ディレクトリ内にある、pwutil というパスワード ユーティリティ 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フォーマットで、標準出力に書き出されます。たとえば、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。

6. SPSS Modeler のインストール先ディレクトリ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。次に例を示します。

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

7. 他からの干渉なしに作業データを格納できるように、SPSS Modeler ユーザーのデータ ディレクトリ内にサブディレクトリを作成することを考慮します。これらのディレクトリは <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。たとえば、ユーザー bob 用の作業ディレクトリを作成するには、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に設定して、そのディレクトリ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
```

### **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 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 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
```

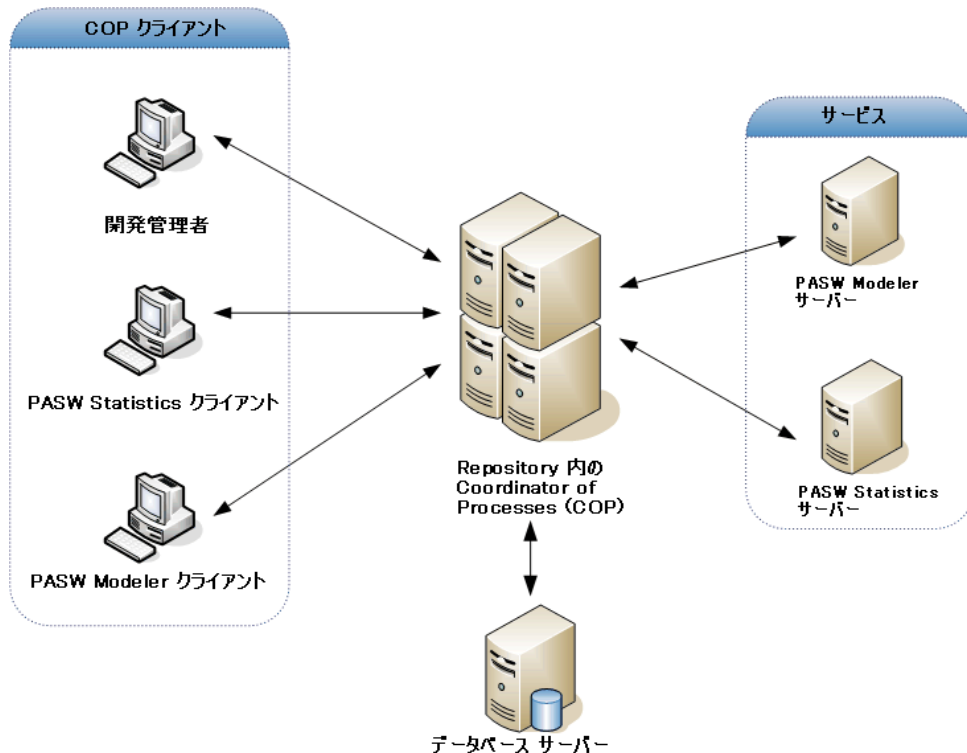
リンクを作成した後、SPSS Modeler Server を再起動します。

# サーバー クラスタ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を送信します。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 D-1  
Coordinator of Processes アーキテクチャ



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 サーバーと本稼動用サーバーなど、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 クラスタが提供されることもあります。

サーバー クラスタとは、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 開発者ガイド』（IBM Corp. から入手可能）を参照してください。

SPSS Modeler でサーバーまたはサーバー クラスタに接続するたびに、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を検索することができます。 [詳細は、3 章 IBM SPSS Modeler Server への接続 in IBM SPSS Modeler 14.2 ユーザー ガイド](#) を参照してください。

# 注意事項

This information was developed for products and services offered worldwide.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive,  
Armonk, NY 10504-1785, U. S. A.

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing, Legal and Intellectual Property  
Law, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa  
242-8502 Japan.

**次の文は、条項が法律と一致しないイギリスなどの国には適用されません。** 本出版物は、SPSS INC., AN IBM COMPANY によって提供され、明示的および暗黙的なあらゆる保証、制限されていない場合を除く商品性や特定の目的への適合性、および無違反に関する暗黙的な保証を含む意思表示と保証を放棄します。特定の取引では明示的または暗黙的な保証の免責が許可されないため、この文が適用されない場合があります。

この情報には、技術的な誤りや誤植を含まれる場合があります。本文では変更が定期的に行われます。これらの変更は本書の次の版に組み込まれます。SPSS は、本文書に記載された製品やプログラムは予告なしに改善または変更される場合があります。

この情報内にある SPSS 以外または IBM 以外の Web サイトに対する参照は、便宜上提供されたものであり、これらの Web サイトを推奨するものではありません。これらの Web サイトの資料は、この SPSS 社製品の使用の一部ではなく、これらの Web サイトの使用は個人の責任によるものです。

IBM または SPSS に情報を送信すると、あなたに対する義務を負うことなく、適切とする方法でその情報を使用または配布する非独占的権利と IBM および SPSS 付与するものとします。

SPSS 以外の製品に関する情報は、これらの製品、公開された通知、公表されているソースの供給者から得たものです。SPSS は、それらの製品をテストしていません。また、SPSS 以外の製品に関連するパフォーマンスの正確性、互換性、またはその他の要求を確認することはできません。SPSS 以外の製品の機能に関する質問は、これらの製品の供給者にお問い合わせください。

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Software Group, Attention: Licensing, 233 S. Wacker Dr., Chicago, IL 60606, USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the

capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM' s future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

この情報には、日常の業務で使用されているデータおよびレポートの例が含まれています。それらを可能な限り詳細に説明するために、例には個人、企業、ブランド、製品の名前が含まれます。これらの名前はすべて架空のものであり、実際の名前や住所に似ているものでも、まったくの偶然によるものです。

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

## 商標

IBM、IBM ロゴ、ibm.com は世界各国の四方に基づく IBM 社の登録商標です。IBM の商標の現在のリストは Web サイト <http://www.ibm.com/legal/copytrade.shtml> を参照してください。

SPSS Inc., an IBM Company の SPSS の商標 は、世界各国の司法に基づく登録商標です。

Adobe、Adobe のロゴ、PostScript、および PostScript ロゴはアメリカ合衆国およびその他各国のアドビシステムズ社の登録商標または商標です。

IT Infrastructure Library は、イギリス商務局の一部である中央電子計算機局の登録商標です。

Intel、Intel のロゴ、Intel Inside、Intel Inside のロゴ、Intel Centrino、Intel Centrino のロゴ、Celeron、Intel Xeon、Intel SpeedStep、Itanium、Pentium はアメリカ合衆国およびその他各国のインテル社およびその子会社の商標または登録商標です。

Linux は、アメリカ合衆国およびその他各国の Linus Torvalds の登録商標です。

Microsoft、Windows、Windows NT、および Windows ロゴは、アメリカ合衆国およびその他合衆国のマイクロソフト社の商標です。

ITIL は、米国特許商標局の登録商標および登録共同体商標です。

UNIX は、アメリカ合衆国およびその他各国の The Open Group の登録商標です。

Cell Broadband Engine は、アメリカ合衆国およびその他各国のソニーコンピュータエンタテインメント株式会社の使用許諾に基づいて使用されています。

Java および Java ベースの商標およびロゴは、アメリカ合衆国およびその他各国のサン・マイクロシステムズ株式会社の商標です。

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

その他の製品およびサービス名は、IBM、SPSS、またはその他の企業の商標である場合があります。



# 索引

- 並行処理
  - 有効化, 52
  - 制御, 33
- 暗号化
  - SSL, 41
- 最適化, 52
  - SQL 生成, 57-58, 61
- 演算子
  - SQL 生成, 69
- 保護
  - LDAP, 47
- 商標, 95
- 接続
  - IBM SPSS Modeler Server へ, 16, 18, 20
  - サーバー クラスタ, 20
- 管理
  - IBM SPSS Modeler Server, 26
  - IBM SPSS Statistics ライセンス  
ロケーション, 41
- 結果
  - Client と Server との違い, 22
  - 丸め, 23
  - レコードの並び順, 22
- 認証, 21
- 例
  - 概要, 4
  - アプリケーション ガイド, 3
- 64-bit のオペレーティング システム, 10
  
- allow\_modelling\_memory\_override
  - options.cfg ファイル, 33
- auto.sh (UNIX)
  - 場所, 76
- cache\_compression
  - options.cfg ファイル, 33
- cache\_connection オプション, 40
- chemsrv.sh (UNIX)
  - 場所, 76
- CLEM 式
  - SQL 生成, 69
- Connect ODBC
  - UNIX 上で構成, 78
- Coordinator of Processes, 20
  - サーバー クラスタ, 91
  - 負荷バランシング, 91
- Coordinator of Processes の構成
  - IBM SPSS Modeler Server の, 37
- COP, 20
  - サーバー クラスタ, 91
  - 負荷バランシング, 91
- COP の構成
  - IBM SPSS Modeler Server の, 37
  
- cop\_enabled
  - options.cfg ファイル, 37
- cop\_host
  - options.cfg ファイル, 37
- cop\_password
  - options.cfg ファイル, 37
- cop\_port\_number
  - options.cfg ファイル, 37
- cop\_service\_description
  - options.cfg ファイル, 37
- cop\_service\_host
  - options.cfg ファイル, 37
- cop\_service\_name
  - options.cfg ファイル, 37
- cop\_service\_weight
  - options.cfg ファイル, 37
- cop\_update\_interval
  - options.cfg ファイル, 37
- cop\_user\_name
  - options.cfg ファイル, 37
- data\_file\_path
  - options.cfg ファイル, 32
- data\_files\_restricted
  - options.cfg ファイル, 32
- DB2
  - SQL 最適化, 57-58
- IBM SPSS Modeler, 1
  - ドキュメンテーション, 3
- IBM SPSS Modeler Administration Console,  
27
  - 管理者アクセス, 28
- IBM SPSS Modeler Server
  - Coordinator of Processes の構成, 37
  - COP の構成, 37
    - password, 16
    - 管理者アクセス, 28
    - エンド ユーザー向けの情報, 21
    - 構成オプション, 30
    - 管理オプション, 26
    - クライアントと比べた結果の違い, 22
    - サーバー プロセス, 38
      - 一時ディレクトリ, 32
      - ドメイン名 (Windows), 16
    - 使用率の監視, 38
    - の管理, 27
      - ファイルの作成, 22
      - 非応答プロセス, 22
    - ホスト名, 16, 18
    - ポート番号, 16, 18, 30
    - ユーザー ID, 16
    - ユーザー アカウント, 21
    - ユーザーの認証, 21

## 索引

- ログ ファイル, 31
- IBM SPSS Modeler Server 接続の追加, 18, 20
- IBM SPSS Modeler Server の停止
  - UNIX 上で, 25
  - Windows 上で, 24
- IBM SPSS Modeler Server の起動
  - UNIX 上で, 25
  - Windows 上で, 24
- IBM SPSS Modeler Server へのログイン, 16
- IBM SPSS Statistics データ アクセス  
テクノロジー, 12
- IBM SPSS Statistics データ ファイル  
インポートとエクスポート, 14
- IBM SPSS Statistics ライセンス  
ロケーション, 41
- IBM SPSS Text Analytics, 2
- io\_buffer\_size
  - options.cfg ファイル, 33
- LDAP, 47
  - 保護, 47
- log\_database\_access
  - options.cfg ファイル, 31
- log\_database\_errors
  - options.cfg ファイル, 31
- log\_directory
  - options.cfg ファイル, 31
- log\_file\_access
  - options.cfg ファイル, 31
- log\_file\_errors
  - options.cfg ファイル, 31
- max\_file\_size
  - options.cfg ファイル, 32
- max\_login\_attempts
  - options.cfg ファイル, 30
- max\_parallelism
  - options.cfg ファイル, 33
- max\_sessions
  - options.cfg ファイル, 30
- max\_sql\_string\_length
  - options.cfg ファイル, 35
- memory\_usage
  - options.cfg ファイル, 33
- messages.log ファイル, 48
- Microsoft SQL Server
  - SQL 最適化, 57-58
- modelling\_memory\_limit\_percentage
  - options.cfg ファイル, 33
- nodes
  - SQL 生成をサポートする, 65
- ODBC データ ソース
  - Connect for ODBC と UNIX スクリプト, 78
  - および UNIX, 78
- options.cfg ファイル, 39
- Oracle
  - SQL 最適化, 57-58, 74
- password
  - IBM SPSS Modeler Server, 16
- performance
  - IBM SPSS Modeler Server, 51
- port\_number
  - options.cfg ファイル, 30
- program\_file\_path
  - options.cfg ファイル, 32
- program\_files\_restricted
  - options.cfg ファイル, 32
- RAM, 11
- rc.modeler (UNIX)
  - 場所, 76
- Secure Sockets Layer, 41
- SPSS Modeler Server, 1
- SQL
  - Oracle を最適化する, 74
  - クエリー, 72
  - 重複する列名, 72
  - 生成のプレビュー, 62
  - モデル ナゲットの表示, 63
- SQL 生成, 52, 57-58, 65
  - 有効化, 61
  - CLEM 式, 64, 69
  - IBM SPSS Modeler Server 用に有効化, 35
  - ストリームのリライト, 64
  - ヒント, 64
  - プレビュー, 54, 62
  - モデル ナゲットの表示, 63
  - ログ記録, 54, 62
- SQL Server
  - SQL 最適化, 57-58
- SQL のプッシュ バック : SQL  
生成を参照してください。 , 57
- sql\_generation\_enabled
  - options.cfg ファイル, 35
- @SQLFN 関数, 72
- SSL, 41
  - 概要, 42
  - 通信の保護, 42
- SSL データ暗号化
  - IBM SPSS Modeler Server 用に有効化, 36
- ssl\_certificate\_file
  - options.cfg ファイル, 36
- ssl\_enabled
  - options.cfg ファイル, 36
- ssl\_private\_key\_file
  - options.cfg ファイル, 36
- ssl\_private\_key\_password
  - options.cfg ファイル, 36
- stream\_rewriting\_enabled
  - options.cfg ファイル, 33

- temp\_directory
  - options.cfg ファイル, 32
- UNC ファイル名, 14
- UNIX
  - ファイル権限の構成, 78
  - ユーザーの認証, 21
  - ライブラリパス, 86
- UNIX カーネルの制限, 25
- UNIX スクリプト
  - 編集, 78
  - auto.sh, 76
  - modelersrv.sh, 76
  - rc.modeler, 76
- UNIX のカーネルの制限, 25
- URL プレフィックス, 47
  
- 管理者アクセス
  - IBM SPSS Modeler Server の, 28
- アプリケーションの例, 3
- アーキテクチャ
  - コンポーネント, 7
  
- 構成オプション
  - 並行処理, 33
  - 概要, 30
  - 自動 SQL 生成, 35
  - Coordinator of Processes, 37
  - COP, 37
  - IBM SPSS Modeler Server, 26
  - SQL 文字列の長さ, 35
  - SSL データ暗号化, 36
  - ストリームのリライト, 33
  - 一時ディレクトリ, 32
  - データ ファイル アクセス, 32
  - 接続とセッション, 30
  - パフォーマンスと最適化, 33
  - ポート番号, 30
  - メモリー管理, 33
  - ログ記録, 31
  - ログインの試行, 30
- オペレーティング システム
  - 64-bit, 10
  
- キャッシュの圧縮, 33
- キャッシング、データベース内, 39
  
- サーバー
  - サーバーの COP の検索, 20
  - 接続の追加, 18
  - ログイン, 16
- サーバー ポート設定
  - options.cfg ファイル, 33
- シングル サインオン, 18
  
- ステータス
  - IBM SPSS Modeler Server の Windows 上で, 24
  - UNIX 上で IBM SPSS Modeler Server の, 25
- 複数ストリームの実行, 35
- ストリーム実行のエラー, 35
- ストリームのリライト, 64
  - 有効化, 52
  
- セキュリティ
  - SSL, 41
  - UNIX 上でファイルの作成を構成, 78
  - ファイルの作成, 22
  
- ゾンビ プロセス、IBM SPSS Modeler Server, 26
  
- ディスク スペース, 10
  - 計算, 11
- 一時ディレクトリ
  - IBM SPSS Modeler Server の, 32
- データ アクセス パック
  - Connect ODBC、UNIX 上で構成, 78
  - UNIX 上での Connect for ODBC のトラブルシューティング, 82
  - UNIX の構成, 78
  - および UNIX ライブラリパス, 86
- データ ファイル
  - IBM SPSS Statistics, 14
  - インポートとエクスポート, 14
- データへのアクセス, 12
- データベース
  - アクセス, 12
  - サポートの段階, 12
- データベース接続
  - 閉じる, 40
- データベース キャッシング
  - options.cfg からの制御, 39
  - SQL 生成, 65
- データベース サーバー, 55
- データベース内キャッシング, 39
- 段階、データベースのサポート, 12
- データベース内マイニング, 55
  
- ドキュメンテーション, 3
- ドメイン名 (Windows)
  - IBM SPSS Modeler Server, 16
  
- 自動的なサーバーの起動
  - UNIX 上で構成, 77

## 索引

法律に関する注意事項, 93

接続の COP の検索, 20

紫色のノード, 52

SQL 最適化, 59

ノード キャッシング

SQL 生成, 65

データベースへの書き込み, 65

パス, 14

ハード ドライブ, 11

ハードウェアの要件

IBM SPSS Modeler Server の, 9

ファイアウォールの設定

options.cfg ファイル, 35

ファイル権限

IBM SPSS Modeler Server 上の, 22

UNIX 上で構成, 78

ファイル名

UNIX, 14

Windows, 14

一時ファイル, 10

権限 (IBM SPSS Modeler Server), 22

プッシュバック, 52, 57-58, 61, 65

CLEM 式, 69

プレビュー

SQL 生成, 62

プロセス、非応答, 26

プロセッサ, 9

複数, 33

ホスト名

IBM SPSS Modeler Server, 16, 18

ポート番号

IBM SPSS Modeler Server, 16, 18, 30

ポート設定

options.cfg ファイル, 33

メッセージ

生成された SQL の表示, 54, 62

メモリー, 11

メモリー管理

管理オプション, 33

モデル作成

メモリー管理, 33

モデル ナゲット

SQL の表示, 63

ユーザー ID

IBM SPSS Modeler Server, 16

ユーザー アカウント

IBM SPSS Modeler Server, 21

ユーザーの認証, 21

ログ ファイル

IBM SPSS Modeler Server の, 31, 48

生成された SQL の表示, 54, 62