

# IBM® SPSS® Modeler Advantage 7



注：この情報とサポートされている製品をご使用になる前に、「表示」（p.104）の一般情報をお読みください。

本版は IBM Analytical Decision Management 7，および新版で指示されるまで後続するすべてのリリースおよび変更に対して適用されます。

アドビ製品の画面コピーは、Adobe Systems Incorporated の承認を得て掲載しています。

Microsoft 製品のスクリーンショットは Microsoft 社の許可を得て転載しています。

Licensed Materials - Property of IBM

**© Copyright IBM Corporation 2010, 2012.**

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

---

# はじめに

## 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、完全で整合性があり正確な情報を実現します。 [ビジネス・インテリジェンス](#)、[予測分析](#)、[財務パフォーマンスと戦略の管理](#)、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が、現在の現在のパフォーマンスと将来の成果を予測する能力の明快、直近、そして、構想可能な洞察を可能にします。豊富な産業用ソリューション、実証された実践とプロフェッショナル・サービスと組み合わせて、あらゆる規模の組織は最高の生産性を推進し、決定を確信を持って自動化し、そして、よりよい結果を実現することができます。

このポートフォリオの一部として、IBM SPSS Predictive Analyticsソフトウェアは組織が将来のイベントを予測し、より良い成果を推進するためのその洞察に沿った積極的なアクションを取ることを支援します。IBM SPSSテクノロジーは、不正やリスクを軽減しつつ、顧客を獲得、維持、拡大するための有益な技術として、世界中の企業、政府、学術関連のお客様から信頼をいただいております。日常の業務にIBM SPSSソフトウェアを組み込むことによって、組織は予測能力を持った企業となります。-決定をビジネスのゴールに合わせて方向づけ自動化し、測定可能な競争力の高い優位性を達成することを可能にします。詳しい情報と営業担当者へのお問い合わせは、次のホームページをご覧ください<http://www.ibm.com/spss>。

## テクニカル サポート

保守契約をんでいるお客様は、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 サポートを受けるには、IBM Corp. 次のウェブサイトを<http://www.ibm.com/support>をご覧ください。ご利用の際には、お客様のお名前、組織名、およびサポート同意契約をご用意ください。

---

# 内容

<b>1</b>	<b>IBM Analytical Decision Managementについて</b>	<b>1</b>
	新機能	1
	IBM Analytical Decision Managementについて	4
<b>2</b>	<b>顧客応答のモデリングの例</b>	<b>5</b>
	IBM SPSS Modeler Advantage	5
	ダイレクトメールキャンペーンの場合の収益の最大化	5
	サンプルファイル	6
	データ	6
	サンプル	6
	アプリケーションの起動	7
	データソースのセットアップ	8
	モデルの構築	10
	モデルの評価	11
	予測ルールモデル	12
	モデルのテスト	13
	モデルのスコアリング	14
	クラスタモデリングの例	15
	アソシエーションモデリングの例	17
<b>3</b>	<b>ワークスペース内の移動</b>	<b>21</b>
	アプリケーションの起動	21
	プロジェクト設定	22
	ホームページ	23
	ギャラリー	23
	環境設定の設定	24
	プロジェクト、モデル、および、ルールの開封及び保存	25
	IBM SPSS Collaboration and Deployment Services Repositoryへの保存	26
	ラベルの操作	27
	プロジェクトのロック	28
	オブジェクトのプロパティ	29

## 4 データソースの管理 30

データのプレビュー	34
測定レベル	34
新規データソースの作成	35
ファイルソース	36
データベースソース	39
エンタプライズビューソース	40
IBM Cognos BIソース	41
追加フィールドの追加	43
入力フィールドの選択	45
式エディタ	45
フィールドの関連付け	46

## 5 ルールの操作 47

ルールの作成	47
セグメントルールの定義	47
選択ルールの定義	50
ルールのエクスポートと再利用	51
注釈を追加	53
ILOG Business Rule Management System のルールの適用	53

## 6 予測モデルの作成 55

予測モデルについて	55
モデルの作成	56
予測モデルの作成	56
予測モデリングの結果	57
予測ルールモデルの作成	58
クラスタリングモデルの作成	60
アソシエーションモデルの作成	62
オプションのモデル設定	64
モデルの評価	68
最大利益をシミュレート	75
のテストmodel.	78
アプリケーション上でのモデルスコアの使用	79

## 7 スコアリング モデル 81

データベース テーブル、ファイル、または、Cognos BI サーバーのモデルのスコアリング	81
スコアリングの宛先の選択	81
出力フィールドの選択	86
データおよびサブセットの選択	87
スコア対象のレコードの選択	87
モデルのスコアリング	88

## 8 アプリケーションの管理 89

一般的なオプション	90
アプリケーション ショートカットの制御	91
ホーム ページを非表示にする	91
データ オプションのロック	92
スコアリング宛先オプションのロック	93
外部ルールの作成に使用するプロジェクト メタデータをダウンロードします	93

## 9 ヘルプの利用 96

利用可能なドキュメント	96
-------------	----

## 付録

### A ストリームの共有 99

### B アクセス機能 102

ヘルプ アクセス機能	102
------------	-----

C 表示	104
索引	107





# IBM Analytical Decision Managementについて

## 新機能

IBM® Analytical Decision Management リリース7では以下の機能が追加されました：

**最適化**最適化の目標は個々のゴールに最適に合致するソリューションを識別することで、たとえば、販促キャンペーンからの売り上げを最大化する、或いは不正または解約のリスクを最小化することです。最適化の結果は、ある質問に対して、指定した母数内で可能性があるすべての組合せにわたって、最適化したい値（たとえば利益）を正に最大化する方法で回答できるソリューションです。

**クラスタリング モデル**。IBM® SPSS® Modeler Advantageでの新機能、クラスタリング モデリング 手法は、類似する値またはパターンのデータ レコードのグループを発見するのに役立ちます。このようなテクニックはマーケティング（顧客のセグメント化）や、その他のビジネス アプリケーション（1つのレコードのクラスタに含まれるレコードはエラーまたは不正のインスタンスを含むかも知れません）に使用されます。クラスタリング モデルは、対象フィールドを使用しないため、また、真または偽を評価できる特定の予測を返さないため、非監視学習モデルと呼ばれることがよくあります。IBM Analytical Decision Management により、クラスタモデルを自動的に作成できます。また、クラスタの結合、名前変更、修正または削除ができ、もしくは、以前のビジネス知識を基にしてカスタム化したクラスタを追加することもできます。

**アソシエーション モデル**。IBM SPSS Modeler Advantageでの新機能、アソシエーション モデリングにより、一緒に表示される可能性の高い項目を予測でき、それらの間の関係の強さを予測できます。アソシエーション ルールは、特定の結果（特定の製品の購入）と条件セット（複数の他の製品の購入）を関連付けます。

**データ**： データ タブには、複数のデータ入力を表示し、アプリケーションビューでプロジェクト データ モデルを定義し、データのようにモデル出力を参照し、ローカル データ ファイルをアップロードし、データをプレビューおよびカウントする機能が含まれています。

**IBM Cognos データ統合の改良点。**IBM® SPSS® Modeler 15 IBM Analytical Decision Management 同様、Cognos プレビルト レポート、Cognos フィルタおよび計算、および自動分解をサポートしています。

**複数の要素。**アプリケーションは、複数の次元階層を含むことができます。アプリケーションを起動すると、選択したアプリケーションが複数の次元階層要素を含むよう設定された場合に、アプリケーションで使用する 1 つまたは複数の使用可能な要素（階層）を選択するよう指示されます。左側の次元ツリーでは、返すことのできる判断または提案の候補の範囲を指定します。アプリケーションに複数の次元階層が含まれる場合、定義タブの次元ツリー アリアの個々のタブにその次元階層が含まれます。

**動的割り当て。**前のリリースでは、何かを割り当てるには、それが次元ツリー内にある必要がありました。今回は、子のない次元項目の場合、プロジェクト データ モデル上でキーによってリンクされたテーブルから項目を読み取ることができます。例えば、子のない 車 という次元項目があるとします。車 での動的割り当てを BMW、Audi および Ford などの表内の参照フィールドに設定できたとします。

**複数割り当て。**今回導入された 複数割り当て オプションにより、各セグメントに対する複数の次元項目（たとえば、複数のオファー）にレコードを割り当てることができます。複数割り当ては、ルールごとの 1 つの割り当ての代わりに、各ルールに対して行うことができます。

**ユーザー インターフェイスの改良点。**すべてのアプリケーションの使いやすさや拡張性を改善するため、IBM Analytical Decision Management ユーザー インターフェイスに対しては、多くの変更および追加がなされています。

**ギャラリの強化** 一度に複数のギャラリー項目を削除できるギャラリーのクリア など、様々な強化がギャラリーになされています。

**ルール エディタの強化。**ルール エディタの起動元に応じて、ユーザーは選択ルール（包含または除外）にルール タイプを設定したり、集計ルールにリスク ポイントを割り当てたり、割り当てルールに次元を選択したり、OR および Split OR コントロールを選択して、ルールないの複数の式を結合したり、OR 式を作成することなどが可能になっています。ルール作成時に、階層ツリーから次元を選択するか（複数のツリーのあるアプリケーションの場合）、または式エディタを起動して、式を作成できます。この機能は以前はルール エディタ ダイアログ ボックスでは使用できませんでした。

**拡張 ILOG Business Rule Management System の統合。**IBM® SPSS® Collaboration and Deployment Services Deployment Managerでは、[新たな外部ルールを追加]ダイアログボックスを使用して、IBM Analytical Decision Managementで使用可能なILOGルールへの参照を定義付けすることができます。以前は、これらのルールはウェブサービスを通してのみアクセスする

ことができました。このリリースによって、ILOGデータベース内でローカルルールを参照することが可能になりました。新たなローカルオプションでは、ルールは IBM® SPSS® Modeler Server (Deployment Managerで定義されたIBM Analytical Decision Managementに使用されるサーバー) 経由で実行されます。

**WhatIf? 強化。**WhatIf? ダイアログ ボックスの機能に様々な強化が加えられています。たとえば、シミュレーション実施日の場合、WhatIf? 評価を実行する日付の範囲を入力できます。

**ストリームファイルをアップロード。**ローカルファイル・システムからSPSS Modelerストリームファイルを開いて、必要に応じて、これをレポジトリに格納できます。

**複数の SPSS Modeler Server 資格証明。**Deployment Managerで定義された複数のSPSS Modeler Server資格証明がある場合、ここでもIBM Analytical Decision Managementで使用するために複数の資格証明を指定できます。次に IBM Analytical Decision Management ユーザーがログオンすると、サーバーがリスト内の 1 番目の資格証明で、ユーザーに権限があるか確認し、適切な権限のある資格証明が見つかるまで、リスト内で順に処理を進めます。ユーザーが順番付けられたリスト内でどの資格証明に対しても権限がない場合、そのユーザーは IBM Analytical Decision Managementを使用できません。これにより、どのユーザーが SPSS Modeler Server ファイル システムのどの部分にアクセスできるかをよりコントロールできるようになります。

以前のリリース、IBM SPSS Decision Management 6.2, では、以下の機能が追加されました：

**Cognos データのサポート。**SPSS Modeler Server 14.2 またはそれ以降と一緒に使用されると、IBM Analytical Decision Management はデータ ソースを定義する際に IBM Cognos Business Intelligence サーバーからデータ読み込む機能を追加し、スコアリング結果を Cognos BI サーバーに戻して書き込みます。Cognos BI サーバーの接続を定義した後、パッケージ、ディメンション、クエリー、そして、レポートを参照して、必要なフィールドを選択します。 [詳細は、4 章 p.41 IBM Cognos BI ソース を参照してください。](#)

**ILOG Business Rule Management Systemの統合。**ILOGなどのBusiness Rules Management Systemで作成されたルールは、現在のIBM Analytical Decision Managementプロジェクトに使用されるのと同じデータモデルをサポートするように開発されていれば、IBM Analytical Decision Managementアプリケーションで参照および使用することができます。これらの外部ルールはIBM Analytical Decision Managementアプリケーションの中で、他のモデルやルールと同じ方法で使用することができます。 [詳細は、5 章 p.53 ILOG Business Rule Management System のルールの適用 を参照してください。](#)

**System ZのLinuxサポート** SUSE Enterprise Server 10 は IBM System z10用の64-bit上でサポートされます。

## IBM Analytical Decision Managementについて

IBM® Analytical Decision Management は予測分析のメリットを実際の業務上の課題に適用します。これを使用すると、顧客や業界向けにカスタマイズされたカスタム アプリケーションを作成することができます。アプリケーションは、固有の課題を解決するように設定されていますが、すべて以下の共通機能をベースにしています。

- ビジネス ルールを使用して自動的に意思決定を行います。
- 予測モデルを使用して洞察を取得します。
- 上記に基づいて最善の判断に到達するために、優先順位付けまたはシミュレーションを使用します。

特定の業務上の課題を解決するようにカスタマイズされた、多数のパッケージ アプリケーションが用意されています。詳細は、お客様の営業担当までお問い合わせください。

# 顧客応答のモデリングの例

## IBM SPSS Modeler Advantage

IBM® SPSS® Modeler Advantage は、ビジネス ユーザーが予測モデリング機能を利用するための、使い易いアプリケーションです。予測モデルを使用すると過去に発生した事象からパターンを特定できるので、それらのパターンを使用すれば将来発生する可能性がある事象を予測することができます。

たとえば、収入、年齢、勤務先、会員資格などの特性に基づいて、解約の可能性が低い顧客や、特定のオファーに応ずる可能性が高い顧客を予測するためにモデルを使用することができます。結果としての予測は、戦略的プランニングに対する入力として、ターゲットとなる顧客または関心領域のリストを生成するために使用できます。また、予測アプリケーション内のルールと統合することもできます。

## ダイレクト メール キャンペーンの場合の収益の最大化

この予測モデルの例では、架空の銀行で、金融商品およびサービスの顧客へのマーケティングを担当していると仮定します。最新の試験的なキャンペーンでは、多数の顧客に複数の新規サービスの 1 つをオファーし、彼らの応答を記録しています。ただし、応答率は期待値よりも低く、キャンペーンを収益源とすることに失敗しています。

IBM® SPSS® Modeler Advantage の使用目的は、予想するコスト、収益、その他の要素のもとで、最適な顧客に的を絞ることで損失を利益に転換することです。

試験的なキャンペーンの結果の要約を以下に示します。

テーブル 2-1  
テスト キャンペーンの結果

テスト キャンペーンの結果	
メール数	3000
応答数	111
応答率	3.7%
コスト	6750
収益	6660
総利益（損失）	(90)

この例では、IBM SPSS Modeler Advantageに付属してインストールされたデータファイルbank\_customer\_data.txtおよびbank\_response\_data.txtを使用します。この例の完全バージョンは、bank\_response\_model.str ファイルで提供されています。インストールおよびサンプル ファイルについての詳細は、必要に応じて、管理者に問い合わせてください。 [詳細は、p.6 サンプルファイル を参照してください。](#)

## サンプルファイル

ユーザーズ ガイドで紹介している事例の再現に使用可能なIBM® Analytical Decision Managementには、データ及び他のサンプルファイルが付属しています。

## データ

インストール時に、多数のサンプル データ ファイルが IBM® SPSS® Modeler Server ¥Demos ディレクトリ（たとえば、C:¥Program Files¥IBM¥SPSS¥ModelerServer¥15¥Demos）にインストールされます。これらのファイルはおそらく事前構築されたサンプル アプリケーションで使用され、以下を含みます：

- insurance\_claim\_data.txt
- insurance\_fraud\_data.txt
- bank\_customer\_data.txt
- bank\_response\_data.txt
- retail\_purchase\_data.txt
- customer\_summary\_data.txt
- customer\_contact\_history.txt

これらのデータファイルを使用するサンプル・アプリケーションファイル、モデル、およびルールは上記のデフォルトの格納場所を前提としています。データ ファイルが別のドライブまたは場所にインストールされた場合は、データ ソース パスをそれに応じて更新する必要があります。

## サンプル

多数のサンプル・アプリケーションファイル、モデル、およびルールが、事前構築されたIBM® Analytical Decision Managementアプリケーションと一緒に使用できるように提供されます。サンプルは、アプリケーションのインストールディスクの¥Demoディレクトリにパッケージ化されているか、ダウンロード可能なファイルとして提供されます。以下のステップに従って、

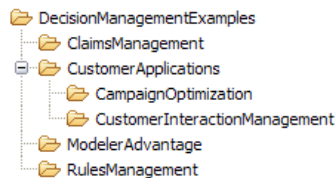
事前にビルドされたアプリケーション用のサンプル ファイルを IBM® SPSS® Collaboration and Deployment Services Repository にインポートします。

- ▶ **DVD.DecisionManagementExamples.pes** ファイルをDVDの¥Demo フォルダーからコピーして一時的な保存場所に保管します。

の要素または

- ▶ **ダウンロード ファイル**。SPSS\_Decision\_Mgmt70\_demo.zip ファイルを一時的な場所に保存して抽出します。
- ▶ IBM® SPSS® Collaboration and Deployment Services Deployment Manager を開き、コンテンツ サーバー接続にログオンします。
- ▶ コンテンツ・レポジトリフォルダーを右クリックして**インポート**を選択します。
- ▶ インポート・ダイアログボックスにて、**参照**をクリックして、以前に一時的な保存場所にコピーした**DecisionManagementExamples.pes** ファイルを選択します。[OK] をクリックします。
- ▶ インポート衝突ダイアログにて、**OK**をクリックします。インポートが完了したら、IBM Analytical Decision Managementにて開く事の出来る事例プロジェクトを含んでいる以下の新しいフォルダーを確認します。

図 2-1  
芸術品サンプルのフォルダー



サンプル のアプリケーション ファイル、モデル、ルール等はすべて、IBM® SPSS® Modeler Server ¥Demos ディレクトリ (たとえば、C:¥Program Files¥IBM¥SPSS¥ModelerServer¥15¥Demos) にインストールされることを前提としています。データ ファイルをそれ以外のドライブまたは場所にインストールする場合は、ローカル構成のデータ ソース パスを変更する必要があります。

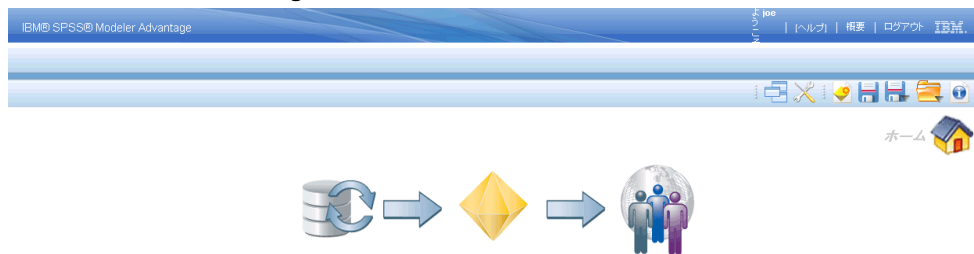
## アプリケーションの起動

IBM® Analytical Decision Management 起動ページには、サイトで利用可能なアプリケーションへのリンクが含まれます。これらのリンクは管理者が設定します。右上にあるリンクをクリックすれば、追加アプリケーションを選択することができます。

- ▶ 例の操作を開始するには、IBM Analytical Decision Managementの起動ページを開きます。将来のアクセスを容易にするため、ブラウザのこのページにブックマークを追加するか、デスクトップにショートカットを作成することができます。
- ▶ 新規モデルを作成するには、IBM® SPSS® Modeler Advantageグループ内のアプリケーション・ドロップダウン・リストで**新規**を選択し、**進む**をクリックします。IBM SPSS Modeler Advantageホームページが表示されます。

図 2-2

## IBM SPSS Modeler Advantage ホームページ



ホームページ上の各アイコンは、作成用のデータ定義から、モデルのスコアリングまで、モデリング プロセス内の異なるステップを表現しています。開始するには、左側のデータ アイコンをクリックします。

## データ ソースのセットアップ

[データ] タブによって例の中で使用されたサンプルデータファイルであるの設定を行うことが可能になってます。データ セットは、管理者が設定するか、またはユーザーが独自に追加します。

プロジェクト データ モデルでは、アプリケーションで使用するフィールドを定義します。これは、ルールやモデルの評価で必要になります。他のすべてのデータセットは、アプリケーションで使用するためには、このデータ モデルにマッピングする必要があります。

この例では、2 つのデータ ファイルを使用します。これらのファイルは共にアプリケーションに付属しています。

- **bank\_customer\_data.txt** には、顧客 ID、返済履歴、保有する商品の数、顧客となってから経過した月数、および年齢や収入のようなデモグラフィック情報など、現在の顧客に関するデータが含まれています。プロジェクト データ モデルは、このファイルをベースにします。
- **bank\_response\_data.txt** には、以前のキャンペーンで 1 つ以上のサービスをオファーした顧客と、その応答を記録した情報が含まれています。アプリケーションは、このデータを分析することで、将来のキャンペーンで類似のオファーに対して応答する確率の高い顧客を特定します。



デフォルトでは、サンプル データは、お使いのアプリケーションからアクセス可能な IBM® SPSS® Modeler Server Demos フォルダにインストールされます（たとえば、C:\Program Files\IBM\SPSS\ModelerServer\15\Demos）。これらのファイルのインストールについての詳細は、必要に応じて、管理者にお問い合わせください。

### プロジェクト データ モデルの定義方法

- ▶ [プロジェクト データ ソース] パネルで、**データソースを追加** をクリックします。
- ▶ bank customer dataをデータソース名として指定してください。
- ▶ **ファイル** を選択し、bank\_customer\_data.txt。
- ▶ **入力フィールドの指定** の下で、不良返済レコードがある フィールドの測定タイプが**フラグ型**になっていることを確認します。フラグの設定が、その属性の有無を示すための、真および偽を表す（一般的にはT and Fもしくは数値フィールドの0と1にコード化されます）2つの値のどちらかを持つことを意味します。
- ▶ **入力フィールドの指定**の下で、性別フィールドの測定タイプが**名目**になっていることを確認します。（このフィールドの値は、単に、性別が男性か女性かを示すだけであり、フラグ型の場合とは異なって、真/偽 を表現するものではありません。）
- ▶ **保存** をクリックします。

これはプロジェクトで最初に定義するデータ ソースなので、プロジェクト データ モデルでは自動的に選択され、アプリケーションで必要なフィールドを定義するために使用されます。その他のデータ セットはすべて、このソースに相対的にマップされます。

### 応答データ ソースを定義するには、以下の操作を行います。

応答データセットは、プロジェクト データ モデルで使用されるフィールドと同じフィールドの多くを含み、さらに 2 つのフィールド キャンペーン ID および 応答 を含んでいます。これらのフィールドは、各顧客に提示したオファーと、応答の有無を示します。これは、将来、応答する確率の高い顧客を予測するためのモデルを作成するとき使用されます。

- ▶ [プロジェクト データ ソース] パネルで、**データソースの追加** を選択します。
- ▶ データ ソース名として bank response data を指定します。
- ▶ **ファイル** を選択し、bank\_response\_data.txt を選択します。

- ▶ **入力フィールドの指定** の下で、**応答** フィールドの測定タイプが**フラグ**になっていることを確認します。
- ▶ **指定された入力フィールド**の下で**キャンペーン ID**フィールドのとなりの**チェック** ボックスの選択を解除することは、このフィールドはモデリング時には入力に使用しないことを示します。（このフィールドは、ある顧客が過去にどのキャンペーンに選ばれたかを示すために使用することができますが、将来の応答を予測するには不便です。）
- ▶ **保存** をクリックします。

データモデル内のいくつかのフィールドは応答データセット内には含まれないため、[データ] タブの **互換** 列の下のこのデータ ソースにはオレンジ色のアイコンが付けられており、マッチするフィールドはあるが、すべてのフィールドがマッチしているわけではないことが示されます。このデータ セットはモデリングには適していますが、シミュレーション、スコアリング、または導入には適していません。（これらの目的で使用するには、データ モデルのすべてのフィールドが揃っている必要があります。）

## モデルの構築

[モデリング] タブを使用すると、モデルの作成と評価を行うためのデータ ソース、ターゲット フィールド、その他のオプションを指定することができます。

- ▶ データソースとして、**bank response data**を選択し、ターゲットフィールドとして、[**応答**]を選択します。
- ▶ **データスキャンが必要**ですというメッセージが表示されたら、アイコンをクリックしてデータスキャンを行います。その後**モデルの作成**をクリックします。

モデルが完成すると、分布図が表示されます。分布図では、観測値が横軸に表示され、値 1 が見込みのある応答を示します。

- 棒の高さは、応答率を反映します。応答データセット内の 3,000 人の顧客のうち、111 人が応答しているので、ヒット率は 3.8% です。
  - 各棒内の色付けされた領域は、モデルが予測した応答を表現しています。すなわち、右上の凡例に示したように、赤は見込みのある予測を示し、青は見込みのない応答を示します。応答率が低いいため、予測した応答（赤の領域）の数は多くありません。しかしながら、「偽」のバー（応答=0）は圧倒的に青になっています。これは、応答しないという予測の圧倒的の大多数が、実際に応答しなかったことを示しています。
- ▶ 如何にモデルが実行されるかについての良いアイデアを得るには、次のセクション（**モデルの評価**）に進みます。

## モデルの評価

[モデリング] タブで、[評価]をクリックし、モデルを顧客データに適用したときの動作を調べれば、モデルが最終損益に与える影響を理解することができます。

- ▶ データ ソースとして、bank response data を選択します。
- ▶ 作成するグラフ タイプに、[差益グラフ]と[利益グラフ]を選択します。
- ▶ [最大利益のシミュレート]を選択し、母集団として3000を指定します。これはテスト キャンペーンの数と一致します。
- ▶ コストに2, 25を、収益に60を指定します。これらの値は、提示する各オファーのコスト（企画、処理、郵送費用など）と見込みのある各応答の期待収益を示しています。
- ▶ [検索する応答]に、1を選択します。
- ▶ [実行]をクリックします。

差益グラフには、モデルを使用することで取得した改善または「上昇分」が示されます。モデルの作成時に（[オプション] 設定の下で）[データを自動的にパーティション分割]オプションを選択していたため、トレーニング パーティションとテスト パーティションには異なるグラフが表示されます。各グラフには、サンプル内のヒット総数と比較した、各パーセンタイルにおけるヒット（見込みのある応答）数が示されます。斜めの線は、モデルを使用しなかった場合の期待される応答をプロットしたものです。半分のヒットで良ければ、半数の顧客に連絡を取れば良いことを示しています。曲線は、モデルを使用した場合の改善を示しています。応答率の高い60%の顧客に的を絞ることで、メールの数を減らすにも関わらず、ほぼすべてのヒットを獲得できることが示されています。

差益グラフは、指定したコストと収益の値に基づいて、総利益の観点から差益を示します。モデル学習に使用した例では、30番目の百分位のすぐ下で収益が最高になります。

この機能を理解するには、[シミュレート] タブをクリックして、指定した母集団のサイズに対する期待最大利益を調べます。期待利益は、次の式を使用して算出されます。

$$\text{期待利益} = (\text{反応率} * \text{収益}) - \text{コスト}$$

ここで反応率は、モデルに基づいて各顧客ごとに決定されます。モデルによって反応率が最も高いと判定された顧客にターゲットを絞ることにより、顧客あたりの収益を最大化することができます。含まれる顧客数を増やせば、ある一定の点までは総収益も増加しますが、反応率の低い顧客が多く含

まれるようになると、顧客ごとの収益は徐々に減ります。シミュレーションは、最大利益を獲得するために線引きをする場所の判断を支援します。

「利益を比較」では、この結果をさらに詳しく説明します。

- **対象を28%に絞る。** この列は、モデルで予測された反応率の高い 28% の顧客に対象を絞った場合の結果を示しています。反応率が最も高い 840 人の顧客を選択的に含めることにより、収益を最大化することができ、結果として期待される収益は \$3,450.00 になります。
- **ランダムに 28%** 同じ 28% の顧客数でも、それをランダムに抽出した場合の結果を示します。この例では送付する数は同じ (840) ですが、モデルで提供される付加情報を利用していないため、期待される応答数は少なくなります (テスト キャンペーンの率に一致)。その結果、全体の応答率はオファー送付コストを回収できるだけの十分な値とならず、損失が発生させることとなります。
- **すべての顧客** 3000 名の顧客をすべて含めた場合の結果を示します。この場合、テスト キャンペーンの結果と正確に一致します。ここでも、モデルの有益な情報を利用していないため、損失が発生しています。

結果を見ると、モデルを使用した場合に応答率が改善され、損失を生みだしていたキャンペーンを利益を生むキャンペーンに転換できることがわかります。予測モデルを使って得られる利点が、\$90.00 の損失と \$3,450.00 の利益の差に表されています。この収益は、テスト キャンペーンで使用した 3000 人の顧客をベースにしたものですが、対象の人数をもっと増やすならば、それだけ受ける恩恵も大きくなります。

- ▶ シミュレーションの結果をレビューした後、[予測利益]グラフの下のスコアリングに使用を選択し、指定されたパーセンテージ (この事例では28%) に基づいて [スコア]タブの設定を更新します。 [詳細は、 p.14 モデルのスコアリング を参照してください。](#)

## 予測ルールモデル

予測されたモデルは、必要とするすべての予測機能を提供しますが、場合によってはモデルの内部的な動作に関してもう少し自由にコントロールしたり、仕組みを理解したいと考えることがあるかもしれません。そのような場合には、予測ルール (対話型) モデル作成を使用できます。予測ルールモデルで、データをセグメントする為にマニュアルでルールを作成出来たり、或いは選んだフィールドの本当の値の高低いずれかの集中があるセグメントを自動的に見つけられるようになります。例えば、キャンペーンに肯定的な反応を示した顧客を探すかもしれませんし、高い確率で反応が返ってくるセグメントを特定します。

- ▶ モデルタブに戻って、**モデルを変更**をクリックして**予測ルールモデル**を選択します。

- ▶ データソースとして、bank response dataを選択し、ターゲットフィールドとして、[応答]を選択します。
- ▶ [検索する応答]に、1を指定します。これは、見込みのある応答を行った顧客を検索することを意味します。
- ▶ [作成を開始]を選択します。
- ▶ [モデルの拡張]を選択し、より高い応答率を持つセグメントを特定します。最高の応答率のセグメントを特定するための3つの最善のルールが、作成され、表示されます。
- ▶ 包含/除外 アイコンをクリックして、各ルールによって選択された顧客の包含または除外を選択します。たとえば、2番目と3番目のルールセグメントは応答率が8%強です。これはキャンペーン全体の3.7%よりは高率ですが、1番目のセグメントの15.6%ほどは高くありません。これらの2つのルールの包含または除外が、全体の収益にどのような影響を与えるを見るためには、両方の収益シミュレーションを行うことができ、その結果を以下の通り比較します。
- ▶ 3つのセグメントをすべて 包含に設定し、[評価]をクリックします。
- ▶ 作成するグラフタイプに、[差益グラフ]と[利益グラフ]を選択します。
- ▶ 最大利益をシミュレートを選択して母集団に3000を、コストに2.25、そして収益に60を指定します。これらは、予測モデルで評価した際に使用した仕様と同じです。[検索する応答]には自動的に値が入力されます。
- ▶ [実行]をクリックすると、グラフの作成が完了した時点で、[シミュレート]タブへ移動します。

このモデルでシミュレートされた最大の収益は \$3172.00 で、自動化されたモデルの同じ事例よりも若干下回ります。必要に応じて、[評価]ダイアログを閉じて、2番目および3番目のセグメントの設定を除外に変更し、収益シミュレーションを再実行して結果がどのように変わるかを確認してみてください。予測ルール（対話型）モデル作成の利点は、ユーザーが動作をコントロールし、理解を深めることができる点にあります。

## モデルのテスト

モデルの作成が完了したら、一部の顧客レコードをテストして、キャンペーンへの応答率を調べることができます。

- ▶ モデルのスコアリングによる特定のレコードの処理をプレビューするには、[モデリング]タブで[テスト]を選択します。[テスト]ダイアログが表示されます。
- ▶ [レコード]領域で、[利益]の数値を100に増やします。

- ▶ [表示フィールドを選択]領域で、[顧客 ID] を選択します。さらに \$XF-Response および \$XFRP-Response フィールドが選択されていることを確認してください。これらのフィールドは、各顧客の応答予測（この事例では、0 が「いいえ」、1 が「はい」）と傾向スコアを表示します。

必要に応じて、[年齢]、[最後のオファーから経過した週の数]、[不良な支払履歴あり]などのフィールドを追加することもできます。これらのフィールドは、特定の顧客がモデルにより一定のスコアリングをされる理由と仕組みを理解するのに役立ちます。

## モデルのスコアリング

モデルをスコアリングするとはモデルを対象のデータソースに適用することです。例えば、どの顧客が最も回答する可能性が高いかを予測するモデルを生成した場合、見込み顧客をスコアリングして、キャンペーンに含めたい人間を特定できます。[評価]ダイアログボックスで、利益を最大化するためにスコアリングを行うレコードのパーセンテージが決定されました。そして、この結果に基づいて、スコアリングするレコードのパーセンテージを指定しました。

- ▶ [スコア] タブの選択
- ▶ [出力フィールドを指定]で、[顧客 ID] フィールドが選択されていることを確認します。
- ▶ スコアリング結果の送信先を指定します。これはデータベース テーブルでもファイルでもどちらでも可能です。（単にスコアをプレビューしたいだけなら、このステップは飛ばしてください。）
- ▶ スコアリングの下では、設定は、モデルを評価した際に実行した利益シミュレーションに基づき、レコードの上位28%をスコアリングするようにすでに構成されているはずです。（シミュレーション結果を参照する際に[スコアリングに使用]ボタンを選択していない場合は、設定をここで更新することができます。）
- ▶ ダイアログの最下部にある[プレビュー]をクリックして、スコアリングで得られた結果のサンプルを表示します。

傾向に基づかないスコアを選択したので、最も応答の傾向が大きい顧客のみが含まれています。顧客IDを顧客データベースに反映するリンクによって、この結果によってメーリング リストを生成することもできたでしょう、

- ▶ [スコア]をクリックします。モデルが顧客データに適用され、ダイアログの下部に確認メッセージが表示されます。

## クラスタモデリングの例

クラスタリングモデルで、類似の特徴のグループごとに記録を分類できます。これは、データ内の自然グループを見つけるのに便利です。例えば、デモグラフィックな特徴や購買態度に基づいて顧客をセグメント化するかも知れません。

この例はbank\_customer\_data.txtという商品に付属してインストールされていたデータファイルを使用しています。この例の完全バージョンは、bank\_cluster\_model.strファイルで提供されています。インストールおよびサンプルファイルについての詳細は、必要に応じて、管理者に問い合わせてください。 [詳細は、 p.6 サンプルファイル を参照してください。](#)

- ▶ アプリケーション起動ページに戻って新しいIBM® SPSS® Modeler Advantageプロジェクトを作成します。
- ▶ データタブにて、プロジェクトデータソースパネルにて、[データ ソースのセットアップ p.8](#) で解説されている通りに、bank customer dataという新しいデータソースを追加します。
- ▶ モデリングタブにて、[モデルを変更](#)をクリックして、[クラスタリングモデル](#)を選択して、[保存](#)をクリックします。
- ▶ データ ソースとして、bank customer data を選択します。
- ▶ オプション設定セクションを展開します。クラスタリングモデルでは、評価フィールドの選択、或いはクラスタリングの必要な範囲の設定の為に、追加の[自動クラスタオプション](#)のオプション設定が使用可能になります。

クラスタリングモデルが作成された時、場面の後ろでは、いくつかの異なるクラスタリングモデルが実際には作成されます。これらのモデルはその後品質の尺度によって比較、そしてランクされ、最高モデルは自動クラスタモデルの使用の為に選択されます。よってこれらの2つの自動クラスタオプション設定はどのモデルが実際に選択されているのか制御できるようにしてくれます。評価フィールドがここで指定される場合は、評価フィールドの値を上手に識別するモデルが選択されます。そして必要な範囲が設定されたら、指定された範囲外にクラスタを発見するモデルは排除されます。

- ▶ この例として、[発見されたクラスタの必要な範囲を設定](#)オプションを選択して最低で3、最高で15のデフォルトを使用します。これでクラスタリングモデルが持つクラスタの数が少な過ぎたり、多過ぎたりする事がないようにします。
- ▶ 残りのオプション設定は他のモデリングタイプのもと同じです。[信頼できるモデル作成のために自動的にデータをクリーンアップ](#)オプションの選択を取り消して、ただ自動クラスタリングの例を簡単に理解できるようにします。
- ▶ オプション設定パネルを閉じてクラスタパネルを展開します。

クラスタパネルには、手動クラスタのセクションと自動クラスタのセクションがあります。手動クラスタはデータの知識に基づくクラスタの定義を可能にしてくれます。この例として、手動クラスタを高価値顧客に、そして手動クラスタを若い独身の顧客に、次のようにして定義します。

- ▶ **新しいルールを作成**アイコンをクリックします。ルールを高価値顧客という名前にして、次を式に追加して、OKをクリックします。

顧客としての月>12

収入>65000

- ▶ 次の式で若い独身の顧客という新しい別のルールを作成します。

年齢<35

結婚暦 = 独身

- ▶ **手動クラスタセクション**の下に2つの新しいルールがリストされます。これらの手動クラスタに何人の顧客が捕らえられているかを知るには、**クラスタレコードカウント**をクリックします。その後手動クラスタを再度展開して、カウント列を見て各クラスタに何人の顧客が見つかったかを確認します。67人の高価値顧客と33人の若い独身の顧客が確認されたことがわかります。
- ▶ 次に**自動クラスタ検出**をクリックして関心のある他のクラスタを自動的に検出します。終了したら、**自動クラスタセクション**を展開します。自動クラスタが3つ確認されたことがわかります。
- ▶ **自動クラスタを見る**をクリックして更に詳細を見ます。自動クラスタ結果ビューアが表示されます。

モデル要約ページは使用された自動クラスタモデルのアルゴリズム（この例ではTwoStep）に関する基本情報、検出された自動クラスタの数、そしてクラスタの品質の基本的な考えを提供します。

他のグラフも使用可能です。各グラフの様々な区域をうろついて更なる詳細を確認したり、いくつかのグラフ（例えば予測の重要度グラフ）と相互に影響できます。

クラスタビュー（一番最近のグラフ）を使って、クラスタを比較して興味深いパターンを探し出す事も出来ます。現実世界の実行においては、自動クラスタを研究してその後より意味のある名称を（クラスタビューの各クラスタラベルをクリックして）与えます。クラスタラベルの名称を更新したら、新しい名称がすべての他のグラフにて使用され、モデリングタブに戻ります。そして**評価**に進んで、モデルを**テスト**し、満足したら最後に採点します。



クラスタモデリングに関する更なる詳細は、[クラスタリング モデルの作成 p. 60](#) をご覧下さい。

## アソシエーションモデリングの例

アソシエーションモデリングは予測分析における重要な技術です。アソシエーション モデリングにより、一緒に表示される可能性の高い項目を予測でき、それらの間の関係の強さを予測できます。例えば、現在の買い物カゴにあるアイテムを元に次の可能性の高い顧客の購入を特定するかも知れません。

この例はbank\_purchase\_data.txtという商品に付属してインストールされていたデータファイルを使用しています。テレビ、パソコン、そしてスマートフォンといった電化製品に関する購入情報で構成されているデータ。

この例の完全バージョンは、retail\_association\_model.strファイルで提供されています。サンプルファイルのインストールについての詳細は、必要に応じて、管理者に問い合わせてください。 [詳細は、p.6 サンプルファイルを参照してください。](#)

- ▶ アプリケーション起動ページに戻って新しいIBM® SPSS® Modeler Advantageプロジェクトを作成します。
- ▶ データタブにて、プロジェクトデータソースパネルにて、retail purchase dataという新しいデータソースを追加します。ファイルを選択し、retail\_purchase\_data.txtを選択します。
- ▶ モデリングタブにて、モデルを変更をクリックして、アソシエーションモデルを選択して、保存をクリックします。
- ▶ データソースとして、retail purchase dataを選択します。
- ▶ データフォーマットは、表形式を選択します。トランザクションデータの設定は包含されたアイテムを表示するだけでなく、すべてのアイテムが単一の列に現れます。しかしこのデータセットにおいては、各アイテムは、購入されたかそうでないかを表示する真、或いは偽のフラグのある別の列にあります。
- ▶ モデルの作成 をクリックします。製作に失敗しました、しきい値が高すぎますというエラーメッセージを受け取ります。エラーを閉じる。

アソシエーションルールを求める時は、一般的に適用可能、信頼できる、そして簡単なルールを捜し求めることが普通は良い考えです。

- ▶ オプション設定セクションを展開します。アソシエーションルールのオプションの下では、最初のしきい値を10から1に変更して、2番目のしきい値を80から1に下げます。普通は、異なるしきい値を実験してみて、そ

の度にモデルを再構築して、ルールが多過ぎたり、足りなかったりしないようにバランスをとります。

- ▶ **モデルの作成** をクリックします。

モデルの結果において、各行は別個のアソシエーションルールを表現しています。リストの一番最初のルールは顧客がスマートフォンを購入する場合、大画面テレビも合わせて購入する場合があると説明します。サポート列は15%（或いは顧客6人につき1人）がスマートフォンを購入し、確信列が購入した顧客を表示し、それらの顧客が大画面テレビを購入する確立は約30%であるということを表示します。

よって、少数のルール（この例では239）でも、ビジネスの文脈から次に顧客が何を購入するかを予測を立てたりするのに使えるかも知れません。

- ▶ 次に、**テスト**をクリックしてその後**実行**をクリックして予測や推薦がどのように成されるのかを確認します。

この場合、モデルは最高確信数値のルールを見て、ユーザが既に購入したのを見て、そしてその後これら2つを元に顧客が何を購入するかの推薦を行います。

- ▶ 最初の記録の横側にあるプレビューアイコンをクリックして、例えば大画面テレビとスピーカーを最初に購入したけど、他の商品はまだ何も購入していない顧客を確認したりします。プレビューを閉じる。

この顧客への推薦は2列目に確認できるように標準テレビの購入である。3列目の値。29とは顧客が大画面テレビとスピーカーを購入した顧客が標準テレビも購入する可能性が29%であるという意味である。

しかしながら、これが例外なのか一般的に適用可能な予測されたトレンドをどうやって判別するのか？ルール番号（この例では115）を記録すれば、戻ってきてルールの更なる詳細を確認できる。

- ▶ テストダイアログを閉じる。
- ▶ 115という数字を確認するまでモデルの結果をページをめくります。結果に関してたくさんのページがある大きなモデルは、IDでルールを検索をクリックしてルールの番号を入力します。

- ▶ 結果列のヘッダーのとなりの矢印をクリックして選択する  
列 > インスタンス

ホームページが表示されます。事例数によると、ルール115が3つの商品と一緒に購入した58人の顧客に基づいているので、これが例外ではないということが明らかになりました。これは我々が将来の推薦の基本として使うことに十分な自信を与えてくれます。

どうやって推薦に使用できるかを確認する単一のルールを見ています。しかしIBM® Analytical Decision Managementはまた、モデルを我々の特定の目的に、より関連付けるようにカスタマイズする強力なオプションも提供します。例えば、どの顧客が大画面テレビの購入に関心があるのかについて予測したいかも知れない。

- ▶ [適用] をクリックします。これはフィルターを定義して、フィルターに一致しているルールのみがモデルに表示されるダイアログボックスを表示します。
- ▶ フィルターを有効化を選択して、アイテムに大画面テレビを選択します。
- ▶ 保存をクリックして、再度保存をクリックしてモデルの結果に戻ります。

ルールの番号は239から34に落ちました。

程よい数の顧客に基づいたルールのみを使用していることも確認したいかも知れませんが、これを達成するにあたって、10人以下の修正に基づいたルールを除外するかもしれません。

- ▶ 事例列のヘッダーのとなりの矢印をクリックして昇順で並び替えるを選択します。
- ▶ 除外列では、10以下の事例のすべてのルールを選択します。
- ▶ この時点で多くのルールがフィルタ、除外されたとはいえ、まだモデル内にあるということに注意して下さい。モデルが採点された時に使用されないように永久に削除するには、新しいモデルを作成 - 削除、除外、&フィルターされたルールをクリックして、OKをクリックします。

返還されたルールの数は10、或いはそれ以上の事例全部、そして大画面テレビを予測した全てを元にした20にまで減少した。

- ▶ 適用可能なルールの数を確認するには、テストをクリックして、その後実行をクリックします。大画面テレビは今ではさらに数名の顧客に推薦されていて、また何人かは他よりも大きな自信を持っています。
- ▶ テストダイアログを閉じる。
- ▶ オプション設定セクションでは、スコアリングオプションのセクションを展開してアソシエーションモデルに特定したスコアリングオプションを表示します。

デフォルトでは、予測最大数は3です。これは最大確信数値に基づいたり、或いは事例の最大数に基づいてルールを選択するルール基準の変更が可能になり、また最大数のリフト（購入された特定のアイテムの確立の上昇）を元にしたルールをします。また大画面テレビを予測を表示させないにより既に1台購入済みの顧客に提供しないようにしたいかも知れません。

デフォルト・スコアリング・オプションはほとんどの状況に適しています。しかしいくつかの場合において、これらのオプションが提供する追加の柔軟性は有利に働くかも知れません。

このプロセスにおける次の段階はスコアリングタブを使ってモデルを判定します。詳細は、[7 章 p.81 データベース テーブル、ファイル、または、Cognos BI サーバーのモデルのスコアリング](#) を参照してください。

アソシエーションモデリングに関する更なる情報は、[アソシエーション モデルの作成 p.62](#) をご覧下さい。

# ワークスペース内の移動

## アプリケーションの起動

アプリケーション起動ページは、インストールしたアプリケーションへの設定可能リンクを備えています。各パネルはインストールされているアプリケーションを表現します。ドロップダウンをクリックして、最近保存されたプロジェクト、モデル、またはルールから選択します。特定のラベルが付いたバージョンを指し示す 1 つ以上のカスタム ショートカットを追加することもできます。

- ▶ 起動ページにアクセスするには、`http://hostname:port/DM` に移動します。ここで、`hostname` はリポジトリ マシンの IP アドレスまたは名前、`port` はアプリケーション サーバーのポート番号です。この情報を知らない場合は、管理者に問い合わせてください。また、ブラウザのこの画面をブックマークに追加したり、デスクトップにショートカットを作成することもできます。
- ▶ 新規プロジェクト、モデル、または、ルールを作成するには、必要なアプリケーションのドロップダウン リストで **新規** を選択し、**進む** をクリックします。
- ▶ 保存したプロジェクト、モデル、または、ルールを開くには、最近使用したバージョンまたはショートカットを[アプリケーション]パネルのドロップダウン リストから選択するか、または、ドロップダウン リストから[ブラウザ]を選択して異なるオブジェクトまたはバージョンを選びます。リスト内の項目のリポジトリの場所を素早く見るには、ドロップダウン リスト内の該当する項目にカーソルを合わせます。
- ▶ ドロップダウン リスト内に常に表示されるカスタム ショートカットを追加するには、アプリケーション パネルの右上にある三角形の矢印アイコンをクリックします。オプションで、表示するショートカットの数を指定することができます。

### アプリケーションを開くための別の方法

- ブラウザのアドレス バーに URL をコピーまたは入力するか、ブラウザの保存されたブックマークかお気に入りのリストに URL を追加します。
- URL を指し示すデスクトップ ショートカットを作成します。
- 会社のイントラネットまたはポータル内からアプリケーションの URL にリンクします。

### 起動ページのカスタマイズ

アプリケーション起動ページの右上にあるコントロールを使用すると、アプリケーションの追加と、パネル ビューとリスト ビューの表示のトグルを行うことができます。また、パネルをドラッグして、ページ内で位置を変更することもできます。

- ▶ 新しいアプリケーション パネルを追加するには、アプリケーション起動ページの右上で **アプリケーションの追加** を選択します。(追加できるアプリケーションが存在しない場合、このリンクはありません。)

利用可能なアプリケーションのリストが表示されます。オプションで、1 つ以上のラベル付き バージョンにショートカットを作成することができます

- ▶ 起動ページのアプリケーション パネルの位置を変更するには、パネルのタイトル バーをクリックし、マウス ボタンを押しながら、パネルを新しい位置にドラッグします。
- ▶ アプリケーションを削除するには、アプリケーション パネルの右上隅にある削除アイコン (赤色の X) をクリックします。そのアプリケーションを指していたショートカットもすべて削除されます。(削除アイコンは、管理者によってアプリケーション パネルがページ上でロックされているときは使用できません。)
- ▶ 保存していないプロジェクト、モデル、および、ルールは起動ページのギャラリーにリストされるかもしれませんが、このギャラリーには、まだ作成中のモデルや、作成は完了したけれどもまだ保存されていないモデルが表示されます。 [詳細は、 p.23 ギャラリー を参照してください。](#)

## プロジェクト設定

定義タブを含むアプリケーションを起動すると、選択したアプリケーションが複数の次元階層要素を含むよう設定された場合に、アプリケーションで使用する 1 つまたは複数の仕様可能な要素を選択するよう指示されます。左側の次元ツリーでは、返すことのできる判断または提案の候補の範囲を指定します。アプリケーションに複数の次元階層が含まれる場合、定義タブの次元ツリー アリアの個々のタブにその次元階層が含まれます。定義タブおよび次元ツリーの詳細については、 [を参照してください。](#)

- ▶ **プロジェクト設定** ダイアログには、アプリケーション テンプレートで定義された使用できるすべての次元階層が表示されます。プロジェクトを定義するのに使用したい 1 つまたは複数の次元階層 (要素) を選択します。各要素の説明は、アプリケーション テンプレートでアプリケーション デザイナーが定義したように使用できる場合があります。

要素名は、最上位の次元の名前になります。次元は 1 つの要素にのみ所属することができます。

- ▶ 複数の要素を選択した場合、アプリケーションに表示する順番や互いに依存できる要素を並び替えることができます。

個々に指定する順番は重要です。要素は順番付けられたタブに表示され、前のタブの階層メンバーは次のタブ内のルールで参照できますが、次のタブは参照できません。例えば、2 番目および 3 番目のタブのルールは、1 番目のタブ内の階層メンバーを参照できますが、1 番目のタブは、2 番目および 3 番目のタブのメンバーは参照できません (2 番目のタブのルールは 3 番目のタブのメンバーを参照できません)。

- ▶ **プロジェクト設定** ダイアログで、OK をクリックします。

## ホームページ

アプリケーションは、ホームページ上に、アイコンで表現したステップバイステップのワークフローを用意しています。アイコンをクリックすれば、該当するステップにジャンプします。

**データ:** アプリケーションが使用するデータ セットの定義

**モデル:** 分析データを使用してモデルを作成し、予測の重要度、分布、ゲイン グラフ等を使用してそのモデルを評価してデータ内に隠されているパターンを発見することによりデータを理解し、その理解を判断の精緻化のために使用します。

**スコアリング:** モデルをスコアリングします。

## ギャラリー

ギャラリーは[アプリケーション起動ページ]に表示されます。ギャラリーには、一晩かけて作成するモデル (モデルの作成を起動し、ブラウザを閉じてから帰宅し、翌朝結果を調べる) のような未保存の作業がリストされます。また、ブラウザを誤って閉じた場合に、作業を検索するためにギャラリーを使用することもできます。ギャラリー内の項目には、アプリケーションに応じて、プロジェクト、ルール、モデル、データ スキャン、評価、その他長時間実行中のタスクが含まれます。

- ▶ 項目の名前に下線が付けられている場合は、名前をクリックすると、関連項目を開くことができます。たとえば、未保存のデータ スキャンがある場合は、ギャラリー内のデータ スキャンのタイトルをクリックすると、[データ] タブを開くことができます。
- ▶ ギャラリー内に項目を表示する必要がない場合は、削除アイコンをクリックして削除します。項目を完全に削除してもいいのかわ確認するための警告メッセージが表示されます。あるいは、幾つかのギャラリー項目を一度

に削除する場合は、[ギャラリーをクリア] をクリックします。[ギャラリーをクリア] ダイアログボックスを使用すると、すべてのオブジェクトを作成するタブを選択できます。

- ▶ ギャラリーのプロジェクトとシミュレーション タブの、[中に作成] ドロップダウンにより、各アプリケーションに対してタブに表示される項目をフィルタリングできます。たとえば、アプリケーションでの作業中に、その他のアプリケーションに対してギャラリー項目を見たくない場合があります。

## 環境設定の設定

ニーズに応じて IBM® Analytical Decision Management をカスタマイズするための調整可能な設定項目が複数あります。このカスタマイゼーションは、主として、作業の作成、テスト、保存を行う際に同じアクションを繰り返し選択することを避けるための独自のデフォルト オプションの設定で構成されます。環境設定は、設定したユーザーに固有であり、そのユーザーが使用するアプリケーションのみに適用されます。

アプリケーション内からユーザー環境設定の設定ダイアログにアクセスするには、ツールバー上の [ユーザー環境設定] アイコンをクリックします。

図 3-1  
ユーザー環境設定の設定



以降のセクションで説明するように、各領域には独自の環境設定を指定することができます。

### 一般的な設定

**アプリケーション ショートカット内の最近開かれたオブジェクトのデフォルト数:** アプリケーション起動ページのドロップダウン リストに表示する項目の数を選択します。

**日付の形式:** 日付を表示または記録するときに使用するフォーマットを選択します。

**時刻の形式:** 時刻を表示または記録するときに使用するフォーマットを選択します。

**通貨:** 新規の各プロジェクトで使用するデフォルトの通貨を選択します。これは、異なる通貨で記録されている既存のデータは上書きしないことに注意してください。



**小数位。**数字上に表示する少数位の数を選択してください。

**プレビューに表示する行の数:** たとえば、インポートするデータ ソースをチェックするときに、プレビューに表示するレコードの最大数を選択します。

## モデルの設定

**デフォルトで作成したタイプをモデル化。** 特定の出力（予測ルールモデル）に関連した高いまたは低い類似性を表示するセグメントを特定するモデルを使用したい場合、レコードを一定数のクラスタに分類するモデル（クラスタモデル）を作成したい場合、または一緒に表示される項目を予測し、その間の関係性の強さを予測するモデル（アソシエーション モデル）を使用したい場合は、デフォルトで、アプリケーションに利用可能なデータ（予測モデル）から自動的に最適なモデルを作成させるか指定します。アプリケーションに表示するよう設定したモデルのタイプのみがこのドロップダウン メニューで利用できます。

**信頼できるモデル作成のために自動的にデータをクリーンアップして準備します** このオプションを選択すると、アプリケーションに各種のデータの準備を行わせることができます。たとえば、問題あるデータの修復、少数の有用なデータへのスクリーニング、あるいは新しい属性の導出などです。

クリーンアップ オペレーションでは、速度と正確さを均等にバランスさせるか、速度より正確さを優先させるか、あるいは、大量のデータセットを操作しているとき、またはクイック結果を検索しているときのように、速度を優先させるかを選択することができます。

**単一モデルの構築に費やされる時間を制限。** このオプションを使用して、各モデルの作成で費やされるデフォルトの時間を指定します。これは、大規模なデータセットを処理する場合に特に役に立ち、モデル作成を完了させるために必要な時間に対して大きな影響を持ちます。

## テストおよびシミュレーションの設定

**テストレコードのデフォルト数:** テストの実行時に対象とするレコードのデフォルト数を指定します。

# プロジェクト、モデル、および、ルールの開封及び保存

作業の開封及び保存がIBM® SPSS® Collaboration and Deployment Services Repositoryで可能です、または、ローカルのファイルシステムにダウンロードすることもできます。ローカルのファイルシステムからIBM® SPSS® Modelerストリームファイルをアップロード、及びIBM SPSS Collaboration

and Deployment Services Repositoryへの保存が可能です。IBM SPSS Collaboration and Deployment Services Repository への保存は、複数バージョンのサポート、ラベル付け、セキュリティなどの重要な利点があります。ローカルの設定によっては、すべてのオプションがサポートされない場合があります。



新規モデルの作成



現在のプロジェクト、モデル、またはルールを、IBM SPSS Collaboration and Deployment Services Repository 内の最後に確認された場所に保存します。オプションで、保存時にラベルを指定することもできます。



現在のプロジェクト、モデル、またはルールを保存およびラベル付けするか、または新しい名前でも保存します。状況に応じて、IBM SPSS Collaboration and Deployment Services Repositoryへの保存、またはローカル ファイル システムへのダウンロードを選択できます。



プロジェクト、モデル、またはルールをIBM SPSS Collaboration and Deployment Services Repositoryから もしくは、適宜、ローカル ファイル システム、あるいは両方から開きます。



プロジェクト、モデル、またはルールをローカルファイルシステムからアップロードします。ファイル システムへのアクセスは、すべてのアプリケーションでサポートされているわけではありません。SPSS Modelerストリームをアップロードする場合は、ファイルを開く、或いはレポジトリへの保存が選択できます。

## IBM SPSS Collaboration and Deployment Services Repositoryへの保存

IBM® SPSS® Collaboration and Deployment Services Repository への保存は、複数バージョンのサポート、ラベル付け、セキュリティなどの重要な利点があります。保存するたびに新しいバージョンが作成されるので、任意の時点で以前のバージョンに戻すことができます。ラベルを使ってバージョンを管理し、セキュリティ設定を適用して、オブジェクトに対してアクセス、表示、または削除できるユーザーを決定することができます。

左側のペイン内のツリー表示には、閲覧権限が与えられている IBM SPSS Collaboration and Deployment Services Repository 内のフォルダが表示されます。フォルダ ツリーをナビゲートするか、検索を実行して探しているオブジェクトを見つけます。ファイル リストを、左下のフィルタ ペイン内に用意されたラベルを使用してフィルタリングすることもできます。たとえば、本番システム ラベルをクリックすると、そのラベルが付いたファイルのみを表示することができます。

使用可能な場合に、ファイルに関する詳細を表示するには、プレビューアイコンまたはその隣にある情報アイコンをクリックします。

**フォルダ:** 利用可能なフォルダを表示し、階層をナビゲートできます。保存先フォルダの選択。また、フォルダ リストの上にあるアイコンを使用して、フォルダを新規作成したりきゾックのフォルダの名称を変更することができます。

**ラベルによるフィルタリング:** 表示を、選択したラベルを持つオブジェクトのみに制限します。フォルダを参照するとき、このラベルが付いたオブジェクトのみが表示されます。

内のアプリケーション保存ルールで使用されるルールをエクスポートする際は、名前を付けて保存 ドロップダウン リストも保存できます。ドロップダウン リストから 単一のルールセット または 個々のルールセット を選択し、ルールの保存方法を指定します。参照したルールしか使用できない場合、ドロップダウン リストは無効になり、ルールは単一のルールセットとして保存されます（参照されたルールは個別に保存できません）。

個別のルールとして保存することを選択した場合、保存したい個別のルールに対して詳細を入力するための表が表示されます。

**名前** 保存するオブジェクトの一意の名前を入力します。

**キーワード:** 保存したオブジェクトを検索するときに使用する任意のキーワードを指定します。たとえば、「Sales」や「1-2009」です。

**説明:** 保存したオブジェクトの、キーワードよりも長い説明を指定します。たとえば、「2009 年第 1 四半期の小売売上高の分析」です。

**ラベル** 保存したオブジェクトに複数のバージョンがある場合は、一意のラベルでバージョンを識別します。既存のラベルを選択するか、または新しいラベルを作成します。

## ラベルの操作

ラベルは特定のプロジェクト、モデル、またはルールを識別するために使用されます。たとえばルールを複数のプロジェクト間で共有する場合、使用するバージョンラベル付き生成を指定します。ルールの新しいバージョンが生成された場合、プロジェクト内で使用したルールは生成ラベルが明示的にそのバージョンへ移動しない限り変化しません。

共有オブジェクトを参照する場合、ラベル付きバージョンの使用を強く推奨します。特に開発環境においては、変更した最新バージョンは意図しない結果を招くことがあるためです。場合によっては、ルール内からルール

を参照した場合などはラベル付きバージョンを使用する必要があります。詳細は、5 章 p. 51 [ルールのエクスポートと再利用](#) を参照してください。

バージョンが指定されていない場合は最新のバージョンを使用し、その結果、ルールに対する後続の変更は自動的におこなわれます。(最新とは新たなバージョン作成されるたびに自動的に公開される特殊なラベルと考えます。)

毎回、新しいオブジェクトまたはバージョンを保存する際、そのバージョンに適用するラベルを選択できます。詳細は、[p. 25 プロジェクト、モデル、および、ルールの開封及び保存](#) を参照してください。

**ラベル名の制限について。** どのような場合に IBM® Analytical Decision Management、文字で使用するラベルを定義するか? # & および : (コロン) をラベル名に使用することはできません。これらの文字が含まれていると、これらのラベルを使用するオブジェクトへの参照をアプリケーションが正しく解読できなくなる可能性があります。例えば、次のラベル名は無効になります。

テスト & 展開

DEPLOY:PHASE1

## プロジェクトのロック

編集のためにプロジェクト、モデル、またはルールを開いた場合、これらのオブジェクトは IBM® SPSS® Collaboration and Deployment Services Repository でロックされ、他のユーザーが編集できないようになります。これは、作業領域の右上のチェックボックスで表示されます。このチェックボックスの選択を解除すると、他のユーザーに当該オブジェクトの編集が許可されます。

- プロジェクトをロックすると、他のユーザーは開いて表示することはできますが、変更は保存できないようになります。
- ロックしたプロジェクトを閉じる (たとえば、ブラウザを閉じる) と、ロックは解除され、他のユーザーが当該オブジェクトを編集できるようになります。
- プロジェクトを閉じずにアンロック (ロック チェックボックスの選択解除) を選択した場合は、それまでの変更を保存するかが尋ねられます。保存していない変更は、オブジェクトのロックを解除すれば失われます。
- プロジェクトをアンロックした後でも作業領域の参照はできますが、変更は保存できません。

## オブジェクトのプロパティ

ルールセット、モデル、またはプロジェクトの編集時には、オブジェクトを後で参照する場合の識別を容易にするために、キーワードや説明を追加することができます。[プロパティの編集] ダイアログを表示するには、アプリケーション内のプロパティ アイコンをクリックします。

図 3-2  
プロパティ アイコン



**キーワード:** 保存したオブジェクトを検索するときに使用する任意のキーワードを指定します。たとえば、「Sales」や「Q1-2012」です。

**説明:** 保存したオブジェクトの、キーワードよりも長い説明を指定します。例えば、「2012 年第 1 四半期の小売売上高の分析」です。

# データソースの管理

[データ] タブを使用すると、分析、シミュレーションおよびテスト、スコアリング、その他の目的でデータソースを定義することができます。データセットは、管理者が設定するか、またはユーザーが独自に追加します。

タブは 3 つのメイン領域で構成されます。

- **プロジェクト データ モデル:**これは、アプリケーションが必要とするフィールドを定義します。その他のデータセットはすべて、このソースに相対的にマップされます。データソースフィールドセクションはプロジェクトで使用される入力フィールドとタイプをリストします。追加フィールドセクションにより、必要に応じて別のデータソースから式やモデル出力を追加してプロジェクト データ モデルを拡張することができます。
- **プロジェクト データ ソース:**現在のプロジェクト、モデル、またはルールで使用するために保存されているデータソースを一覧表示します。
- **マイ データ ソース:**自分で定義したデータソースが一覧表示されます。この一覧のデータセットは異なるプロジェクト間でコピーできます。これにより、複数のアプリケーション間でデータを共有することが可能になります。

## プロジェクト データ モデル

プロジェクト データ モデルは、プロジェクトで使用される入力フィールドおよびタイプを備えたテンプレートのような役割を持っています。

**データソース:**アプリケーション用に選択されたデータソースを表示します。

**キー:**アプリケーション用に選択されたデータソースがアプリケーションビュー データソースで、アプリケーションビューキーがデータソーステーブルに定義されている場合は、このドロップダウンから、選択元となるキーのリストが使用できます。アプリケーションビューキーはIBM® SPSS® Collaboration and Deployment Services Deployment Managerで作成されるIBM® SPSS® Collaboration and Deployment Services Enterprise Viewの一部となります。アプリケーションビューテーブル用に定義されたキーがない場合、もしくは、アプリケーション用に選択されたデータソースがアプリケーションビュー データソースでない場合は、代わりにキードロップダウンの中からデータソースのすべての操作データフィールドのリストを使用して選択できます。

選択されたキー フィールドは、追加フィールドセクションで定義されたすべての追加データ ソース テーブルのマッピング用に使用されます。実行時には、キーは1次プロジェクト データ モデル テーブルと2次テーブルを結合するために使用されます。プロジェクト データ モデルのキー フィールドは、一意であることが前提です（たとえば、顧客ID、クレームID、など）各2次テーブルには一致するコラムがあることが必要です。

キー フィールドは、データ ソース用に定義された追加テーブルがない場合には、任意です。

**運用中:** このボックスにチェックを付けると、フィールド内のデータがルールと式、および予測で使用可能なことを意味します。このフィールドを分析のターゲット フィールドとして使用する場合は、チェックボックスからチェックを外します。たとえば、ダイレクト メールによるキャンペーンに対する顧客の応答を調べるクエリーを実行する場合は、応答値を前もって知ることはできないので、結果フィールドは分析的になります。

注：IBM SPSS Collaboration and Deployment Services Enterprise View データ ソースをプロジェクト データ モデルとして選択した場合は、この列のボックスは無効になります。

**フィールド名** データ ソースにリストされている各フィールドの名前を、ストレージ タイプを識別するアイコンと共に表示します。

**設定:** データ ソース内の各フィールドの測定タイプを表示します。測定タイプを変更した場合は、データ スキャンを更新して、データ互換性を保証する必要があります。詳細は、[p. 34 測定レベル](#) を参照してください。

**値:** データ ソース内の各フィールドの値をリストします。たとえば、特定の範囲の最小値および最大値をリストします。フラグ フィールドで、マウス ポインタを値の上にかざすと、そのフィールドで定義されている「真」または「偽」の値を示すツール ヒント が表示されます。

**追加フィールドの追加/編集** このセクションは、必要に応じて追加フィールドとテーブルの追加/編集追加フィールドの追加/編集をクリックすることにより、別のデータ ソースから式やモデル出力を追加してプロジェクト データ モデルを拡張することができます。

## プロジェクト データ ソース

プロジェクト データ ソースには、管理者によって定義済みのものや、ユーザーによって追加されたものを含むことができます。オプションで、管理者は 1 つ以上のデータ ソースをロックして、ユーザーがそれらを修正または削除するのを防止したり、またはすべてのデータ オプション

をロックしてユーザーが新しいデータ ソースを作成できないようにすることができます。

データ ソースの入力がプロジェクト データ モデルのものと直接一致しない場合は、前者を後者にマップして矛盾を修正できます。たとえば、プロジェクト データ モデルが、はい および いいえ（測定タイプはフラグ型）の値を持つ 購入 という名前のフィールドを必要とする場合、使用するデータ ソースはそのようにマッピングが可能な、同等のフィールドを備えている必要があります。

**名前** データ ソース名を表示し、ファイル タイプを識別するアイコンを示します。

**レコード数:**この列のアイコンをクリックして、テーブル上のどのデータ ソースのレコード数も表示させることができます。数はアイコンのとなりに表示されます。

**プレビュー:** これを使用すると、ソースに含まれているデータのサンプルをプレビュー表示することができます。 [詳細は、 p.34 データのプレビュー を参照してください。](#)

**互換:** データ ソースがプロジェクト データ モデルとして使用されているというメモを表示するか、またはデータ ソースとプロジェクト データ モデルの互換性の度合を示す、緑、オレンジ、赤のボールを表示します。

- 緑のボールは、データ ソースがプロジェクト データ モデルのデータ ソースと、操作上は互換であることを意味します。操作上は互換のデータ ソースとは、プロジェクト データ モデルのすべての操作フィールドを含み、さらに追加フィールドを持つデータ ソースのことです。このデータ ソースは、ルール、スコアリング、シミュレーション、およびテストの操作に適しています。
- オレンジのボールは、データ ソースが、プロジェクト データ モデルのデータ ソースと同じ名前と型を持つ互換フィールドを少なくとも 1 つ持っていることを意味します。このデータ ソースは追加フィールドを持つ場合があり、モデルの作成および評価に適しています。
- 赤のボールは、データ ソースがプロジェクト データ モデルとは非互換であり、フィールドはアプリケーションで使用する前にマッピングが必要であることを意味します。非互換のデータ ソースとは、対応するプロジェクト データ モデルの型とは非互換の型を持つフィールドが少なくとも 1 つあるデータ ソースです。

**フィールドをマッピング:** このオプションを使用すると、データ ソース フィールドをプロジェクト データ モデルのデータ ソース フィールドと比較し、互換フィールドをマッピングまたはマッピング解除することにより、プロジェクト データ モデルが必要とするデータ ソース フィールド



ドにマッチさせることができます。 [詳細は、 p. 46 フィールドの関連付け](#) を参照してください。

**コピー:** これを使用すると、データ ソースを **マイデータソース** 領域にコピーすることができます。

### マイデータソース

この一覧のデータ ソースはユーザー アカウントとともに保存されるため、そのユーザーが再度ログインしたときにはすぐに利用可能となり、またそのユーザーが開いたあらゆるプロジェクト、モデル、またはルールにコピーすることができます（ただしそれを行う権限が管理者によってそのユーザー付与されていることが前提です）。タブ内でこの場所に存在するフィールドは、**プロジェクトデータソース** 領域のフィールドと同じ機能を持ちます。ただし、**互換** 列はありません。

**コピー** 列を使用すると、データ ソースを **プロジェクトデータソース** 領域にコピーすることができます。

### データソースの操作

- ▶ 新規データ ソースを **プロジェクトデータソース** または **マイデータソース** のいずれかのリストに追加するには、**データソースを追加** を選択します。 [詳細は、 p. 35 新規データ ソースの作成](#) を参照してください。
- ▶ **プロジェクトデータモデル** のデータ ソース内のフィールドの測定レベルまたは型を変更するには、関連するレベルを選択します。 [詳細は、 p. 34 測定レベル](#) を参照してください。
- ▶ **プロジェクトデータソース** リスト間でデータ ソースをコピーするには、[コピー] 列内で適切な矢印をクリックします。
- ▶ データ ソース用のフィールド名を**プロジェクトデータモデル**にマッピングするには、[互換] 列の下の適切なリンクをクリックします。（一度フィールド名がマッピングされると、リンクの表示は消えます。） [詳細は、 p. 46 フィールドの関連付け](#) を参照してください。
- ▶ データ ソースをプレビューするには、**プレビュー** アイコンをクリックします。 [詳細は、 p. 34 データのプレビュー](#) を参照してください。
- ▶ **追加フィールドを追加** をクリックすることで、追加フィールド（式、モデル出力を異なるデータソースから）を追加してデータモデルを拡張することができます。 [詳細は、 p. 43 追加フィールドの追加](#) を参照してください。

## データのプレビュー

正しいデータ ソースを選択したことを確認するには、プレビュー をクリックし、ソースに含まれているデータのサンプルを表示します。このプレビューには、データ タブの追加フィールドセクションにリストされるすべての追加フィールドが含まれます。

## 測定レベル

測定レベルは、各データ フィールドで表される情報の特徴を強調するのに役立ちます。また特定のフィールドがルール、モデル作成、または他のアプリケーションでどのように使用されているかを判定するのに役立ちます。測定レベルは、データ ソースのプロジェクト データ モデル内で指定できます。たとえば、0 か 1 の値を持つ整数フィールドの測定レベルにフラグ型 を設定して、0 = 偽 で 1 = 真であることを示すようにできます。また、使用する入力フィールドの選択時に、データソースエディタ ダイアログでレベルを変更することもできます。 [詳細は、 p. 45 入力フィールドの選択 を参照してください。](#)

次の測定レベルを使用できます。

- **連続:**0-100 や 0.75-1.25 のように、数値の範囲を記述するために用いられます。連続値は、整数、実数、または日付/時刻になります。
- **カテゴリ:** 実際の DISTINCT 値の数字が不明な場合に、文字列値に用いられます。これはインスタンス化されていないデータ型で、すべてのストレージに関する情報やデータの使用法に関する情報が不明であることを示します。データが読取られると、測定レベルはフラグ、名義型、またはデータ型不明
- **フラグ型:**true and false, Yes and No or 0 and 1のように、ある特徴の存在または不存在を示す2つの異なる値をもつデータに使用されます。IBM® Analytical Decision Managementにおいて、カテゴリ リストの最初の値は「偽」の値として使用され、2番目の値は「真」の値として使用されます。これらは、[データ] タブに表示されたときに、マウス ポインタをかざしたときのツール ヒントとして示されます。データは、テキスト、整数、実数、日付、時刻、またはタイムスタンプとして表わすことができます。
- **名義:**複数の DISTINCT 値を持つデータを記述するために用いられ、それぞれがsmall/medium/largeのようなセットのメンバーとして扱われます。名義データには、任意のストレージ (数値、文字列、または日付/時刻) を利用できます。測定レベルに 名義 を設定しても、値が自動的に文字列ストレージに変わるわけではないことに注意してください。
- **序数:** 固有の順位を持つ複数の DISTINCT 値を持つデータを記述するために用いられます。たとえば、給与区分や満足度ランキングには、序数データ型を割り当てることができます。序数は、そのデータ要素の自

然な並べ替え順で定義されます。たとえば、1,3,5は整数セットのデフォルトの並べ替え順で、HIGH, LOW, NORMAL（アルファベット昇順）が文字列セットの序数です。序数測定レベルを使用すると、可視化、モデル作成、および序数データを DISTINCT 型として認識する他のアプリケーション（IBM® SPSS® Statistics など）へエクスポートするため、一連のカテゴリ データを順位データとして定義できます。名目フィールドを使用できる場所ならどこでも、序数フィールドを使用できます。また、任意のストレージ型（実数、整数、文字列、日付、時刻など）のフィールドは序数型として定義できます。

- **型なし:** 上記のいずれのデータ型にもあてはまらないデータ、単一値を持つフィールド、またはセットが定義された最大値よりも多くのメンバーを含む名目データに用いられます。これは、測定レベルが多くのメンバー（アカウント番号など）を持つセットのような場合に役に立ちます。

## 新規データソースの作成

- ▶ データ ソース エディタにアクセスするには、[プロジェクト データ ソース] 領域または [マイ データ ソース] 領域の下で **データソースを追加** を選択します。
- ▶ データ ソースの名前を入力します。
- ▶ ソース タイプを選択します。
  - **ファイル:** サポートされているファイルタイプは多数あり、例えばテキスト (\*.txt、\*.csv、など)、Excel、IBM® SPSS® Statistics で使用される \*.sav フォーマット、および SAS ファイルがあります。 [詳細は、p. 36 ファイル ソース を参照してください。](#)
  - **データベース:** Microsoft SQL Server、DB2、Oracle など ODBC をサポートする任意のデータベースです。 [詳細は、p. 39 データベース ソース を参照してください。](#)
  - **エンタープライズ ビュー** IBM® SPSS® Collaboration and Deployment Services Enterprise View への接続を、IBM® SPSS® Collaboration and Deployment Services での定義に従って作成します。 [詳細は、p. 40 エンタープライズ ビュー ソース を参照してください。](#)
  - **IBM Cognos BI サーバー。** IBM Cognos BI サーバーをサポートするデータベース。 [詳細は、p. 41 IBM Cognos BI ソース を参照してください。](#)
- ▶ ソース タイプに適切なオプションを指定します。

レコード数アイコンをクリックして現行のデータ ソース内のレコード数を表示させることができます。もしくはプレビュー アイコンをクリックしてデータ ソースから最初の 10 個のレコードを表示させることができます。

## ファイル ソース

アプリケーションは、いくつかのデータ ファイル タイプをサポートします。ファイル名を入力するか、またはファイルを参照すると、アプリケーションによって自動的に、タイプの検出、データ ソース エディタのオープン、およびそのタイプに関連する追加フィールドの表示が行われます。ファイル タイプが不正な場合、または特定のバージョンが必要な場合（たとえば、Microsoft Excel 1997-2003 など）は、必要なタイプに変更することができます。

参照をクリックしてリポジトリからファイルを選択することができます。もしくは、**ローカル ファイルのアップロード**をクリックして使用するローカル ファイル システムからファイルを選択することができます（有効にしている場合）。

ローカル ファイルをアップロードする場合は、IBM® SPSS® Modeler Server にアップロードされます。アップロード ファイル ダイアログの中で、アップロードするローカル ファイルを選択し、ファイルをアップロードするSPSS Modeler Server宛先を参照し選択します。宛先を選択する際には、適宜ファイル名を変更することができます。

ローカル ファイルをアップロード機能はデフォルトでは無効になっています。管理者は、ブラウザベースの IBM® SPSS® Collaboration and Deployment Services Deployment Manager を使用して、この機能の有効化または無効化、および許容されている最大ファイル サイズ制限の増減を行うことができます。

### テキスト ベースのデータ ファイル

テキスト ベースのデータ ソースを選択すると、さらに詳細を入力するように求められます。

**最初の行は列名:** 各列の名前がデータ ソースの見出し行として含まれている場合は、これを選択します。

**文字コード:** 使用するテキストのエンコード方法を指定します。システムデフォルトまたは UTF-8 から選択できます。

**小数点記号:** データ中で小数点記号をどのように表すかを指定します。

- **ピリオド (.):** 小数点区切り文字として、ピリオドを使用します。
- **カンマ (,):** 小数点区切り文字として、カンマを使用します。

**区切り文字:** このコントロール用に表示されたチェック ボックスを使用して、カンマ (,) などの、ファイル内のフィールドの境界を定義する文字を指定できます。複数の区切り文字を使用するレコードの場合、「, |」

のように複数の区切り文字を指定することもできます。デフォルトの区切り文字はカンマです。

注：カンマが、小数点記号としても定義されている場合、ここでのデフォルト設定は使用されません。この場合、カンマはフィールドの区切り文字と桁区切り記号の両方であるため、区切り文字リストから[その他]を選択します。次に、手動で入力フィールドにカンマを指定します。

隣接する複数の空白文字を単一の区切り文字として認識する場合は、[複数の空白区切り文字を許可]を選択します。たとえば、あるデータ値の後に 4 つのスペースが続き、その後に別のデータ値が続いている場合は、5 つのフィールドではなく、2 つのフィールドとして扱われます。

### 詳細オプション

**EOL コメント文字:** データ内で注釈を示す文字 (# や ! など) を指定します。データ ファイル内でこれらの文字がある場所から次の改行文字のある場所までは、すべて注釈になります。ただし、その改行文字は注釈に含まれずに無視されます。

**入力フィールドの指定:** 各レコードで使用される入力フィールドの数を指定します。

**ヘッダー文字をスキップ:** 最初のレコードの先頭で無視する文字数を指定します。

**スキャン行:** 指定したデータ型をスキャンする行数を指定します。

**前後のスペースを削除** インポート時に文字列の前後のスペースを破棄する場合は選択します。右側、左側、両側からスペースを破棄したり、破棄なしも選べます。

**不正な文字:** データ ソースから不正な文字を除去する場合、[破棄]を選択します。不正な文字を指定した記号 (1 文字だけ) で置換する場合は、[置換値]を選択します。ヌル文字または指定したエンコード方式内に存在していない任意の文字が不正な文字になります。

**引用符:** ドロップダウン リストを使用すると、インポート時の一重引用符および二重引用符の取り扱い方法を指定することができます。すべての引用符の **破棄**、引用符をフィールド値に **テキストとして包含**、またはペアとなっている引用符を探して、それを破棄するために **ペアで破棄** を選択することができます。ペアとなる引用符がない場合は、エラー メッセージが表示されます。**破棄** と **ペアで破棄** では、フィールド値を文字列として (引用符なしで) 保存します。

**入力フィールドの指定。** 詳細は、 [p. 45 入力フィールドの選択](#) を参照してください。

## Excel データ ファイル

Excel データ ソースを選択すると、さらに詳細を入力するように求められます。

**最初の行は列名:** 各列の名前がデータ ソースの見出し行として含まれている場合は、これを選択します。

**名前付き範囲を使用:** これを使用すると、Excel ワークシート内で定義された、名前の付いたセルの範囲を選択することができます。名前付き範囲を使用すると、その他のワークシートおよびデータ範囲の設定が以後適用不可能になり、結果として無効になります。

**ワークシートの選択:** インポートするワークシートを名前またはインデックスのいずれかで指定します。

- **名前で指定:** インポートするワークシートの名前を選択します。
- **インデックスで指定:** インポートするワークシートのインデックス値を、最初のワークシートの 0 から始まり、2 番目のワークシートは 1 というように指定します。

**ワークシート上の範囲:** 最初の空白でない行から始まるデータ、または明示したセルの範囲のデータをインポートすることができます。

- **最初の非空白行から範囲開始:** 最初の空白でないセルの位置を探して、これをデータ範囲の左上隅として使用します。
- **セル範囲の明示:** 行と列を使用して、範囲を明示します。たとえば、Excel の範囲 A1:D5を指定するには、最初のフィールドに A1、2 番目のフィールドに D5と入力します（または、R1C1 と R5C4を入力します）。指定した範囲内のすべての行が、空白行も含めて、返されます。

**空白行を返す:** 空白行があると、行をスキップして無視するように選択するか、または **空白行を返す** を選択し、その空白行も含めてワークシートの最後まですべてのデータの読み込みを続行させることができます。

**入力フィールドの指定。** 詳細は、 [p.45 入力フィールドの選択](#) を参照してください。

## IBM SPSS Statistics データ ファイル

IBM® SPSS® Statistics データ ソースを選択すると、さらに詳細を入力するように求められます。

**変数名:** SPSS Statistics の .sav ファイルからのインポート時の、変数名とラベルの処理方法を選択します。

- **名前とラベルを読み込み:** 変数名とラベルの両方を読み込む場合に選択します。これはデフォルトのオプションです。ラベルは、グラフ、モデルブラウザ、その他のタイプの出力中に表示されることがあります。
- **ラベルを名前として読み込み:** 短いフィールド名ではなく、SPSS Statistics の .sav ファイルの詳細な変数ラベルを読み込み、このラベルを変数名として使用する場合に選択します。

**値:**SPSS Statistics の .sav ファイルからのインポート時の、変数とラベルの処理方法を選択します。

- **データとラベルを読み込み:** 実際の値と変数ラベルの両方を読み込む場合に選択します。これはデフォルトのオプションです。
- **ラベルをデータとして読み込み:** 値を表現するために使用される数値コードまたはシンボル値コードの代わりに、.sav ファイルの値ラベルを使用する場合に選択します。たとえば、男性および女性をそれぞれ 1 および 2 の値で表す性別フィールドを持つデータに対してこのオプションを選択すると、フィールドが文字列に変換され、男性および女性が実際の値としてインポートされます。

このオプションを選択する場合は、SPSS Statistics データの欠けている値について考慮する必要があります。たとえば、数値フィールドで欠損値 (0 = 回答なし, -99 = 不明) にだけラベルを使用している場合、このオプションを選択すると値ラベル 回答なしと不明だけがインポートされ、フィールドが文字列に変換されます。このような場合は、値を直接インポートする必要があります。

**入力フィールドの指定。** 詳細は、 [p.45 入力フィールドの選択](#) を参照してください。

## データベース ソース

データベース ソースを選択すると、さらに詳細を入力するように求められます。

**テスト接続:** データベースのユーザー名とパスワードを入力し、クリックしてデータベースへのアクセス権を持っているか確認します。

**テーブル名:** 使用するデータを含んでいるテーブルを選択します。

**前後のスペースを削除** インポート時に文字列の前後のスペースを破棄する場合に選択します。右側、左側、両側からスペースを破棄したり、破棄なしも選べます。

**テーブルおよび列名を引用符で囲む:** クエリーをデータベースに送信するときにテーブル名と列名を引用符で囲むかどうかを指定します (たとえば、テーブル名と列名にスペースや句読点が含まれているような場合)。

- [必要に応じて] オプションを選択すると、非標準文字が含まれている場合のみ、テーブル名とフィールド名が引用符で囲まれます。非標準文字とは、非 ASCII 文字、スペース文字、およびピリオド (.) 以外の非英数文字を指します。
- テーブル名とフィールド名を引用符で囲まない場合は、[しない] を選択します。
- すべてのテーブル名とフィールド名を引用符で囲む場合は、[常時] を選択します。

**入力フィールドの指定。** 詳細は、 [p. 45 入力フィールドの選択](#) を参照してください。

## エンタプライズ ビュー ソース

エンタプライズ ビュー ソースを選択すると、さらに詳細を入力するように求められます。

**アプリケーション ビュー:** ドロップダウンリストに、選択したアプリケーション ビューが表示されます (存在する場合)。現在のセッションで他のアプリケーション ビューに対して接続を行った場合でも、ドロップダウン リストにはこれらのアプリケーション ビューが表示されます。参照をクリックして、リポジトリ内の他のアプリケーション ビューを検索します。

**テーブル:** 使用するデータを含んでいるアプリケーション ビュー テーブルを選択します。

**データ プロバイダ:** ドロップダウン リストは選択したアプリケーション ビューを参照するデータ プロバイダ定義の名前を表示します。データ プロバイダ定義を選択するか、[データソースなし] オプションを選択します。[データソースなし] オプションを選択すると、データ ソースはアプリケーション上で実質的に使用不可能となり、テスト、シミュレーション、WhatIF? 分析などでは使用できなくなります。

**環境:** IBM® SPSS® Collaboration and Deployment Services で設定される環境値は、使用可能なデータ プロバイダ定義を決定します。

**入力フィールドの指定。** 詳細は、 [p. 45 入力フィールドの選択](#) を参照してください。

IBM® SPSS® Collaboration and Deployment Services Enterprise View についての詳細、およびこのデータの設定方法についての詳細は、IBM SPSS Collaboration and Deployment Services のドキュメントを参照してください。



## IBM Cognos BI ソース

IBM Cognos BI データベース データまたは単一のリスト レポートを IBM® Analytical Decision Management で使用するためにインポートできます。

注：IBM® SPSS® Modeler ServerをIBM Analytical Decision Managementと一緒に使用した場合は IBM Cognos BI サーバーをサポートしませんので、このオプションは使用できません。

IBM Cognos BI サーバーを選択して[接続] をクリックすると、サーバー接続の詳細の入力を求めるメッセージが表示されます。

- ▶ データをインポートまたはエクスポートする IBM Cognos サーバーのサーバー URL をキーインします。使用する URL が不確かな場合は Cognos システム管理者に問い合わせてください。
- ▶ 接続するモードを選択してください。特定のユーザー（たとえば、管理者）としてログインするには、資格証明の設定を選択して、IBM Cognos の名前空間、ユーザー名、および、パスワードを入力してください。
  - サーバーへのログインに使用する、IBM Cognos セキュリティ認証プロバイダ名前空間を指定します。認証プロバイダは、ユーザー、グループ、および、役割りの定義と保守や、認証プロセスの管理に使用されます。
  - サーバーログオンに使用する、IBM Cognos ユーザー名を入力します。
  - 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。
- ▶ または、ユーザー資格証明なしでログインする場合は、名前空間、ユーザー名、および、パスワード フィールドに入力することはできませんが、匿名接続の使用を選択します。

注：一部のサーバー接続には匿名接続は使用できません。

## Cognos オブジェクト詳細の選択

IBM Cognos BI サーバーへの接続設定後、モードを選択します。データ モードにより、データをインポートする元のパッケージを参照および選択できます。リストレポート モードにより、データをインポートする元の定義済みリスト レポートを参照および選択できます。

ひとつのパッケージには IBM Cognos モデルとすべてのフォルダ、レポート、ビュー、ショートカット、URL、および、そのモデルに関連するジョブの定義が含まれています。ひとつの IBM Cognos モデルはビジネス ルール、データ説明、データの相関関係、ビジネス ディメンションと階層、それに、その他の管理作業を定義します。

- ▶ パッケージを選択するには、**データ** モードを選択し、[参照] をクリックします。内容をインポートできるすべての使用可能なパッケージを一覧表示するダイアログ ボックスが表示されます。
- ▶ あるいは、単一のリスト レポートを選択するには、**リストレポート** モードを選択し、[参照] をクリックします。レポートを選択できるパッケージおよびフォルダを一覧表示するダイアログ ボックスが表示されます。注：インポートできるリスト レポートは 1 つだけです。複数のリストはサポートされていません。
- ▶ 必要な場合、リストを展開して、必要なパッケージまたはレポートを検索します。項目を選択したら、**OK** をクリックします。パッケージまたはレポート詳細は**データ ソース エディタ**に表示されます。
- ▶ パッケージ選択後、インポートするオブジェクトを選択します。たとえば、パッケージから使用するパッケージ オブジェクトのクエリ主題（データベースのテーブルを表します）または 個々の問い合わせ項目（テーブルの列を表します）を選択します。注：データは UTF-8 フォーマットでなければなりません。
- ▶ そのままのデータではなく、集計したデータをインポートしたい場合は、[インポート前にデータを集計] オプションを選択します。集計したデータは一般に、そのままのデータよりもレコード数が多くなります。

**オブジェクトを選択します。** パッケージまたはリスト レポートから選択できるオブジェクト（名前空間とも言います）を表示します。1つまたは複数のオブジェクトを展開してインポート可能な入力フィールドを表示させます。オブジェクトを選択すると、自動的にその入力フィールドがすべて選択されます。

**パラメータのレビュー** パラメータが定義済みの場合、データ オブジェクトまたはレポートのいずれかに対して、そのオブジェクトまたはレポートのインポート前にこれらのパラメータに値を指定できます。このテーブルには、フィルタや計算に必要なパラメータのリストが含まれます。これらはそれぞれ異なる値を持っている場合があります。1 つ以上のパラメータにデフォルト値がない場合、警告が表示されます。[値] 列で必要に応じて、パラメータを設定します。パラメータを使用して、データをインポートする前に調整を行うことができます（たとえば、パラメータ化計算実行のためなどに）。

**フィルタのレビュー** パッケージにフィルタを定義している場合、1 つ以上のフィルタをインポートできます。インポートしたフィルタがインポートしたデータに関連している場合、そのフィルタは、データのインポート前に適用されます。フィルタはパラメータ化される場合とされない場合があります。

**入力フィールドのレビュー**—IBM® Analytical Decision Managementにインポートされて処理されるすべての入力フィールド（データベース オブジェクトとも呼ばれます）がリストされます。必要がないフィールドがある場合は、フィールド名の隣のチェック ボックスのチェックを外します。

## Cognos オブジェクト アイコン

Cognos BI データベースからインポートできる各種オブジェクトは、次のテーブルが示すように、各種アイコンによって表現されます。

テーブル 4-1  
Cognos オブジェクト アイコン

アイコン	オブジェクト
	パッケージ
	名前空間
	クエリ主題
	クエリ アイテム
	測定ディメンジョン
	測定
	ディメンジョン
	レベル階層
	レベル
	フィルタ
	レポート
	スタンドアロン計算

## 追加フィールドの追加

ほとんどのIBM® Analytical Decision Managementアプリケーションで、プロジェクト データ モデルに式、モデル出力などの追加フィールドを追加することができます。プロジェクト データ モデルの下のデータ タブの追加フィールドセクションは、データの一部に含めることができるあらゆる式やモデル出力表示します。追加フィールドセクションは、IBM® SPSS® Rules Managementアプリケーションにはありませんのでご注意ください。

- ▶ このプロジェクトデータモデルセクションでは、画面の右側の追加フィールドの追加/編集。

### 式を追加するには

- ▶ 式の追加をクリックします。式エディタが表示されます。 [詳細は、 p. 45 式エディタ を参照してください。](#)

### モデル出力を追加するには

- ▶ モデル出力の追加をクリックしてモデルを参照し選択します。選択されたモデルの入力と出力ダイアログが表示されます。
- ▶ モデル入力フィールドが互換している場合は、ダイアログの**モデル入力マッピング**セクションは閉じられます。モデル入力フィールドが互換していない場合は、セクションが開かれ、選択されたモデルを入力するために使用可能なテーブル入力をマッピングする必要があります。
- ▶ **モデル出力**セクションに、選択されたモデルの使用可能なすべてのモデル出力がリストされます。使用データに追加するモデル出力フィールドを選択します。 **モデル出力**名が、モデルにより生成された出力の名前を表示します。

モデルが出力するどのフィールドも選択することができます。モデル出力フィールドはストリームから発生するすべてのフィールドですが、ストリームには入力されません（すべてのフィールドはストリームにより付加されます）。入力フィールドには、プロジェクト データ モデルのすべては含まれない場合があります。そのため、モデル出力フィールドをプロジェクト データ モデル フィールドと同じ名前にすることも可能です。したがって、ここでデータ フィールド名を変更することも可能です。

測定コラムはモデル出力の測定タイプを適切なタイプの接頭辞アイコン付きで表示します。値列はモデル出力の値を表示します。この列は一部のフィールドの結果を表示しますが、その他は表示しないのでご注意ください。通常は、測定タイプと値はデータ スキャンなしの既知の実際のモデル出力（予測データ、確率、など）です。したがって、既知の場合に表示されます。ただし、測定タイプまたは値は、ストリームの他の出力フィールド（たとえば、派生フィールド）にはデータスキャンなしでは未知の場合もあります。

- ▶ **モデル パラメータ** セクションは、選択されたストリームにスコアリング パラメータが含まれる場合には、ダイアログの一番下にあります。モデルパラメータを入力し、ダイアログを終了させる際には**保存**をクリックします。選択されたモデル出力はデータ タブの**追加フィールド**セクションに付加されます。

## 入力フィールドの選択

データが大量にある場合は、ソースからインポートされるデータ量を減らすか、または調整を行ってください。たとえば、関心のある領域に関連するフィールドのみを使用することが考えられます。

- ▶ データ ソース エディタのダイアログの最下部で、**入力フィールドの指定** をクリックします。ダイアログが拡張され、ソースに含まれるすべてのデータフィールドが表示されます。さらに、定義したすべての式も表示されます。 [詳細は、 p. 45 式エディタ を参照してください。](#)
- ▶ 使用する各フィールドおよび式を選択します。

データの選択に加えて、必要に応じて、1 つ以上のデータ フィールドの測定レベルを変更することもできます。 [詳細は、 p. 34 測定レベル を参照してください。](#)

## 式エディタ

式マネージャを使用すると、アプリケーションで使用する追加フィールドまたは属性を導出することができます。たとえば、銀行データを使用している場合は、顧客の収入と顧客が持つローン アカウント数の比率を計算する式を作成することが考えられます。式は、測定タイプが連続型の数値である必要があります。これは変更できません。

注：式は、同一アプリケーション内ではどの場所にあっても運用データとして使用することができます。たとえば、ルールとモデルの予測で使用可能です。

- ▶ 既存の式を表示するには、データタブで**追加フィールドの追加/編集**をクリックし、それから**式を追加**をクリックします。

### 式の作成

- ▶ 式エディタダイアログで、式の固有の名前を入力します。
- ▶ データソースで利用可能なインプットが全て表示されます。必要な入力値および数学記号をダブルクリックするか、**方程式** フィールドにドラッグして、式を作成します。
- ▶ **説明** に式の使用目的を入力します。
- ▶ **保存** をクリックします。

## フィールドの関連付け

新しいデータ ソースを追加するとき、すべてのフィールドをマッピングして、プロジェクト データ モデルとの互換性を保証する必要があります。たとえば、プロジェクト データ モデルが、はい および いいえ（測定タイプはフラグ型）の値を持つ 購入 という名前のフィールドを必要とする場合、使用するデータ ソースは同等のフィールドを備えている必要があります。フィールド名が同一でない場合は、必要に応じて、マッピングすることができます。入力およびマッピングした関連フィールドでは、同じデータ タイプが使用されている必要があることに注意してください。

逆に、プロジェクト データ モデルを変更した場合は、多くのデータ ソースが正常にマッピングできなくなります。その場合は、プロジェクト データ ソース 領域内の互換インジケータ ボールは、各データ ソースに該当する色に変化します。したがって、すべてのデータ ソースを新しいプロジェクト データ モデルにマッピングし直す必要があります。

- ▶ **プロジェクト データ ソース 領域内で、フィールドをマッピング** をクリックします。ダイアログが開き、システムが最適なマッピング フィールドを推測します。任意の時点で、**リセット** をクリックすることにより、システムの最適な推測に戻ることができます。
- ▶ **利用可能な入力 列のフィールドとそれに対応する 必須入力フィールド** を選択し、**マップ** をクリックして、2 つのフィールドをマッピングします。**利用可能な入力 列**のすべてのフィールドをマッピングするまで、この手順を繰り返します。
- ▶ 完了したら、**OK** をクリックします。

# ルールの操作

ルールは、ビジネス ロジックまたは予測モデルの出力に基づいて、判断を自動化するために使用されます。アプリケーションに応じて、支払い履歴に基づいて顧客を除外するため、リスクな請求について調査部門に問い合わせるため、または予測モデルによって特定された解約、購入、不正を試みる傾向の大きい顧客をターゲットとするためにルールを使用します。ルールには自動化機能があり、モデルには予測機能があります。そしてこれらの機能を結合する機能が、予測アプリケーションを他の分析ツールから差別化する主要な要因の 1 つとなっています。

ルールは、各アプリケーション内の異なる場所で、作成、適用、保存を行うことができます。

## ルールの作成

- ▶ アプリケーション内のルールを追加する場所で、[新規ルールを作成] アイコンをクリックします。
- ▶ あるいは、**既存ルールを検索** アイコンをクリックし、現在のアプリケーションで追加または再利用できる既存のルールを参照します。 [詳細は、p. 51 ルールのエクスポートと再利用](#) を参照してください。
- ▶ 共有ルールを作成するには、アプリケーション起動ページで IBM® SPSS® Rules Management を起動します。

## セグメント ルールの定義






セグメント ルールは、各種のアプリケーションで、レコードの選択と割り当てを行うために使用されます。セグメントは、真または偽の結果を返す 1 つ以上の式で定義されます。例えば、1 つのセグメントで年齢フィールドの値が 18 未満であるときに真を返すといったルールを持つことができます。セグメントの定義は、そのデータで定義されるフィールドまたは式に基づいて、あるいはモデルが返す傾向スコアや予測値に基づいて行うことができます。複数の次元階層（要素 として参照される）を含むよう設定された定義タブのあるアプリケーションでは、その他の次元メンバーからも選択できます。 [詳細は、3 章 p. 22 プロジェクト設定](#) を参照してください。

各セグメントは、1 つ以上の式で定義され、すべての条件が満たされれば真の値を返します。たとえば、年齢と性別で定義されたセグメントは、両方の条件に一致するレコードのみを含み、それ以外のすべてのレコードに対しては偽を返します。

### セグメント ルールの作成または編集方法

- ▶ ルールを追加したいアプリケーションで、**新規ルールの作成** アイコンをクリックします。
- ▶ ルールの名前を指定します。
- ▶ 必要に応じて、入力列で下矢印アイコンをクリックし、フィールド、ルール、またはモデルを指定します。

テーブル 5-1  
入力列オプション

アイコン	使用方法
	有効にすると、前に定義した要素から次元メンバーを選択できます。
	現在のデータ セットからフィールドを選択することができます。
	リポジトリから既存のモデルまたはルールを選択することができます。
	新しいモデルを作成することができます。
	式を作成できます。

- ▶ 使用する機能を選択します。使用可能オプションについては以下のテーブルをご参照ください。



テーブル 5-2  
使用可能オペレーター

値のデータ型:	使用可能オペレーター
タイムスタンプ、日付、日付/時刻	<ul style="list-style-type: none"> <li>■ = (等しい)</li> <li>■ != (等しくない)</li> <li>■ &gt; (より大きい)</li> <li>■ &lt; (より小さい)</li> <li>■ &gt;= (以上)</li> <li>■ &lt;= (以下)</li> <li>■ 間</li> <li>■ ISNIL</li> <li>■ IS NOT NIL</li> </ul>
範囲	<ul style="list-style-type: none"> <li>■ = (等しい)</li> <li>■ != (等しくない)</li> <li>■ &gt; (より大きい)</li> <li>■ &lt; (より小さい)</li> <li>■ &gt;= (以上)</li> <li>■ &lt;= (以下)</li> <li>■ 間</li> <li>■ ISNIL</li> <li>■ IS NOT NIL</li> <li>■ ONE OF</li> <li>■ NOT ONE OF</li> </ul>
セット	<ul style="list-style-type: none"> <li>■ = (等しい)</li> <li>■ != (等しくない)</li> <li>■ ISNIL</li> <li>■ IS NOT NIL</li> <li>■ ONE OF</li> <li>■ NOT ONE OF</li> </ul>
フラグ	<ul style="list-style-type: none"> <li>■ = (等しい)</li> <li>■ != (等しくない)</li> <li>■ ISNIL</li> <li>■ IS NOT NIL</li> </ul>

- ▶ 入力タイプに該当する値を指定します。非数値の値には、選択肢のドロップダウン リストが表示されます。

あるいは、[既存ルールの検索](#) を選択し、定義済みのルールを参照します。  
詳細は、[p. 51 ルールのエクスポートと再利用](#) を参照してください。

### セグメント ルール セット

- セグメント ルールは、結合してセットにすることができます。セグメント ルールセットは、個々のセグメントのいずれかが真の場合に、真の値を返します。これは、セグメント内部の式とは異なることに注意してください。式の場合は、すべてが真の必要があります。
- セグメント ルールセットは、ルールを使用している任意のアプリケーションからエクスポートすることができます。あるいは、利用可能であればIBM® SPSS® Rules Management からエクスポートすることもできます。 [詳細は、 p.51 ルールのエクスポートと再利用 を参照してください。](#)

## 選択ルールの定義

選択ルールとは、それぞれに包含または除外属性が割り当てられたセグメント ルールまたはセグメント ルール セットで構成されるセットです。選択ルールは、アプリケーションのモデリング、スコアリング、その他の処理でレコードのサブセットを選択するために使用されます。

- ▶ ルールを追加したいアプリケーションで、**新規ルールの作成** アイコンをクリックします。 このアイコンは、選択ルールが適用できる場所で利用できます。
- ▶ ルールの名前を指定します。このルールを包含または除外のいずれに使用するか（または後で行うか）を選択し、必要に応じて、1 つ以上の式を追加します。 [詳細は、 p.47 セグメント ルールの定義 を参照してください。](#)
- ▶ OK をクリックして、選択ルールを保存します。
- ▶ 包含/除外列のアイコンをクリックすれば、各ルールに対して、この設定をトグルできます。参照されている選択ルールの場合、この設定は変更できません。 同一ルール内で、包含ルールおよび除外ルールを混在させることはできません。

あるいは、**既存ルールの検索** を選択し、定義済みのルールを参照します。[詳細は、 p.51 ルールのエクスポートと再利用 を参照してください。](#)

### 選択ルールの評価

**選択ロジックの表示** を選択し、選択の内容（論理積か論理和か）を表示します。

- 選択ルールの評価では、複数の包含セグメントが指定されている場合は、レコードは、セグメントのいずれかに一致すれば包含されます。たとえば、顧客は、信用格付けまたは保有期間に基づいて包含されます。

- 複数の除外セグメントが指定されている場合は、レコードは、セグメントのいずれかに一致すれば除外されます。たとえば、顧客は、年齢または信用格付けに基づいて除外されます。
- 競合が発生した場合、除外ルールが優先されます。たとえば、除外セグメントのいずれかに該当した顧客は、適格性が認められるセグメントの存在の有無にかかわらず、除外されます。
- IBM® SPSS® Rules Management で選択ルールを作成する場合、すべてのルールが、同一タイプ（包含または除外のいずれか）である必要があります。同一ルールセット内で、包含ルールおよび除外ルールを混在させることはできません。

## ルールのエクスポートと再利用

デフォルトでは、ルールはアプリケーション内で作成され、そのアプリケーションで保存されます。別の方法として、ルールを IBM® SPSS® Collaboration and Deployment Services Repository にエクスポートまたは保存し、複数のアプリケーションで使用できるようにすることもできます。このケースでは、各アプリケーションは特定の、ラベル付けされたバージョンの共有ルールを参照します。このルールバージョンの変更は、それを使用しているすべてのアプリケーションに適用されます。

### ルールのエクスポート

- ▶ 共有したいローカルルールを含む任意アプリケーションで、**エクスポート** アイコンをクリックし、ルールを IBM SPSS Collaboration and Deployment Services Repository に保存して、必要に応じて再利用ができるようにします。
  - 複数のルールをエクスポートする場合は、それらすべてを単一のルールセットとしてエクスポートするか、または個々のルールとしてエクスポートするかを指定します。個々のルールとしてエクスポートする場合は、各セグメントが独立したルールとしてエクスポートされます。
  - 複数の選択ルールをエクスポートする場合は、**[保存選択]**を選択すると、選択ルールセットがエクスポートされます。このケースでは、包含または除外のいずれかをエクスポートできますが、両方をエクスポートすることはできません。包含と除外を共にエクスポートする場合は、別々のルールセットとしてエクスポートする必要があります。**[保存選択]**を選択しない場合は、セグメントルールセットがエクスポートされます。
  - 集計ルールエディタでエクスポートする場合は、**リスクポイントと共にエクスポート**を選択すると、集計ルールセットが作成されます。このオプションを選択しない場合は、セグメントルールセットがエクスポートされます。

### 既存ルールの使用法

- ▶ 共有ルールを利用する任意のアプリケーションで [既存ルールを検索] アイコンをクリックし、IBM SPSS Collaboration and Deployment Services Repository を参照して、現在のアプリケーションで再利用できる既存のルールを検索します。
- ▶ 特定のラベル付きバージョンを参照するには、ルール名のとなりの矢印をクリックし、必要なバージョンを選びます。バージョンが選択されていない場合、最新バージョンが使用されます。（最新とは新たなバージョン作成されるたびに自動的に公開される特殊なラベルと考えます。）新しいバージョンのルールが作成された際に発生し得る予期しない結果を避けるために、特定のラベルを付けたバージョンの使用をお奨めします。

参照したルールは、多くの場合、ローカルルールと結合できますが、参照したルールは、太字テキストおよびルール（またはルールセット）のタイプを示すアイコンによって、ローカルルールとは視覚的に区別できます。

- ▶ これまでとは逆に、必要に応じて、参照しているルールをローカルルールに変換し、現在のアプリケーションに保存することができます。そのためには、[ルールプレビュー] ダイアログボックスで、参照しているルールの名前をクリックし、[単一ルールへ変換] を選択します。

### 共有ルールのガイドライン

共有ルールを参照し、操作する場合は、以下のガイドラインが適用されます。

- ルールは再利用のため任意のアプリケーションからエクスポートできますが、一度エクスポートすると、編集できるのは IBM® SPSS® Rules Management のみとなり、他のいかなるアプリケーションでも編集できなくなります。注 IBM SPSS Rules Management はすべてのサイトにインストールされているとは限らず、またすべてのユーザーが利用可能とは限りません。
- 特定のラベル付きバージョン（最新以外）、特に運用環境のものを使用を強く推奨します。特に開発環境においては、継続した変更されている最新バージョンは意図しない結果を招くことがあるためです。場合によっては、ルール内からルールを参照した場合などはラベル付きバージョンを使用する必要があります。
- 単一ルールは、参照によって IBM SPSS Collaboration and Deployment Services Repository から追加できます。または [ルールプレビュー] ダイアログボックスで、ローカルルールに変換できます。
- 単一ルールは、任意のルールエディタを使用して、参照したり、ローカルルールと結合することができます。ルールセットとしてエクスポートした場合は、単一ルールへのネストした参照はそのまま保たれます。

- セグメント ルールセットと選択ルールセットは、多くの場合、単一ルールと同様に参照や結合が可能です。結合したルールセットは、ネストしたルールセットを単一ルールに変換した後でないと、エクスポートすることはできません。したがって、エクスポートしたルールセットは、ネストしたルールへの参照を含むことはありますが、ネストしたルールセットへの参照を含むことはありません。
- 集計ルールセットも参照は可能ですが、他のルールとの結合はできず、他のルールへのネストした参照を含むこともできません。集計ルールセットを参照する場合は、集計ルール エディタの内容全体が参照するルールで置き換えられます（逆に、集計ルールセットを参照しない場合は、集計エディタでは、ローカル ルール、参照セグメント ルール、セグメント ルールセットを混在させることができます）。
- IBM SPSS Rules Management を使用する場合、エクスポートするルールには同じ制限が適用されます。ネストしたルールは許可されますが、ネストしたルールセットは禁止されます。これは、IBM SPSS Rules Management が（他のアプリケーションからのエクスポートと同様に）共有ルールの作成および編集を目的として設計されていることが理由です。
- IBM SPSS Rules Management 等で、参照されているルールセットに関する制約を受けないようにするには、ルールセットは独立したローカル ルールに変換しておく必要があります。

## 注釈を追加

集計または割り当てで、**注釈を追加** アイコンをクリックし、フリー テキスト フィールドの列を追加して、そこにルール出力に含める注釈を記述します。たとえば、これは、コール センタのオペレータに対して、保険請求に問い合わせが必要というフラグが付けられている場合に問い合わせ先を知らせるための追加情報です。

注釈は、ディメンジョンで指定された **値を返す** の値と共に使用され、追加情報を示します。**値を返す** の値はディメンジョンまたは返される判断に関連付けられていますが、注釈はその判断への入力として使用されたルールを示します。

## ILOG Business Rule Management System のルールの適用

ILOG などのビジネス ルール マネージメント システムで作成されたルールは、IBM® Analytical Decision Management アプリケーションで参照し、使用することができます。これらの外部ルールは、他のルールと同じ方法で選択し、使用することができます。アプリケーションの必要な個所で**[既存ルールを検索]** アイコンをクリックします。外部ルールはアイコンによって区別することができますが、それ以外の点では他は標準方法で使用できます。

IBM Analytical Decision Management に使用する外部ルールの作成に関する情報に関しては、Application Designers Guide をご覧ください。

外部ルールは、他のルールと同じ方法でアプリケーションに挿入されます。ルールの出力はルールを使用するコンテキストに一致しなければなりません。たとえば、決定を入力するために、真/偽、または、はい/いいえ（ブーリアン）の値が割り当てられる必要があります。

他のルールと同様に、アプリケーションで使用されるいずれかの外部ルールの名前をクリックして、入力やフィールド、それに、ルールの説明、その他のプロパティを含むルールの詳細を表示させます。

外部ルールに複数の出力フィールドがある場合には、プロンプトが表示されて使用したいフィールドを選択することができます。

モデルの代わりに外部ルールが使用されている場合には、ターゲットのフィールドは表示されません。そして、ドロップダウンリストに構成されたすべての出力が表示されます。

#### 外部ルール使用のヒント

- 真 または 偽（または はい/いいえ もしくは 0/1）の値を返すルールは、いずれの選択または割り当てとしても、ルール エディタ上のいずれかのルールの入力としても、または、個々のディメンション項目に関連するフィールドの返すとしても、使用することができます。
- カテゴリー出力（赤、白、または青、もしくは北、南、東、西）を返すルールは、[定義]タブの **決定のためにモデルを使用する** セクションで、ルール エディタの入力として、または、個々のディメンション項目に関連するフィールドの返すとしても、使用することができます。
- 数値出力を返すルールは、[定義]タブの **決定のためにモデルを使用する** セクションで、優先順位付けへの入力として、ルール エディタの入力として、または、個々のディメンション項目に関連するフィールドの返すとして使用することができます。

# 予測モデルの作成

## 予測モデルについて

モデルは、過去のデータに基づいて、将来発生する可能性がある事態を予測するために使用します。たとえば、モデルを使用すれば、収入、年齢、および加入している組織やメンバーシップなどの特性に基づいて、解約の確率が低い顧客や、特定のオファーに応ずる確率が高い顧客を予測することができます。

モデルはルールと同様に使用できますが、ルールが企業のポリシー、ビジネス ロジック、その他の前提条件に基づくのに対し、モデルは過去の結果の実際の観察に基づいて作成されるので、観察しなければわからなかったパターンを検出することができます。ルールによってアプリケーションに日常のビジネス ロジックが付与され、モデルによって洞察および予測能力が付与されます。

### モデルの作成およびスコアリングのためのデータ

モデリング プロセスでは、次の 2 種類のデータが使用されます。

- モデルを作成するには、予測対象についての情報が必要です。たとえば、解約を予測する場合は、過去に解約した顧客についての情報が必要です。これは、しばしば、履歴データまたは分析データと呼ばれ、プロジェクト データ モデルの一部またはすべてのフィールドと、予測の成果すなわち結果を記録する追加フィールドを含む必要があります。この追加フィールドは、モデリングの**ターゲット**として使用されます。
- 将来の結果を予測するためにモデルを使用するには、たとえば、潜在顧客、受け取るクレームのような関心のあるグループまたは母集団に関するデータが必要です。。これは、**運用**データまたは**スコアリング** データと呼ばれます。プロジェクト データ モデルは、通常、このファイルをベースにします。
- データ モデルにターゲット フィールドが含まれる場合は、このフィールドを **運用** 列として選択しないでください。モデルのスコアリング時には利用できないからです。スコアリングの目標は、履歴データを使用してモデルの作成を完了した後、結果がわかっていない新しいデータにモデルを適用することです。

## モデルの作成

[モデリング] タブを使用すると、データ ソース、ターゲット フィールド、その他のオプションを指定することができます。また作成するモデルのタイプを選択できます。

パラメータがIBM® SPSS® Modeler中のアプリケーション（ストリーム ファイル）に追加された場合、**パラメータ リンク**が [モデル] タブならびに [スコア] タブ上で利用できるようになることに注意してください。**パラメータ リンク**をクリックすると、定義されたパラメータとその現在の値が表示され、モデル作成またはスコアに影響するパラメータ値の変更が可能となります。

新しいモデルの作成は、次のいずれかの方法で行います。

### 予測モデルの作成

自動化モデルを作成する場合に指定する必要があるのはデータ ソースとターゲット フィールドのみです。データを準備してパーティション分割するためのオプションや、使用するフィールドやレコードを選択するオプションを選択することもできます。ターゲット タイプに適したメソッド範囲を適用すると、選択したデータに最もふさわしい技法またはその組み合わせが自動的に選択され、モデルが生成されます。

スコアリングでは、モデルは、予測およびその予測に関連付けられた確信度の値を含むフィールドを 1 つ以上返します。たとえば、ターゲット フィールドが **回答** という名前のフラグ型フィールドの場合、モデルは、そのフィールドの予測値を含む **XF-回答** という名前のフィールドを返します。

- ▶ データ ソースの指定:これには、予測対象の結果を記録する任意の分析データ ソースを選択できます。
- ▶ 予測対象の結果を記録するターゲット フィールドを指定します。たとえば、ターゲット フィールドには、解約した顧客、過去のオファーに応じた顧客、不正請求を行った顧客等が表示されます。

他のフィールドは、この値の予測を補足する入力として、すべて自動的に残されます。

- ▶ 必要に応じて、オプションの設定を指定します。 [詳細は、 p.64 オプションのモデル設定 を参照してください。](#)
- ▶ **モデルの作成** をクリックします。

モデルの作成中は、ブラウザを閉じてしまったり、その他の作業を行うことができます。モデルは、作成が完了すれば、ギャラリーからアクセスできます。 [詳細は、 3 章 p.23 ギャラリー を参照してください。](#) モデルの作成



が予想したより時間がかかるようであれば、**停止** ボタンをクリックして、前に保存したバージョンに戻すこともできます。

モデル作成が完了すると、分布図と予測の重要度グラフが表示されます。詳細は、[p. 57 予測モデリングの結果](#) を参照してください。

- ▶ オプションで、評価とテスト機能を使用して、モデルのサンプル データに対する効果を表示することができます。詳細は、[p. 68 モデルの評価](#) を参照してください。

## 予測モデリングの結果

予測化モデルの実行が完了すると、モデルのパフォーマンスと、最終結果に対する各予測因子の貢献度が視覚的に評価できるように、分布図と変数の重要度グラフが表示されます。

### 分布図

モデルの分布図では、観測値は横軸に表示され、予想値は縦軸またはオーバーレイで表示されます。これにより、それぞれの値すなわち回答を予測した回数と、予測が正しかった頻度を確認することができるため、モデルの精度が視覚的に評価できることとなります。

- フラグ型ターゲットまたはセット ターゲットの場合、分布は棒グラフを使用して表示されます。このとき、ターゲット フィールドのカテゴリごとに 1 つの棒が使用されます。各棒の高さは、結果がモデルによって予測された回数を示します。各棒の色のついた領域は、マッチした結果が実際にテスト データで観察された回数を示します。モデルが合理的に正確な場合は、各棒内の最大の領域は、対応する回答にマッチしているはずです。
- 範囲ターゲットの場合、分布はビン散布図で表現されます。ここでも、予想値は横軸にプロットされ、観測値は縦軸にプロットされます。点が対角線に沿って集まるのが理想的です。点（結果）が散在すればするほど、そして線形分布から遠ざかれば遠ざかるほど、モデルの精度はその分低下していることとなります。

### 予測の重要度グラフ

通常、ユーザーはモデル作成の目標を最も重要な予測フィールドに焦点を当て、最も重要でない予測フィールドを削除または無視したいと考えます。予測の重要度グラフは、モデルを推定する際に各予測の相対重要度を示すことによってこれを支援します。値は相対的なもので、表示されるすべての予測の値の合計は 1.0 です。予測の重要度はモデルの精度と関係はありません。予測する際の各予測の重要度と関係するだけで、予測が正確かどうかには関係しません。

予測の重要度の計算には、特に大きなデータセットを使用する場合、モデル構築よりもはるかに長い時間がかかることがあります。

モデルの作成時にデータの自動パーティション分割オプションを選択した場合、評価ページではテストパーティションが自動的に選択されます。

データのパーティション分割を有効にしていない場合は、モデル評価用のデータソースを選択する必要があります。モデルの作成で使用したデータソースとは異なるデータソースを選択することをお勧めします。データのパーティション分割はデフォルトで有効になっているため、この設定を変更していない限り、改めて有効にする必要はありません。

## 予測ルールモデルの作成

予測ルールモデルを使用すると、特定の結果に関して、高い確率値または低い確率値を示すセグメントを特定することができます。たとえば、解約の確率が低い顧客や、特定のオファーに応ずる確率が高い顧客を探すことができます。結果のモデルは、選択ルールのリストに似ており、各セグメントのスコアリング方法を指定する包含属性または除外属性を含んでいます。

予測ルールモデルは選択ルールに似ていますが、スコアリング方法は異なります。選択ルールの場合は、除外したレコードは削除されるため、出力内のレコードは少なくなります。モデルの場合は、削除されるレコードはなく、各レコードについて予測および傾向の値を示す追加フィールドが追加されます。包含されたセグメントは、真としてスコアリングされます（または、予測対象によっては、偽または真偽いずれかとしてスコアリングされます）。除外されたセグメントは、どのルールにもマッチしないため「残り」に分類されるレコードになるので、NULLとしてスコアリングされます。

各レコードには、ターゲットの回答の確率値を示す傾向スコアも付属しているため、応ずる確率が高い顧客を特定することができます。このため、予測ルールモデルはマーケティングリストの作成に適しており、コールセンターやマーケティングアプリケーションを含む、カスタマリレーションシップ管理（CRM）で広く使用されます。

- 予測ルールモデルを適用できるのは、カテゴリーで分類可能なターゲットのみです。
- スコアリング目的で包含または除外するセグメントを選択するため、セグメントを表示、変更、再構成することができます。たとえば、顧客の特定のグループを将来のオファーから除外し、他のグループを包含してみた場合のヒット率に与える影響を調べることができます。

- 選択ルール使用時と同様に、必要に応じて、セグメントの手動での追加または編集が可能です。モデルにオプションの選択を委ねることは多くの場合に簡単な方法ですが、多数用意されているオプションを使用すれば結果を調整することができます。
- ルールはリストされている順番に適用され、最初にマッチしたルールで、特定のレコードに対する結果が決定されます。個々を取り上げれば、重複するルールあるいは条件が存在する可能性があります。このあいまいさは、ルールの順番によって取り除かれます。マッチするルールがない場合、レコードは残りのルールに割り当てられます。

### 予測ルールモデルを使用するには

- ▶ データ ソースの指定: データ ソースとしては、予測対象の結果を記録する任意の分析ソースを選択することができます。
- ▶ 予測対象の結果を記録するターゲット フィールドを指定します。
- ▶ 検索対象の特定の値または回答を指定します。たとえば、ターゲットフィールドが 回答 という名前の場合は、回答のあった顧客または回答のなかった顧客を検索するために、真 または 偽 をそれぞれ選択することができます。

ターゲットフィールドには、例えば、解約した顧客、過去のオファーに応じた顧客、または不正請求を行った顧客などの情報が示されます。他のフィールドは、この値の予測を補足する入力として、すべて自動的に残されます。

- ▶ 必要に応じて、オプションの設定を指定します。 [詳細は、 p. 64 オプションのモデル設定 を参照してください。](#)
- ▶ [作成を開始] をクリックします。
- ▶ 指定した目標よりも高い確率または低い確率のセグメントを検出するかどうかを指定します。たとえば、解約の確率が高いまたは低いグループまたは顧客を捜し出し、それを包含することができます。
- ▶ 検出するセグメントの最大数を指定します。一般にこの値は、単純さとパフォーマンスのために小さく保つ必要があります。
- ▶ [モデルの拡張] をクリックし、高い確率または低い確率を持つグループを特定する、1 つ以上のセグメント ルールで構成されるリストを取得します。

- ▶ **包含/除外** アイコンをクリックして、ルールの使用法と、関連セグメントのスコアリング方法を指定します。

包含したセグメントには、ターゲットの値（たとえば、回答 = 真）にマッチするスコアが割り当てられます。除外したセグメントは NULL としてスコアリングされますが、ファイルから削除されることはありません。

- ▶ オプションで、必要に応じて、セグメントの追加または編集が可能です。[詳細は、5 章 p.47 セグメント ルールの定義 を参照してください。](#)
- ▶ オプションで、評価とテスト機能を使用して、モデルのサンプル データに対する効果を表示することができます。

### ヒント

- 上/下の矢印を使って、ルールを上下に移動します。各レコードが最初のヒットに基づいてスコアリングされるため、ルールの順序が結果を変化させることとなります。たとえば、複数のルールにマッチするレコードでも、スコアリングされるのは最初に適用されたルールのみです。
- 1 つ以上のオプションを変更し、再び **モデルの拡張** をクリックして、追加ルールを検出します。たとえば、最小セグメントのしきい値を下げたり、信頼間隔を狭くすることにより、追加セグメントを特定します。
- **モデルの拡張** をクリックしても、テーブルはクリアされず、モデルが最初から再構築されるわけではありません。単に、テーブルにセグメントルールが追加されるだけです。

## クラスタリング モデルの作成

クラスタリング モデルにより、レコードを一定の数のクラスタに分類できます。これは、データ内の自然グループを特定するのに便利です。

クラスタリング モデルは、類似したレコードのグループを識別し、そのグループに従ってレコードにラベルを付けます。この操作には、グループやその特性に関する事前の知識は必要ありません。実際には、検索するグループ数が正確にわからない場合もあります。これが、クラスタリング モデルと他のマシン学習技法との違いであり、-クラスタリング モデルには、モデルが予測する定義済みの出力フィールドや対象フィールドはありません。クラスタリング モデルは、モデルの分類性能を判定する外部標準がないので、**非監視学習モデル**と呼ばれることがよくあります。これらのモデルには、正、誤という回答はありません。モデルの価値は、データのグループ構成を把握し、それらのグループについて役に立つ説明を提供できるかどうかで決まります。

クラスタリング手法は、レコード間およびクラスタ間の距離の測定に基づいています。レコードは、同じクラスタに属するレコード間の距離を最小にするようにして、クラスタに割り当てられます。

クラスタリング モデルは、クラスタやセグメントを作成するためによく利用されます。このクラスタやセグメントは、後の分析で入力として使用されます。一般的な例として、マーケット セグメントがあります。これは、マーケティング担当者がマーケット全体を等質のサブグループに細分化するために使用します。各セグメントには、そのセグメントを対象としたマーケティングの成果に影響する特性があります。データ マイニングを使用してマーケティング戦略を最適化している場合は、適切なセグメントを識別し、そのセグメント情報を予測モデルで使用することで、モデルを大幅に改善できます。

### クラスタリング モデルの作成

- ▶ データ ソースの指定: データ ソースとしては、予測対象の結果を記録する任意の分析ソースを選択することができます。
- ▶ 必要に応じて、オプションの設定を指定します。 [詳細は、 p. 64 オプションのモデル設定 を参照してください。](#)
- ▶ [セグメントの検出] をクリックします。
- ▶ オプションで、手動クラスタを追加できます。 [詳細は、 p. 61 手動クラスタの使用 を参照してください。](#)
- ▶ オプションで、評価とテスト機能を使用して、モデルのサンプル データに対する効果を表示することができます。

### 手動クラスタの使用

必要に応じて、クラスタを手動で定義でき、自動化クラスタリングを使用して、手動で定義したどのクラスタでも特定されないレコードのクラスタを検索できます。「最初にヒットした」セグメント ベースの割り当てルール セットを介して、手動クラスタを作成します。作成したルールにより、クラスタが定義され、自動クラスタ モデリングに対してリメインダが使用されます。

手動クラスタは結果要約図には含められません。

- ▶ [手動クラスタ] で、[ルールアイコンを新規作成] をクリックし、1 つ以上のルールを追加して、手動クラスタを定義します。
- ▶ 必要に応じて、クラスタのラベル名を変更します。
- ▶ [手動クラスタ内のレコード] アイコンをクリックして、手動クラスタ内に含まれるレコードの合計数のレコード カウントを実行します。カウント 列が定義した各手動クラスタに対して更新されます。

## アソシエーション モデルの作成

アソシエーション モデリングにより、一緒に表示される可能性の高い項目を予測でき、それらの間の関係の強さを予測できます。

アソシエーション ルールは、特定の結果（特定の製品の購入）と条件セット（複数の他の製品の購入）を関連付けます。たとえば、次のルール

ビール <= 缶詰野菜 & 冷凍食品 (173, 17.0%, 0.84)

は、缶詰野菜と冷凍食品を同時に購入するときに、ビールがよく購入されることを示しています。このルールは信頼度 84% で、データの 17 %、つまり 173 個のレコードにあてはまります。アソシエーション ルールのアルゴリズムは、ユーザーがなどの視覚化手法を使用して手動で見つけていた連関を、自動的に見つけ出します。

アソシエーション ルールのアルゴリズムは、任意の 属性の間にアソシエーションを成立させることができるという点で、より一般的なディシジョン ツリーのアルゴリズム (C5.0 や C&R Trees など) より勝っています。ディシジョン ツリーのアルゴリズムは、一つの結果にいたるルールを構築するのに対し、アソシエーション ルールのアルゴリズムは、それぞれ異なる結果にいたる多数のルールを見つけようとします。

アソシエーション ルール アルゴリズムは、パターンの検索範囲が非常に大きくなる可能性があり、そのためディシジョン ツリーのアルゴリズムを実行するよりもはるかに時間がかかるという欠点があります。アソシエーション ルール アルゴリズムでは、ルール検索に**生成と検定**の手法を使用します。最初に簡単なルールが生成され、データセットに対して妥当性が検証されます。有効なルールは保存され、その後でさまざまな制約に従って、すべてのルールが特殊化されます。**特殊化**とは、条件をルールに追加する処理のことです。次に、データに対して新しいルールの妥当性が検証され、この繰り返しによって、最善または最も関心の高いルールが保存されます。通常、ユーザーは、ルールで許可する推定数を制限し、情報理論に基づく各種の手法や効果的なインデックス作成方法を使用して、広大になる可能性がある検索範囲を縮小します。

### アソシエーション モデルを取得するには

- ▶ データ ソースを選択します。データ ソースとしては、予測対象の結果を記録する任意の分析ソースを選択することができます。
- ▶ データ形式を指定します。詳細は、[p. 63 表形式データ対ランザクシオンデータ](#) を参照してください。
- ▶ 必要に応じて、オプションの設定を指定します。詳細は、[p. 64 オプションのモデル設定](#) を参照してください。
- ▶ **モデルの作成** をクリックします。

- ▶ オプションで、評価とテスト機能を使用して、モデルのサンプル データに対する効果を表示することができます。評価はアソシエーション モデル タイプには使用できません。

## 表形式データ対ランザクションデータ

Wアソシエーション モデルの作成時、表 または トランザクション データ形式のいずれかを選択できます。デフォルトでは、表形式が選択されます。トランザクション を選択したら、表示されるドロップダウン メニューから ID および 内容 のフィールドを選択します。

### トランザクション形式のフォーマット

トランザクション形式のデータには、各トランザクションまたは項目に対応する独立したレコードがあります。たとえば、顧客が複数の買い物をした場合、それぞれが顧客 ID にリンクされた項目に関連付けられた、独立したレコードになります。これは、時にtill-roll形式としても知られています。

顧客	購入品
1	ジャム
2	牛乳
3	ジャム
3	パン
4	ジャム
4	パン
4	牛乳

### テーブル形式のデータ

テーブル形式のデータ（バスケットまたは**真理値表**データとも呼ばれる）には、フラグで区切られて表現された項目があります。各フラグ型フィールドで、特定の項目の有無が表現されます。各レコードで、関連付けられている項目の完全セットが表現されます。フラグ型フィールドは、カテゴリまたは数値とすることができます。ただし、ある種のモデルでは、さらに特定の要件があります。

顧客	ジャム	パン	牛乳
1	T	F	F
2	F	F	T
3	T	T	F
4	T	T	T

## オプションのモデル設定

モデルの作成時に、次のオプション設定を選択することができます。既存のモデルに対してこれらの設定を変更した場合は、変更を適用するためにモデルを作成し直す必要があります。

**検証およびテスト用のデータソースの作成についてのモデル検証を有効にするために、自動的にデータを区分します。** データを、モデルのトレーニング用およびテスト用の独立したサブセット、つまりサンプルに分割します。モデルを特定のサブセットに基づいて作成し、それを別のサブセットでテストすれば、他のデータセットに対して一般化する方法がわかります。

**信頼できるモデル作成のために自動的にデータをクリーンアップして準備します**  
データの問題を特定して修復することで、モデリングの時間を短縮し、予測性能を向上させ、信頼性を高めることができます。欠けている値および極端な値の処理、必要に応じた新しい属性の作成、インテリジェントなスクリーニングとサンプリング手法によるパフォーマンスの改善などにより、問題のあるフィールドや役に立たない可能性があるフィールドをスクリーニングします。新規データソースを使用してモデルを初めて作成する場合は、分析では最初のパスのみを実行して、パフォーマンスを劣化させる可能性がある問題とその修正案を特定します。以降の実行では、修正案が適用されますが、分析はデータソースが変更されるまで繰り返して行われることはありません。この設定は、エキスパートユーザーによって指定されたカスタムデータの準備設定があるモデルを含め、一部のモデルに対して無効にできます。

**使用する入力を指定:** 使用するフィールドを選択することができます。通常、選択するフィールドは、年齢や収入など、予測する項目と実際に関係があるフィールドです。データセットが大きい場合、フィールドの数を制限することは、モデルを簡略化するための 1 つの方法です。データに顧客 ID または連絡先電話番号のようなフィールドが含まれていても、通常モデル作成には役立たないため選択されません。他のデータと重複するフィールドも除外されます。

リンクされた入力フィールド (式) をクリックすると、その式のための式ビューアが開きます。式を編集するには、[データ] タブを使用します。  
詳細は、4 章 p. 45 式エディタ を参照してください。

**使用する選択を指定:** モデルの作成時に包含するレコードまたは除外するレコードを指定します。既存のルールを検索するか、または新しいルールを作成します。詳細は、5 章 p. 50 選択ルールの定義 を参照してください。



また、グローバル選択が定義されていた場合は、ここに表示されるので、モデル作成中にそれも適用する必要があるかどうかを指定することができます。グローバル選択ルールをすべて使用するか、あるいは全く使用しないかのいずれかを選択できます。サブセットの選択はできません。

## セグメント化のオプション

予測ルールモデルを作成する場合、多数のセグメント化オプションを利用できます。他のタイプのモデルで利用可能なオプションに加えて、これらのオプションが利用できます。

**最小セグメント サイズ:** 以下の設定は最小セグメント サイズを指示します。2つの値の大きいほうが優先されます。たとえば、パーセンテージ値が絶対値よりも大きい場合は、パーセンテージの設定が優先されます。

- **%の前のセグメント:** グループの最小サイズをレコードのパーセンテージで指定します。設定可能な最小値は 0 です。設定可能な最大値は 99.9 です。
- **絶対値:** グループの最小サイズをレコードの絶対数として指定します。設定可能な最小値は 1 です。最大値はありません。

**最大属性数:** 1つのセグメント ルールについて、条件の最大数を指定します。設定可能な最小値は 1 です。最大値はありません。

**属性の再使用を許可:** 有効にすると、前のサイクルで使用された属性も含めて、すべての属性が各サイクルで評価の対象になります。各サイクルで新しい条件が追加されるため、セグメントの条件はサイクルごとに累積していきます。サイクルの数は **最大属性数** 設定を使用して定義します。

**新しい条件の信頼区間:** セグメントの有意性をテストするための信頼レベルを指定します。この設定は、返されるセグメントがある場合はその数と、「1つのセグメントについての条件数」のルールに大きな影響を与えます。この値を大きくするほど、返される結果セットが少なくなります。設定可能な最小値は 50 です。設定可能な最大値は 99.9 です。

## 自動クラスタ オプション

クラスタリング モデルを作成する場合、多数の自動クラスタ オプションを利用できます。他のタイプのモデルで利用可能なオプションに加えて、これらのオプションが利用できます。

**評価フィールド。** 評価に使用するフィールドを指定します。評価フィールドが選択されると、IBM® SPSS® Modeler 自動クラスタ作成ノードでのランキング方法が、[評価の重要度] に設定されます。評価フィールドが選択され

ないと、ランキング方法はデフォルトのままになります。分析フィールドはここで評価フィールドとして選択できます。

**検索されたクラスタの数に必要な範囲を設定します。** 自動クラスタ モデルによって見つけられたクラスタの数に対して、必要な最小および最大範囲を指定します。

## アソシエーション ルールおよびスコアリング オプション

アソシエーション モデルを作成する場合、多数のアソシエーション ルールおよびスコアリング オプションを利用できます。他のタイプのモデルで利用可能なオプションに加えて、これらのオプションが利用できます。

### アソシエーション ルール オプション。

- **範囲(%)** ルールの全体的な適用範囲を管理するために最小範囲割合を指定してください。例えば、値を80%に設定すると、少なくとも80%のケースに適用されるということです。より高い値は、より少ないルールを生成または全く生成しないということを留意してください。
- **確信度(%)** 最小確信度割合を指定します。
- **最大事前条件数:** 最大事前条件数を指定します。
- **データはすでに ID 別に並び替えられています (トランザクション形式専用)** このオプションを選択すると、モデル作成時にデフォルトで行われる並び替えが無効になり、ID が連続するようになります。このオプションを選択すると処理時間が減ります。データ形式が票の場合、このオプションは効果はありません。

### スコアリング オプション。

- **最大予測数:** 含まれる予測最大数を指定します。このオプションは、ルール基準ドロップダウンとともに使用され、「最上位」の予測を行います。ここで、最上位というのは、確信度、サポート、リフトなどについて、最も高いレベルであることを示しています。
- **ルール基準。** ルールの強さを決定するために、測定を選択します。アイテムセットに最上位の予測を返すために、ここで選択した基準の強さによって、ルールがソートされます。利用可能な基準は次の通りです。

確信度は、ルール条件サポートに対するルールサポートの比率を表示します。これは、指定したルール条件を持つレコードの一部で、ルール予測も真 (true) となるものを示しています。例えば、学習データの50% がパン (ルール条件サポートです) を含むだけでなく、20% がパンとチーズの両方 (ルールサポートです) も含んでいる場合、パン<チーズ>というルールの予測は、ルール サポート / ルール条件サポートで計算され、この場合は、40% となります。

**範囲(%)**は、ルール条件サポート、すなわち、学習データを基準にしてルール条件が真 (true) である ID の比率を表示します。たとえば、学習データの 50% がパンを購入していれば、**パン>チーズ**というルールは、50% のルール条件サポートとなります。

**ルール サポート**は、ルール全体、ルール条件、ルール予測が真 (true) となるレコードの比率を表示します。たとえば、学習データの 20% がパンとチーズの両方を含む場合、**パン > チーズ**というルールのルールサポートは、20% となります。

**リフト**は、ルール予測が得られる事前確率に対するルールの確信度の比率を表示します。たとえば、母集団全体の 10 パーセントがパンを買う場合、パンを買うかどうかを 20 パーセントの確信度で予測するルールのリフトは、 $20/10 = 2$  になります。11 パーセントの確信度でパンを買うとするルールが他にあれば、そのルールのリフトは 1 に近くなります。これは、そのルール予測を持つ確率に対するルール条件の影響が大きいことを意味します。一般に、ルールのリフトが 1 に近い場合より、1 から離れた場合の方が、より興味深い結果が得られます。

**展開性**は、ルール条件を満足しつつルール予測を満足しない学習データの割合を示す尺度です。製品購入については、全顧客ベースで見て、ルール条件を所有し (あるいはすでに購入し) かつ未だルール予測を購入していない人の割合を基本的に意味しています。展開性の統計量は、 $(\text{レコード数中のルール条件サポート} - \text{レコード数中のルール サポート}) / \text{レコード数} * 100$ で定義されます。ここで、**ルール条件サポート**とは、ルール条件が真 (true) であるレコード数であり、**ルール サポート**とは、ルール条件とルール予測の両方が真 (true) となるレコード数のことです。

■ **トランザクション全体をチェック。** 次のオプションから選択します。

[予測を表示させない] により、ルール予測が表示されなくなります。たとえば、スコアリングの目的が、家庭で使う家具製品を推奨することであれば、**ダイニング テーブル**を既に含む**バスケット**が別のものを購入するケースはほとんどありません。このような場合、このオプションを使ってください。一方、製品が腐りやすかったり、使い捨てのものである場合 (チーズ、粉ミルク、ティッシュペーパーなど)、バスケットにすでに結論が入っているルールは、価値があります。後者の場合、最も便利なオプションは下[バスケットに予測があるかどうかを検査しない]です。

[予測を表示させる] により、ルール予測はバスケット内に表示されたままになります。このアプローチは、既存の顧客やトランザクションに対する洞察を得ようとする場合に役立ちます。例えば、最上位のリフトを持つルールを識別したり、さらにどの顧客がそのルールに適合するのかを調べたい場合があります。

[バスケットに予測があるかどうかを検査しない]により、バスケットの中にルール予測があってもなくても、スコアリングする場合には、すべてのルールが含まれます。

- **予測の繰り返しを許可:** このオプションを選択すると、スコアリング時の同じルール予測に複数のルールが含まれるようにします。たとえば、このオプションを選択すると次のルールのスコアリングができるようになります。

パン & ワイン > チーズ

チーズ & 果物 > ワイン

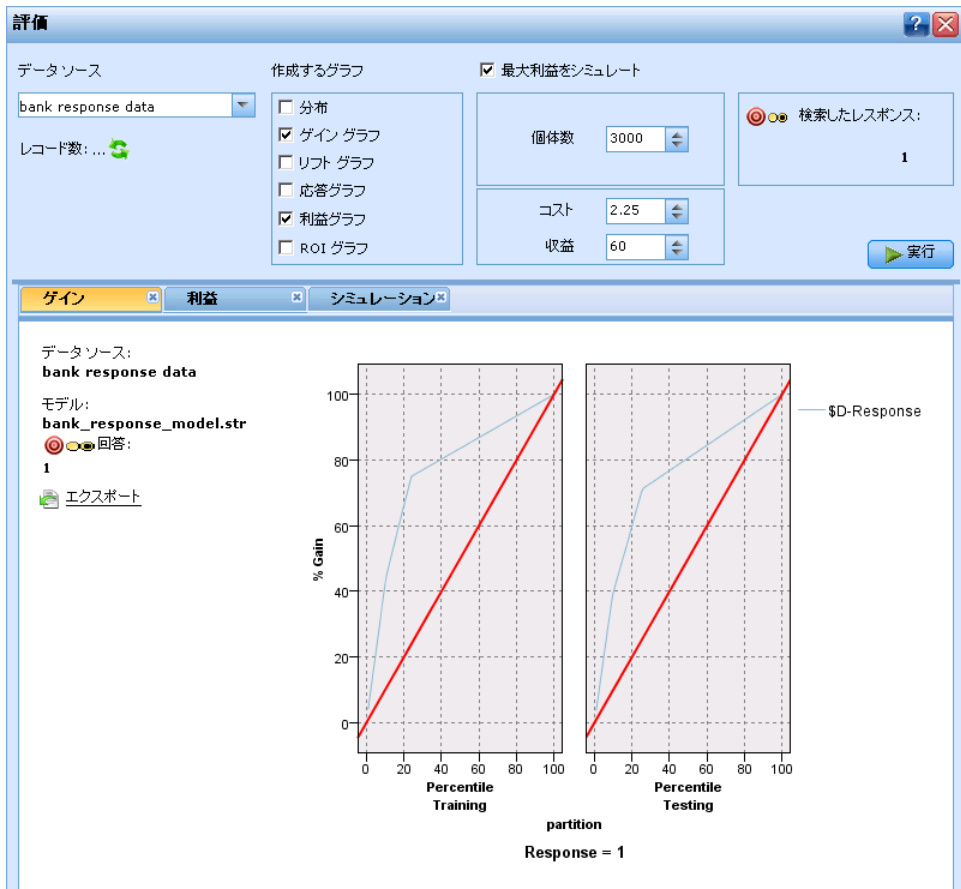
スコアリング時に予測の繰り返しを除外するには、このオプションをオフにします。

- **一致しない項目を無視。** このオプションを選択すると、アイテム セットに追加されたアイテムの存在は無視します。たとえば、このオプションが[テント&寝袋& やかん]を含むバスケットに使用された場合、たとえバスケットに余計なアイテム (やかん) があったとしても、テント&寝袋 > ガスストーブというルールが、適用されます。

## モデルの評価

完成したモデルを作成した後、または開いた後、[評価] を選択して、分布、ゲイン、リフト、回答、利益、および ROI の各グラフを取得し、最大可能利益をシミュレートします。このダイアログが表示されるのは、選択したモデルが完成している場合のみです。評価はアソシエーション モデル タイプには使用できません。

図 6-1  
ゲイン グラフ



- ▶ 上部のパネルで、評価データ セットを選択します。これは、モデルの作成と同様に、実際の回答が分かっている分析データ セットである必要があります(そうでないと、実際の回答を、モデルによって予測された回答と比較する方法がなくなります)。
- ▶ 作成するグラフを選択します。
- ▶ 利益をシミュレートする場合は、**最大利益をシミュレート** ボックスを選択し、利益計算のベースとなる母集団を入力します。 [詳細は、 p.75 最大利益をシミュレート を参照してください。](#)
- ▶ 利益または ROI グラフを要求している場合、または利益をシミュレートしている場合は、コストと収益の値を指定します。これらの値はそれぞれ、各オファーのコストと、好意的な各回答に対する期待収益の額に相当します。

- ▶ 「ヒットした」と見なす回答値を選択します。これは、積極的または見込みのある回答を示す 回答 = 真 のように、モデルの作成で使用されるターゲット フィールドの値の 1 つである必要があります。
- ▶ [実行]をクリックして結果を表示します。

### 分布図 (すべてのターゲット)

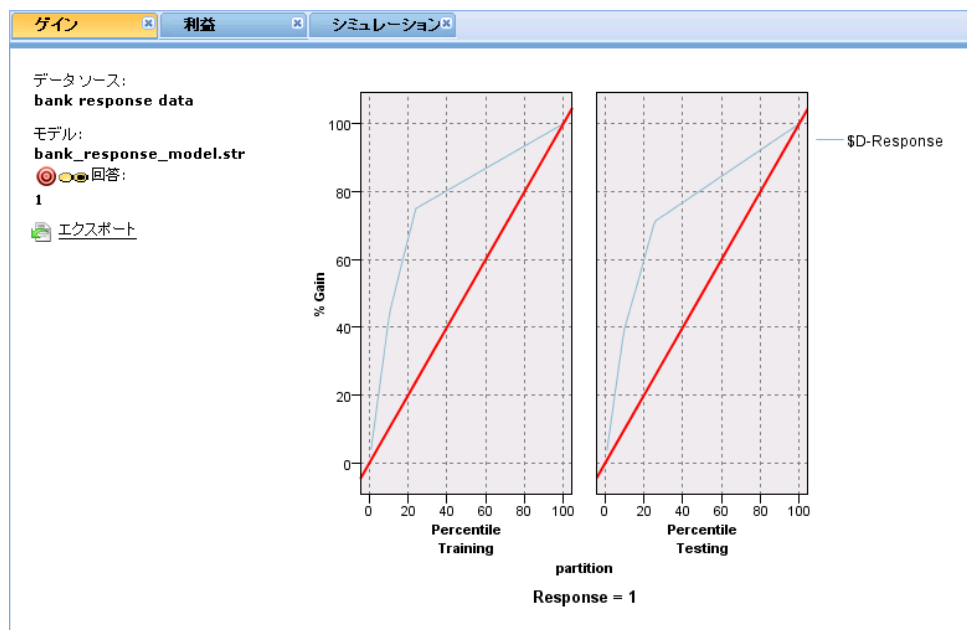
分布図は、モデルの完成時にデフォルトで表示され、予測値に対応させた観察値がプロットされます。[評価] ダイアログボックスを使用すると、データ セットをテストするだけでなく、データ セットの分布図を取得することもできます。範囲がターゲットの場合は、用意されるグラフは分布図のみです。

フラグ型ターゲットおよびセット ターゲットの場合は、次の追加グラフから選択することができます。

### ゲイン グラフ

ゲイン グラフは、モデルを使用して取得するゲインまたは「リフト」を表示するために使用され、利用可能なヒットの合計数と相対的な各増分内のヒットの割合として定義されます。斜めの線は、モデルが使用されなかった場合に、すべてのサンプルで期待される回答をプロットしたものです。この場合、1 人が別の人と同様に応答するため、レスポンス割合は定数になります。売り上げを 2 倍にするには、2 倍の人に質問する必要があります。曲線は、応ずる確率の高い顧客に的を絞ることで改善できる結果を示しています。カーブが急になればなるほど、ゲインも高くなります。

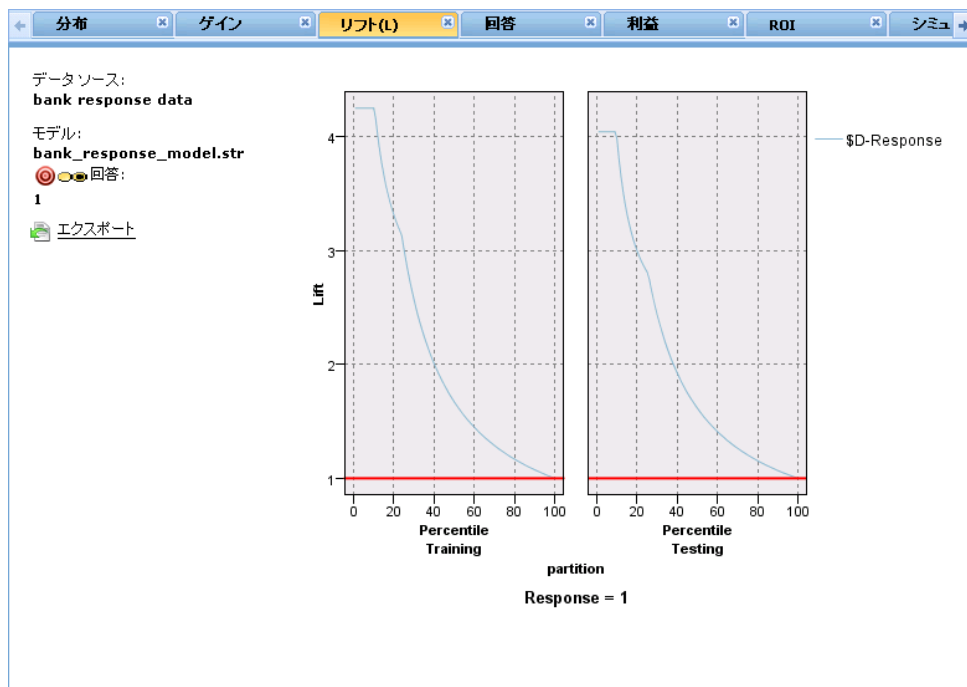
図 6-2  
ゲイン グラフ



## リフト グラフ

リフト グラフでは、トレーニング データ セット内の全体的なヒット割合に対して、各増分内のヒットしたレコードの割合がプロットされます。

図 6-3  
リフト グラフ

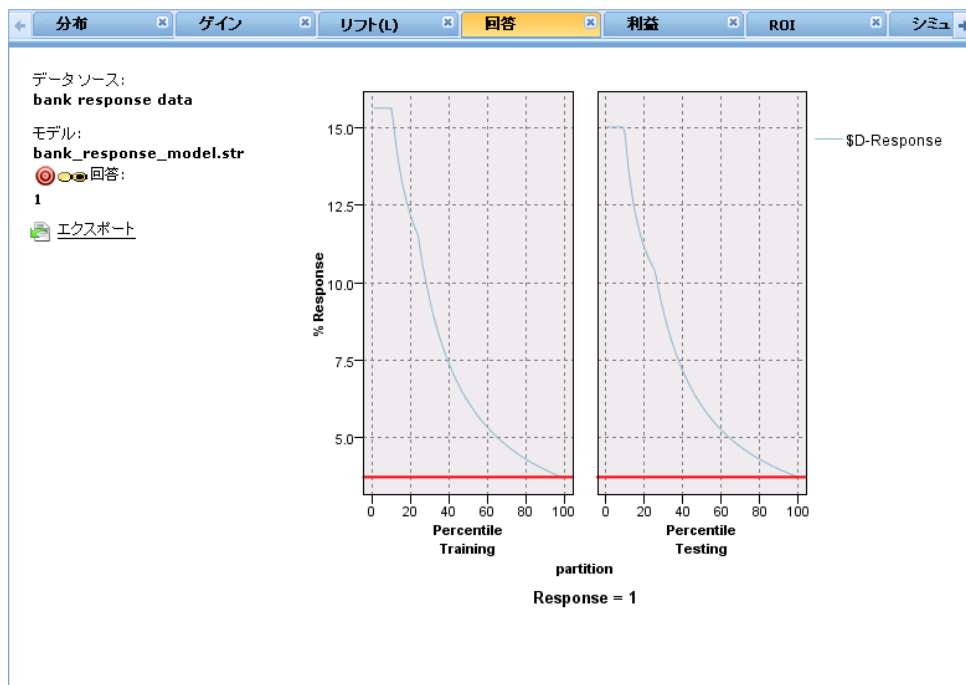


### レスポンス グラフ

回答グラフは、ヒットした増分内のレコードの割合をプロットします。



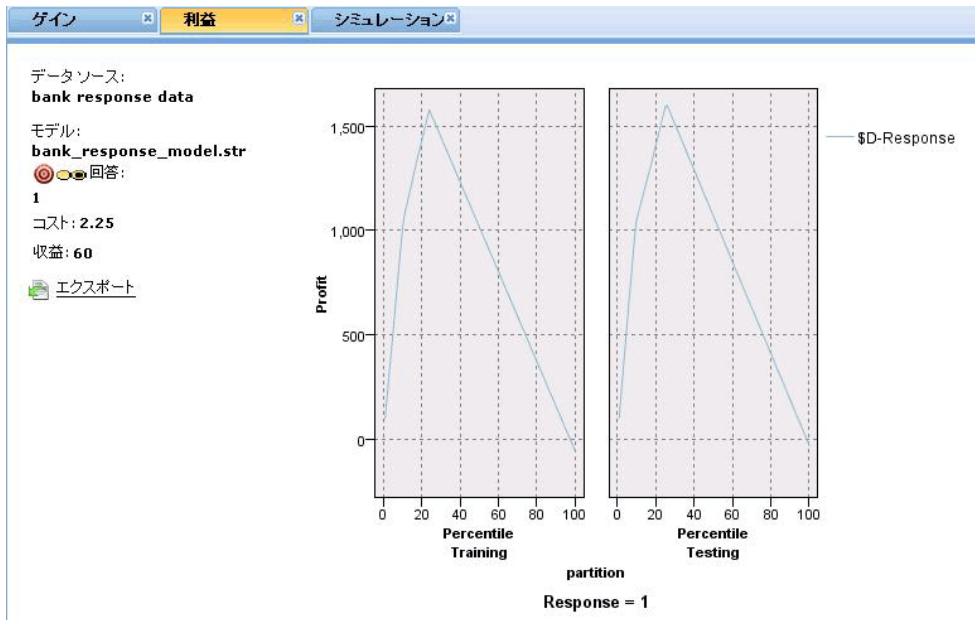
図 6-4  
応答グラフ



## プロフィット グラフ

利益は、各レコードの収益から、そのレコードのコストを引いた値です。分割の利益は、その分割内の全レコードの利益を単純に合計したものです。収益はヒットのみに適用されることが前提ですが、コストはすべてのレコードに適用されます。 [詳細は、 p. 75 最大利益をシミュレート](#) を参照してください。

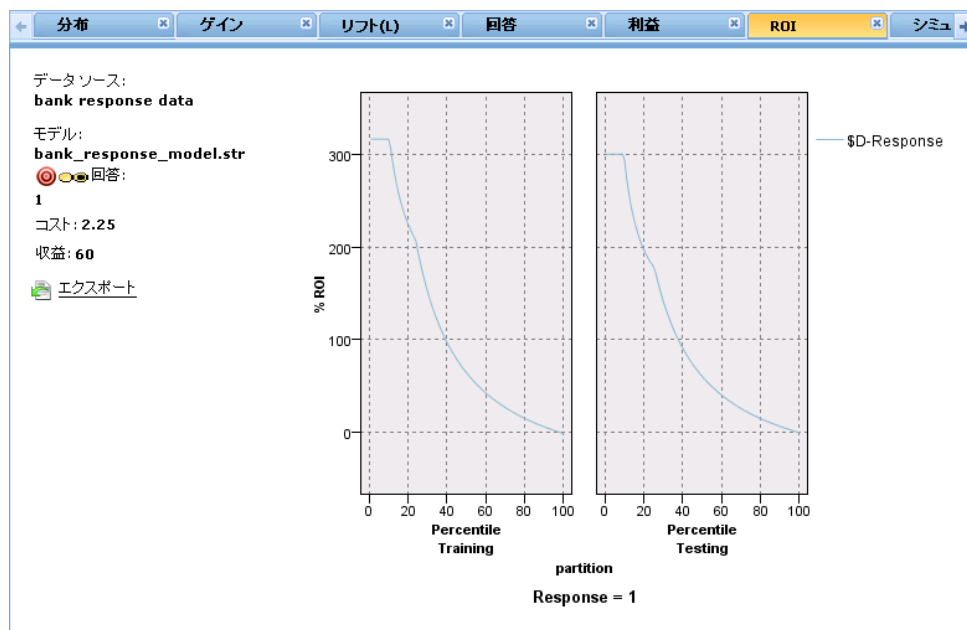
図 6-5  
利益グラフ



## ROI グラフ

ROI（投資収益率）グラフは、収益とコストを定義するという点で利益グラフに似ています。ROI グラフでは、各期間の利益とコストが比較されます。

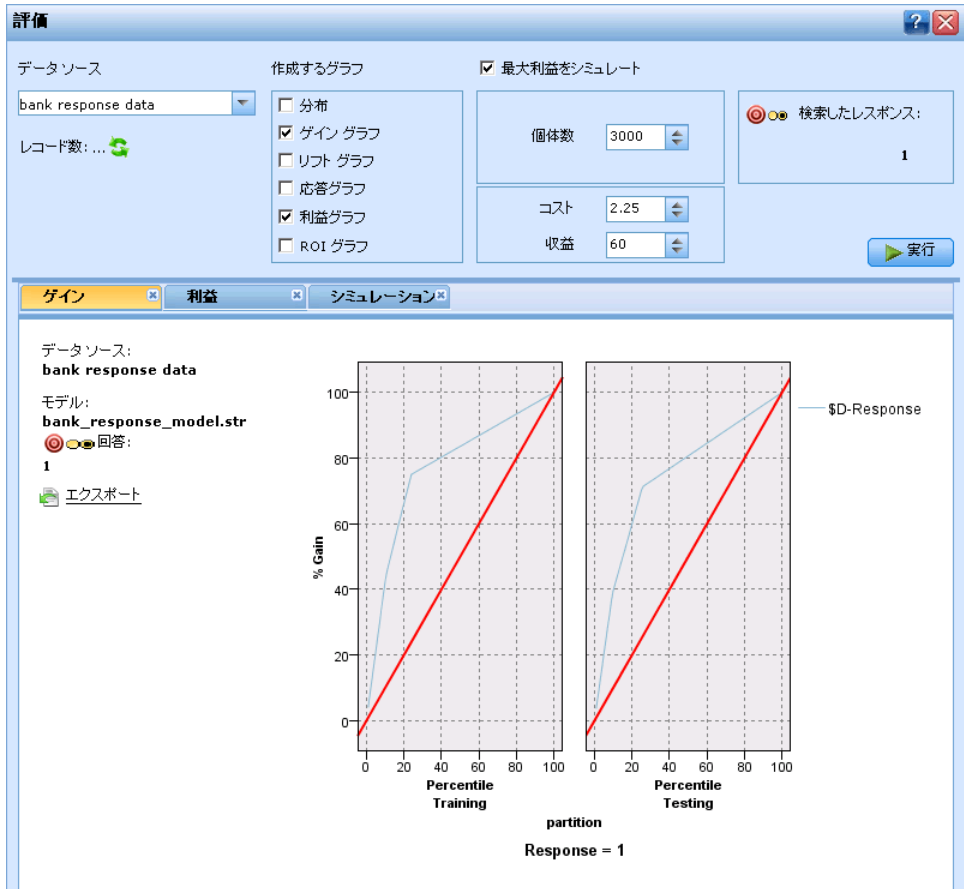
図 6-6  
ROI グラフ



## 最大利益をシミュレート

モデルで予測した回答率に基づいて、特定の母集団についての最大利益をシミュレートすることができます。そのためには、コストと収益の値、および母集団のサイズを指定する必要があります。

図 6-7  
シミュレーションの設定



最大利益を予測するには、以下の操作を行います。

- ▶ 予測のベースとなる母集団またはレコード数の合計を指定します。注：これは、モデルのスコアリング時に、自動的にデータソースまたはサブセット内のレコードの合計数に設定されます。
- ▶ 各レコードに関連付けるコストを指定します。たとえば、キャンペーン用の宣伝パンフレットの一人当たりの印刷費および送料などです。
- ▶ 見込みのある回答から生み出される予想収益を指定します。たとえば、キャンペーン内のあるアイテムの小売原価などです。
- ▶ 関心のある回答を選択します。
- ▶ [実行]をクリックします。

予測結果のグラフは、最大利益に関する最善の結果を示し、モデルで特定した回答傾向に基づいて、この結果を達成するためにターゲットとする必要のあるレコードのパーセンテージを特定します。 [詳細は、2章 p. 11 モデルの評価 を参照してください。](#)

ターゲット レコードのグラフの下にあるスライダー バーを 0 ~ 100 % の範囲で移動させると、それが利益にどのような影響を与えるかを確認できます。

数値が許容できる場合は、レコードのスコアリングにその設定を使用するように選択できます。これを行うには、[スコアリングに使用] をクリックして、[スコア] タブの [傾向オプション] 領域にある [上位 %] フィールドに目的のレコードのパーセンテージを表示します。 [詳細は、7章 p. 81 データベース テーブル、ファイル、または、Cognos BI サーバーのモデルのスコアリング を参照してください。](#)

### 利益を比較

選択したレコードが優れた選択であることを確認するために、これらのレコード、レコードから同じパーセンテージで抽出した無作為のグループ、およびレコード全体からそれぞれ見込める利益を比較することができます。

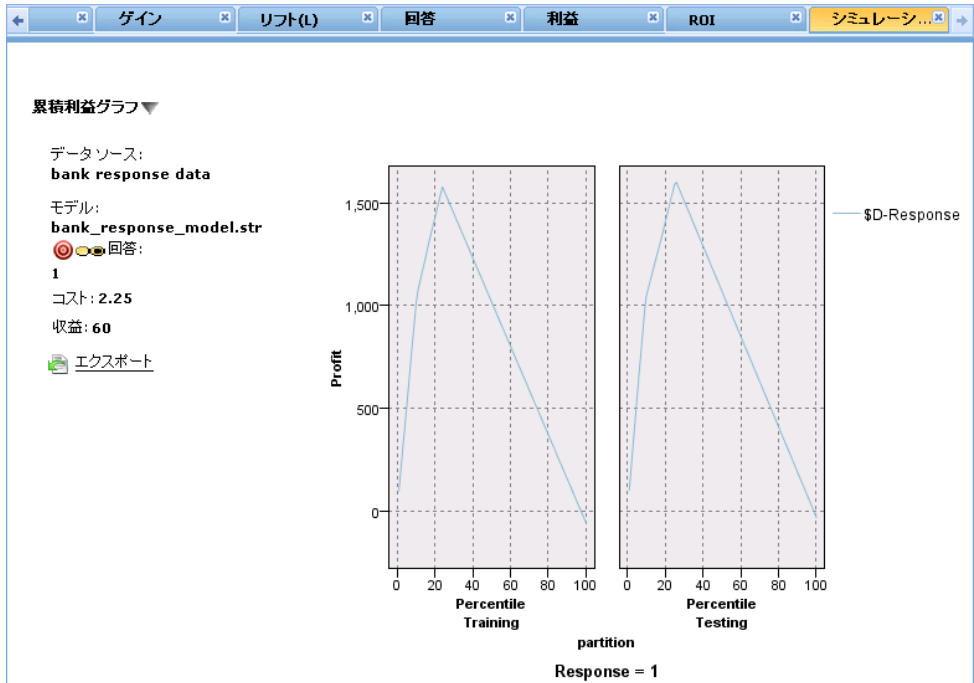
表示される詳細には、ターゲットとしたレコード数、予想される回答率、コスト、収入、総利益が含まれます。このHTML形式の情報は、このアプリケーションで使用させる為に、エクスポートすることができます。

### 利益グラフ

選択したレコードから、予測する累積/非累積利益グラフを表示できます。

注：このグラフは評価を初めて実行したときに生成されますが、利益シミュレーション セクションのスライダーを移動させて変更しても、自動的に更新されません。

図 6-8  
累積利益グラフ



## のテストmodel

[テスト] ダイアログボックスを使用すると、サンプル レコードのスコアリングを調べることができます。結果はテストを実行したコンテキストに特有であり、モデルまたはルールにより返されたスコアリング反映するかも知れません。たとえば、少数の顧客レコードについて、ダイレクト メール キャンペーンに応答する確率の高い顧客を調べることができます。

- ▶ [モデリング] タブで、**テスト** をクリックします。[テスト] ダイアログが表示されます。
- ▶ 使用するデータ ソースを選択し、[レコード] セクション内の **カスタム データの指定** を選択し、テスト レコードを入力します。
- ▶ 必要に応じて、使用する選択を指定します。たとえば、テストで特定のレコードについて、包含または除外を行います。既存のルールを検索するか、または新しいルールを作成します。詳細は、5 章 p.50 **選択ルールの定義** を参照してください。
- ▶ データ ソースを使用する場合は、返してほしいレコードの数を指定します。最初のNレコードを最大で100まで返したり、上位のデータから、或いは無作為のサンプル レコードから読み取る事が可能です。トランザク

ションデータを使用してアソシエーションモデルをテストする場合は、入力データはレコードレベルではなく、トランザクションレベルにてサンプルされます。

- ▶ カスタム データ セットを使用する場合は、**カスタム データの指定** を選択し、**カスタム レコードの追加** をクリックします。プロジェクト データ モデル内で定義されているすべてのフィールドがリストされるので、各フィールドに値を指定することができます。完了したら、**保存** をクリックします。
- ▶ **表示するフィールドの指定** の下で、テスト出力内に含めるフィールドを選択します。
- ▶ **[実行]** をクリックします。
- ▶ テストの要約には、各レコードへの成功した割り当てのリストが含まれます。
- ▶ 各行の先頭にある情報アイコンをクリックすると、その行の詳細を表示することができます。

## アプリケーション上でのモデル スコアの使用

モデルをスコアするとは、結果または決定への入力として使用できる予測を得るために、それを何らかのデータまたは母集団に適用することを意味します。たとえば、すべての顧客をデータベースにスコアリングしてダイレクト メール キャンペーンの見込みを識別するかもしれませんし、もしくは、顧客がコール センターにコンタクトした際にその都度スコアリングして彼らにどの製品またはサービスをオファーするかを決定するかも知れません。アプリケーションに応じて、スコアリングの結果をデータベース テーブルまたはフラット ファイルに書き込むことができ、または、セグメント、選択、および、アプリケーション上の決定を促すアクション ルールへの入力として使用することができます。

モデル スコアは以下の方法で使用することができます：

- ルールへの入力として。 [詳細は、5 章 p.47 ルールの作成 を参照してください。](#)
- バッチ スコアリングをサポートするアプリケーションには、モデルスコアをデータベース テーブルまたはフラットファイルに書き込むことができます。 [詳細は、7 章 p.81 データベース テーブル、ファイル、または、Cognos BI サーバーのモデルのスコアリング を参照してください。](#)

## モデル出力フィールド

表記法により、これらのフィールドの名前は対象フィールドをベースとし、接頭辞を付加します。たとえば、対象フィールド名が Responseであれば、出力フィールドの名前は \$XF-Response あるいは \$XFRP-Responseのようになります。以下に詳述する通り、具体的なフィールドの詳細は、モデルのタイプ、および選択された対象フィールドの測定レベルに依存します。詳細は、4 章 p.34 測定レベル を参照してください。

**フラグ型対象:**フラグ型対象の自動化されたモデルは、以下のフィールドを出力します。

- **\$XF-<target>**.各レコードの予測値。そのフィールドに定義された「真」および「偽」の項目で報告されます。
- **\$XFRP-<target>**.傾向スコア。各レコードの「真」値の尤度を示します。確信度スコアとは異なる傾向スコア。通常はそのレコードの予測項目として報告されます。たとえば、確信度の高い「偽」の予測は、応答しない高い尤度を示すので低い傾向と解釈されます。

**カテゴリ対象。**カテゴリ対象（名目または序数）のある自動化モデルは、以下のフィールドを出力します：

- **\$XS-<target>**.各レコードの予測値
- **\$XSC-<target>**.その予測に関連づけられた確信値

**連続型対象。**連続数値型対象の自動化されたモデルは、以下のフィールドを出力します。

- **\$XR-<target>**.各レコードの予測値
- **\$XRE-<target>**.その予測の標準誤差

**予測ルール(対話型)モデル。**予測ルールモデル（フラグ型またはカテゴリ型対象のみ）は以下のフィールドを出力します。

- **\$D-<target>**.各レコードの予測結果。望ましい応答を示すために使用されるフィールドのために定義された「真」の値、および、他の何らかの値を示す空白を伴います。
- **\$DP-<target>**.各レコードの希望する結果値の確率。モデル作成時に結果表のセグメントに表示される確率に一致します。詳細は、p.58 予測ルールモデルの作成 を参照してください。
- **\$DI-<target>**.レコードが属するセグメントを識別するインデックス番号（モデル結果の一番左の列に表示）
- **\$DRP-<target>**.希望する応答の尤度を示す傾向スコアフラグ型対象の場合のみ可。



# スコアリング モデル

## データベース テーブル、ファイル、または、Cognos BI サーバーのモデルのスコアリング

モデルのスコアリングとは、モデルをデータ サンプルまたは対象の母集団に適用することです。たとえば、既存の顧客を使用して、キャンペーンに最も反応する可能性のある顧客を予測するモデルを生成して、このモデルを使用して、現在は顧客ではないけれども販売促進の送付先を含めたい人々のレコードをスコアリングすることができます。

モデルのスコアリングにおける基本手順は次のとおりです。

- ▶ スコアリングされたデータの宛先を選択します。
- ▶ 出力する情報とその格納方法を選択します。
- ▶ データ ソースとそのデータのスコアリングの対象のサブセットを選択します。
- ▶ 格納するレコードを指定し、期待する結果のプレビューを表示します。
- ▶ **スコアリング** をクリックし、モデルに対して自らの設定を使用してスコアリングを実行します。

スコアは、データに追加した 1 つまたは複数のフィールドに書き込まれます。表記法により、これらのフィールドの名前は対象フィールドをベースとし、接頭辞を付加します。たとえば、対象フィールド名が Response であれば、出力フィールドの名前は \$XF-Response あるいは \$XFRP-Response のようになります。 [詳細は、 p. 86 出力フィールドの選択 を参照してください。](#)

## スコアリングの宛先の選択

スコアリング データをデータベース、ファイル、または、Cognos BI サーバーに出力することを選ぶことができます。サポートされるファイル タイプには、テキスト (\*.txt, \*.csv)、Excel、SAS で使用される \*.sav、IBM® SPSS® Statistics フォーマット、および、IBM® SPSS® Data Collection にサポートされるファイルがあります。 [詳細は、 p. 82 ファイルの宛先 を参照してください。](#)

## データベースの宛先

データベースにスコアリングの出力を送信すると、さらに詳細を入力するように求められます。

- ▶ **データベース名** を入力するか、または必要なデータベースを参照します。
- ▶ スコアの格納方法を指定します。新しいテーブルの作成を選択することもできます。その場合、そのテーブルの名前を入力するように要求されます。その代わりに、既存のテーブルに追加するか、または既存のデータソースに上書きすることもできます。この場合は、関連するテーブルまたはデータソースを選択するように要求されます。

既存のデータベースに付加したり、上書きする場合は、スコアリングフィールドを既存の宛先内のフィールドにマッピングすることができます。詳細は、[p. 85 スコアリング フィールドをマッピング](#) を参照してください。

## ファイルの宛先

スコアはいくつか異なるファイルタイプに出力することができます。タイプごとに、既存のファイルを参照するか、スコアリングデータを追加する新規ファイルを作成します。

既存のファイルに付加したり、上書きする場合は、スコアリングフィールドを既存の宛先内のフィールドにマッピングすることができます。詳細は、[p. 85 スコアリング フィールドをマッピング](#) を参照してください。

### フラット ファイル

スコアリングの出力をフラットファイルに送信することを選択した場合は、さらに詳細を入力してファイルの内容を制御することができます。

**保存モード:**[上書き] を選択すると、指定したファイル中の既存のデータが上書きされます。[レコード追加] を選択すると、このノードからの出力データが既存ファイルの末尾に追加され、既存データはそのまま保存されます。

- **フィールド名を包含する:** このオプションを選択すると、出力ファイルの1行目にフィールド名が書き込まれます。このオプションは、保存モードで [上書き] を選択した場合にだけ利用できます。

**各レコードの後に改行を入れる:** このオプションを選択すると、各レコードが出力ファイル中の新しい行に書き込まれます。

**フィールド区切り文字:** 生成するテキストファイルで、フィールド値の間に挿入する文字列を選択します。オプションはカンマ、タブ、スペース、および [その他] です。[その他] を選択した場合は、テキストボックスに適切な区切り文字を入力してください。

**シンボル値の引用符:** シンボル値フィールドの値に対して使用する引用符の種類を選択します。オプションは[なし] (値は単一引用符で囲まれていない)、シングル ('), **ダブル** ("), および[その他]です。[その他]を選択した場合は、テキスト ボックスに適切な引用文字を入力してください。

**文字コード:** 使用するテキストのエンコード方法を指定します。システムデフォルトまたは UTF-8 から選択できます。

**小数点記号:** データ中で小数点記号をどのように表すかを指定します。

- **ピリオド (.):** 小数点区切り文字として、ピリオドを使用します。
- **カンマ (,):** 小数点区切り文字として、カンマを使用します。

### Excel ファイル

**データ型:** エクスポートする Excel ファイルの形式を選択します。

**フィールド名を包含する:** フィールド名をワークシートの最初の行に表示するかどうかを指定します。

### SAS ファイル

**データ型:** 作成する SAS ファイルの形式を選択します。次の 3 種類の SAS ファイル形式を選択することができます。SAS for Windows/OS2 (\*.sd2)、SAS for UNIX (\*.ssd)、または SAS バージョン 7/8/9 (\*.sas7bdat)。

**フィールド名をエクスポート:** SAS で使用するために、フィールド名とラベルをエクスポートするオプションを選択します。

- **名前と変数ラベル:** フィールド名とフィールド ラベルの両方をエクスポートする場合に選択します。名前は SAS の変数名としてエクスポートされ、ラベルは SAS の変数ラベルとしてエクスポートされます。
- **変数ラベルとして命名:** フィールド名を SAS で変数ラベルとして使用する場合に選択します。SAS で無効な名前が作成されることを防止するには、代わりに **名前と変数ラベル** を選択します。

### IBM SPSS Statistics ファイル

**フィールド名をエクスポート:** IBM® SPSS® Statistics .sav ファイルへのエクスポート時に変数名とラベルを処理する方法を指定します。個々の名前は、自動的に無効な文字を置き換えて修正されることに注意してください。

- **名前と変数ラベル:** フィールド名とフィールド ラベルの両方をエクスポートする場合に選択します。名前は SPSS Statistics の変数名としてエクスポートされ、ラベルは SPSS Statistics の変数ラベルとしてエクスポートされます。
- **ラベルとしての名前:** フィールド名を SPSS Statistics で変数ラベルとして使用する場合に選択します。SPSS Statistics で無効な名前が作成されることを防止するには、代わりに **名前と変数ラベル** を選択します。

## Cognos BI サーバーの宛先

Cognos BI サーバーにスコアリングの出力を送信すると、さらに詳細を入力するように求められます。

注：エクスポートできるのはリレーショナル データのみで、OLAPデータはできません。

Cognos BI にデータをエクスポートするには、以下を指定する必要があります：

- **Cognos 接続。** Cognos BI サーバーへの接続。
- **ODBC 接続。** Cognos BI サーバーが使用する Cognos データ サーバーとの接続。

これらの接続は同じデータベースを指定し、接続に同じユーザー名を使用しなければなりません。また、ODBC 接続する Cognos サーバーのパスワードは ODBC の詳細と同じでなければなりません。

データサーバーには実際のデータをエクスポートし、Cognos BI サーバーにはパッケージ メタデータをエクスポートします。

- ▶ Cognos BI サーバー[宛先]を選択して**接続** をクリックします。サーバー接続の詳細の入力を求めるメッセージが表示されます。
- ▶ データをインポートまたはエクスポートする IBM Cognos サーバーの**サーバー URL**をキーインします。使用するURLが不確かな場合は Cognos システム管理者に問い合わせてください。
- ▶ 接続する**モード**を選択してください。特定のユーザー（たとえば、管理者）としてログインするには、**資格証明の設定**を選択して、IBM Cognos の名前空間、ユーザー名、および、パスワードを入力してください。
  - サーバーへのログインに使用する、IBM Cognos セキュリティ認証プロバイダ**名前空間**を指定します。認証プロバイダは、ユーザー、グループ、および、役割りの定義と保守や、認証プロセスの管理に使用されます。
  - サーバーログオンに使用する、IBM Cognos **ユーザー名**を入力します。
  - 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。
- ▶ または、ユーザー資格証明なしでログインする場合は、名前空間、ユーザー名、および、パスワード フィールドに入力することはできませんが、**匿名接続の使用**を選択します。

注：一部のサーバー接続には匿名接続は使用できません。

- ▶ **データソース**を選択します。
- ▶ エクスポート パッケージを作成するCognos BI サーバーの**フォルダ**のパスと名前をキーインします。あるいは、必要なものを表示させることもできます。

- ▶ エクスポートされるメタデータを含むパッケージの**パッケージ名**をキーインします。これは新しいパッケージでなければなりません。既存のパッケージはエクスポートできません。 [詳細は、4章 p.41 Cognos オブジェクト詳細の選択](#) を参照してください。
- ▶ ODBC宛先を選択するには、**データベース名**を入力するか、必要なものを表示させます。
- ▶ スコアの格納方法を指定します。新しいテーブルの作成を選択することもできます。その場合、そのテーブルの名前を入力するように要求されます。その代わりに、既存のテーブルに追加するか、または既存のデータソースに上書きすることもできます。この場合は、関連するテーブルまたはデータ ソースを選択するように要求されます。

既存のデータベースに付加したり、上書きする場合は、スコアリングフィールドを既存の宛先内のフィールドにマッピングすることができます。 [詳細は、p.85 スコアリング フィールドをマッピング](#) を参照してください。

## スコアリング フィールドをマッピング

既存のデータベース テーブルまたはファイルにスコア付加したり、上書きする場合は、スコアリング フィールドを既存のテーブルまたはファイル内のフィールドにマッピングすることができます。出力およびマッピングした関連フィールドでは、同じデータ タイプが使用されている必要があることに注意してください。

- スコアを既存のデータベース テーブルに書き込むときは、新しいスコアを上書きする場合も追加する場合も、スコアリングが正しく処理されるようにすべてのフィールドが既存のフィールドにマッピングされる必要があります。必要であれば、**フィールドをマップ**をクリックして、存続するすべてのフィールド用の出力を指定します。
- 既存のファイルへ書き込みの際は、マッピングが必要なのは、出力ファイルのオプションで**追加**が選択されている場合だけです。既存のファイルを上書きする場合は、新規のファイルが既存のファイルを置き換えるので、マッピングは必要ありません。
- ▶ **フィールドをマッピング** をクリックします。ダイアログが開き、システムが最適なマッピング フィールドを推測します。任意の時点で、**リセット** をクリックすることにより、システムの最適な推測に戻ることができます。
- ▶ 利用可能な出力宛先から選択して、**マッピング** をクリックし、それらを必要なスコアリング フィールドにマッピングします。
- ▶ 終了したら、**保存** をクリックします。

## 出力フィールドの選択

[スコア] タブで、[出力フィールドを指定] を選択して、スコアリング出力を含めるフィールドを選択します。通常は、実際のスコアに加えて、各レコードを識別できるようなフィールド、たとえば顧客ID フィールドのようなフィールドが少なくとも 1 つは必要でしょう。必要に応じて、モデルで使用している入力フィールドの全部または一部を含めることもできます。

### モデル出力フィールド

表記法により、これらのフィールドの名前は対象フィールドをベースとし、接頭辞を付加します。たとえば、対象フィールド名が Response であれば、出力フィールドの名前は \$XF-Response あるいは \$XFRP-Response のようになります。以下に詳述する通り、具体的なフィールドの詳細は、モデルのタイプ、および選択された対象フィールドの測定レベルに依存します。詳細は、4 章 p. 34 測定レベル を参照してください。

**フラグ型対象:**フラグ型対象の自動化されたモデルは、以下のフィールドを出力します。

- **\$XF-<target>**.各レコードの予測値。そのフィールドに定義された「真」および「偽」の項目で報告されます。
- **\$XFRP-<target>**.傾向スコア。各レコードの「真」値の尤度を示します。確信度スコアとは異なる傾向スコア。通常はそのレコードの予測項目として報告されます。たとえば、確信度の高い「偽」の予測は、応答しない高い尤度を示すので低い傾向と解釈されます。

**カテゴリ対象。**カテゴリ対象（名目または序数）のある自動化モデルは、以下のフィールドを出力します：

- **\$XS-<target>**.各レコードの予測値
- **\$XSC-<target>**.その予測に関連づけられた確信値

**連続型対象。**連続数値型対象の自動化されたモデルは、以下のフィールドを出力します。

- **\$XR-<target>**.各レコードの予測値
- **\$XRE-<target>**.その予測の標準誤差

**予測ルール(対話型)モデル。**予測ルールモデル（フラグ型またはカテゴリ型対象のみ）は以下のフィールドを出力します。

- **\$D-<target>**.各レコードの予測結果。望ましい応答を示すために使用されるフィールドのために定義された「真」の値、および、他の何らかの値を示す空白を伴います。

- **\$DP-<target>**.各レコードの希望する結果値の確率。モデル作成時に結果表のセグメントに表示される確率に一致します。詳細は、6章 p.58 [予測ルールモデルの作成](#) を参照してください。
- **\$DI-<target>**.レコードが属するセグメントを識別するインデックス番号（モデル結果の一番左の列に表示）
- **\$DRP-<target>**.希望する応答の尤度を示す傾向スコアフラグ型対象の場合のみ可。

## データおよびサブセットの選択

[スコアリング対象のデータセット] リストから、必要なデータ ソースを選択します。

### データ サブセット

データ ソース全体のスコアリングを行いたくない場合は、既存のルールを使用するか、このモデルに対してのみ使用する新しいルールを作成し、データのサブセットを選択することができます。詳細は、5章 p.50 [選択ルールの定義](#) を参照してください。

## スコア対象のレコードの選択

傾向スコアを返すモデルでは、これらの値を使って、スコア対象のレコードを選択できます。たとえば、顧客の 30% を高い応答傾向を持つとスコアリングし、これを使ってマーキング リストを作成することができます。傾向スコアは、対象に関する「真」値の尤度を示すもので、フラグ型対象の予測モデル、および予測ルール（インタラクティブ）モデルに使用することができます。（予測ルールモデルの場合は、選択された応答は「真」値を、その他はすべて「偽」を示します。その他のタイプのモデルでは、これらのオプションは使用できません）。

傾向に基づいてレコードを選択するには、次の 3 つの方法があります。

- **上位のパーセンテージ:** デフォルトでは、これにより上位 10 パーセントが選択されます。このオプションでは、利益に基づいてスコアリング結果をシミュレートすることができます。詳細は、6章 p.75 [最大利益をシミュレート](#) を参照してください。
- **上位レコードの数:** デフォルトでは、これにより上位 1000 レコードが選択されます（存在する場合）。
- **傾向の最大値と最小値を指定:** たとえば、75 ~ 100 % の応答傾向があるすべてのレコードです。

## モデルのスコアリング

スコアリングの詳細を指定した後、モデルのスコアのサンプルをプレビューすることができます。これは大きなデータセットを持っている場合に非常に役に立ちます。つまり、データセット全体のスコアリングが終わることを待つことなく、スコアリングの結果が期待通りに表示されるか簡単にチェックできるからです。

注：データが無効または不完全と判断されると、プレビューを試みた段階で警告が表示されます。

プレビューには、データが、行にレコードを、列に変数を示すテーブルとして表示されます。

### スコアリング

スコアリングの対象となるレコード数を確定するには、画面下部にある[更新]アイコンをクリックします。これにより、データソース内のレコード合計数と指定したオプション設定に基づいて、モデルの作成で使用されるレコード数が表示されます。

[スコアリング] をクリックすると、タブの下部に進捗バーが表示されます。表示中、いつでもスコアリングを中止することができます。

注：データが無効または不完全と判断されると、スコアリングを試みた段階で警告が表示されます。

データのスコアリングが正常に終了すると、結果を格納した場所（ユーザーが指定したデータベースまたはファイル）を知らせるメッセージが表示されます。



# アプリケーションの管理

管理者は、ビジネス ユーザーが IBM® Analytical Decision Management アプリケーションを操作する方法を制御することができます。このレベルの制御によりビジネス ユーザーのジョブが簡素化されるため、ビジネス ユーザーはビジネス上の課題を解決するためにアプリケーションを一層迅速かつ効率的に使用できるようになります。また管理者は特定の機能へのアクセス制限を制御することもできます。

IBM® SPSS® Collaboration and Deployment Services Deployment Manager を使用し、必要に応じて特定のユーザーまたはグループに管理者権限を許可します。Decision Management の管理者 アクションを付与されたユーザーだけが、このセクションで説明する各種機能にアクセスできます。

- ▶ Deployment Manager を開きます。
- ▶ [ツール] メニューから **サーバー管理** を選択します。
- ▶ [サーバー管理] タブで、サーバーに接続します。
- ▶ **ユーザーおよびグループ** の下の、**ローカル ユーザー リポジトリ** を開きます。
- ▶ **新規グループ** をクリックし、管理者の**新規グループ**と**ビジネス ユーザー**の**新規グループ**を作成します。例：
  - Decision Management の管理者
  - Decision Management のユーザー：

この時点で必要に応じて、新規ユーザーを作成することもできます。グループの作成時に、各グループに適切なユーザーを追加します。例えば、管理者グループのメンバーは、組織のIBM Analytical Decision Management アプリケーションを管理する権限を付与されます。
- ▶ [ロール] に移動し、**新規ロール** をクリックして、**新規管理者ロール**と**新規ビジネス ユーザー ロール**を作成します。例：
  - IBM Analytical Decision Management 管理者
  - IBM Analytical Decision Management ユーザー。

新規ロールの作成時に、少なくとも以下のアクションをロールに割り当てます。

テーブル 8-1  
ロール アクション

役割	必須のアクション
Decision Management 管理者	コンテンツおよびフォルダへのアクセス Decision Management の管理者 レポートを動的に実行 モデルのスコアリング 作業を送信 全部のバージョンを表示
Decision Management ユーザー	コンテンツおよびフォルダへのアクセス レポートを動的に実行 モデルのスコアリング 作業を送信 最新の*ORを表示 全部のバージョンを表示

\*注：ユーザーが自分で作ったプロジェクト バージョン以外にアクセスする必要がある場合には、**Show latest** または **Show All Versions** の何れかのアクションが必要です。さもなければ、現在のユーザーに最新版のアクセス権がない場合には、プロジェクトは読み取り専用モードで開かれます。

- ▶ Decision Management の管理者ロールを選択し、**ユーザーおよびグループの編集** をクリックします。直前で作成したDecision Management の管理者グループを追加します。
- ▶ Decision Management のユーザー ロールを選択し、**ユーザーおよびグループの編集** をクリックします。直前で作成したDecision Management のユーザーを追加します。

## 一般的なオプション

管理者は、IBM® Analytical Decision Management ユーザー インタフェースのすべてのタブで、特定の一般的なオプションを使用することができます。これらのオプションを使用すれば、管理者はビジネス ユーザーに対してタブを非表示にしたり、ロックすることができます。

- 各タブの先頭には、**非管理者のタブの<タブ名>を非表示にする**というオプションがあります。ここで、<タブ名> はタブの名前です。このオプションを使用すると、タブ全体がビジネスユーザーには非表示になります。例えば、ユーザーに対してアプリケーションの [データ] タブを表示させる場合は、アプリケーションにログオンし、[データ] タブに移動して、**非管理者からの [データ] タブを非表示にする** を選択します。
- 各タブの先頭には、**すべての<タブ名> オプションをロックする**というオプションがあります。ここで、<タブ名> はタブの名前です。このオプションを使用すると、タブ上のすべてのコントロールをロックできるので、ビジネス ユーザーには読み取り専用になります。例えば、ビジネスユーザーに対して、現在のデータ設定の表示を許可し、変更は禁止す

る場合は、アプリケーションの [データ] タブ上のこのオプションを使用します。この設定は、ビジネス ユーザーのジョブを簡素化させる場合に役立ちます。

- 各タブの個々の設定に南京錠アイコンが用意されている場合は、管理者はそれをクリックして個々の設定をロックまたはロック解除することができます。例えば、[定義] タブのアプリケーション内において、管理者は次元階層ツリーをロックすることができ、ビジネスユーザーは変更できません。南京錠アイコンをクリックすると、それに関連付けられたコントロールをロックまたはロック解除します。管理者がロックしたコントロールに対応するロックされた南京錠アイコンは、ビジネス ユーザーに対して表示されますが、ロック解除された南京錠アイコンは表示されません。

ロックされた個々の項目は、淡色表示の背景またはくすんだ色で表示されて、選択できないことが示されます。ユーザーが選択を試みると、エラー メッセージが表示されます。タブ全体をロックすると、タブの先頭に、ロックされていることを示すメッセージが表示されます。



## アプリケーション ショートカットの制御

アプリケーション起動ページには、ビジネス ユーザーがアプリケーションを開いて、特定のラベル付きバージョンのアプリケーションに対するカスタム ショートカットを追加するための簡単な方法が用意されています。管理者は、すべてのユーザーに対して、起動ページで表示するアプリケーション ショートカットを指定することができます。管理者権限でログオンした後、以下の操作を行います。

- ▶ すべてのユーザーに対して、起動ページで表示するアプリケーションの **デフォルト時のランチャー** を選択します。
- ▶ 必要に応じてアプリケーション ショートカットの南京錠アイコンをクリックし、ビジネス ユーザーが起動ページから削除できないようにします。この場合でも、他のアプリケーション ショートカットの追加または削除はビジネス ユーザーに許されていることに注意してください。ロックされたショートカットは、ユーザーに対して、閉じたアイコンではなく、ロックされた南京錠アイコンとして表示されます。 [詳細は、3 章 p.21 アプリケーションの起動](#) を参照してください。

## ホーム ページを非表示にする

ホーム ページには、各アプリケーションのステップバイステップのワークフローがグラフィックで表示されます。グラフィックをクリックすると、ユーザー インタフェースの当該セクションにジャンプします。ただし、ビ

ビジネス ユーザーの場合は、アプリケーション ワークフローのすべてのステップへのアクセスが必要なわけではありません。

管理者は、ホーム ページをビジネス ユーザーには非表示にすることができます。例えば、特定のアプリケーションでビジネス ユーザーが使用する可能性があるのは、[レポート] タブのみだとします。この場合、管理者はホーム ページと他のすべてのタブをビジネス ユーザーには非表示にすることができます。そうすることで、ビジネス ユーザーはログオンすると、直接、[レポート] タブに導かれ、他のタブが表示されることはありません。

- ▶ ホーム ページで、**アプリケーション ホーム ページを非表示にする** を選択します。
- ▶ ドロップダウンで、ユーザーがアプリケーションを初めて開いたときに表示するタブを指定します。

ここで選択したタブをその後非表示にした場合は、アプリケーションには次の非表示ではないタブが表示されることに注意してください。

## データ オプションのロック

### データ

[データ] タブを使用すると、アプリケーションが、分析、シミュレーションおよびテスト、スコアリング、またはその他の運用目的で使用するデータ セットを定義することができます。これらのデータ ソースは、サーバーに相対的に定義されます。したがって、データ ファイルや odbc ソース等へのパスは、すべてサーバーに相対的になります。

管理者は、ビジネス ユーザーがアプリケーションを使用する前に、データの一部またはすべてを準備することができます。管理者は、以下の操作を行うことができます。

- **新規データ ソースを作成し、作成後にロックする場合は、データソースのロック** を選択します。
- テーブル内または [データ ソース エディタ] ダイアログ内でデータソースの横にある南京錠アイコンをクリックして、個々のデータ ソースをロックまたはロック解除します。
- **プロジェクト データ モデル** ドロップダウンでデータ モデルを選択し、南京錠アイコンをクリックすることにより、アプリケーションで使用するデータ モデルを設定し、ロックします。

ビジネス ユーザーは [プロジェクト データセット] セクションから [マイ データセット] セクションに、ロックされたデータ ソースをコピーすることができますが、その場合は、両方のセクションでロックされた状態

のままであることを注意してください。詳細は、4章 p.30 データソースの管理 を参照してください。

## スコアリング宛先オプションのロック

[展開] タブまたは [スコア] タブの [直ちにスコアリング] セクションを使用するように設定されたアプリケーションには、ビジネス ユーザーにスコア データの格納先（データベースまたはファイル）を指定させるセクションがあります。例えば、スコアをデータベースに格納する場合は、データベース名とスコアの格納方法（例えば、データベース内に新規テーブルを作成する）を指定します。

管理者は、スコアリングの宛先オプションをロックすることができます。これを使用すると、例えば、ビジネス ユーザーに対して、スコアを特定のデータベースの既存のテーブルに追加することのみを許可し、既存のデータへの上書き、既存のテーブルに新規データとして追加、新規テーブルの作成、あるいはファイルへの格納は禁止することができます。

- ▶ すべてのスコアリング宛先オプションをロックするには、[展開] タブの [宛先] ドロップダウンの横にある南京錠アイコンをクリックします。ビジネス ユーザーは、宛先オプションの表示はできますが、編集はできません。

## 外部ルールの作成に使用するプロジェクト メタデータをダウンロードします

ILOG などの Business Rules Management System で作成されたルールは、現在の IBM® Analytical Decision Management プロジェクトに使用されるのと同じデータモデルをサポートするように開発されていれば、IBM Analytical Decision Management アプリケーションで参照および使用することができ、アプリケーション使用のためのウェブ サービスとして展開することができます。この作業を達成するために、IBM Analytical Decision Management 管理者は、現行プロジェクトのメタデータが入った .ZIP ファイルをダウンロードできます。ダウンロードするファイルは、管理者がダウンロードを開始するために使用するシステムに、ローカルで保存することができます。

図 8-1  
メタデータ アイコンをダウンロードします



以下のファイルは、ZIP ファイルに入っています：

- **XML schema definition (\*.XSD)**. 現行の IBM Analytical Decision Management プロジェクトのデータ モデル中のフィールド タイプの定義が含まれています。このファイルは、ILOG Rules Studio などの外部開発ツールにインポートして、現行のプロジェクトに使用するルールを開発することができます。
- **ウェブ サービス記述言語テンプレート (\*.WSDL)**。データ モデルを使用するウェブ サービスの開発支援用として含まれます。.WSDL ファイルは、IBM Analytical Decision Management アプリケーションと外部ルール サービスの間のメッセージをメディエートする役割りのウェブ サービスを構築するために使用できるテンプレートとして提供されます。ウェブ サービス メディエートの詳細は、インテグレーターによりマニュアルとして提供されます。

IBM Analytical Decision Management にさまざまなタイプを返す ILOG ルール（選択に使用するブーリアン、たとえばルールの集計に使用する数値、および文字列）を作成し使用することができることにご注意ください。ただし、文字列 dataresponse ルール以外のものを作成するには、XML スキーマ定義 (.XSD) はマニュアルで編集する必要があります。

同様に、ルール プロジェクトには各々異なるタイプを返す複数ルールを持たせることができます。このことをサポートするためには、複数バージョンの .XSD を維持する必要があります（たとえば、dataresponse 変数が数値に設定されたバージョン、文字列の別バージョン、ブーリアンの別バージョン、など）。

IBM Analytical Decision Management からダウンロードされた .XSD ファイルは2つの複雑なデータ タイプを持ちます：1つは ILOG ルール プロジェクトの入力パラメータ用に使用する IBM Analytical Decision Management プロジェクト データ モデルを表し、もう1つは ILOG ルール プロジェクトの出力パラメータ用に使用するタイプを返すルールを表します。返されるタイプ（その名前には必ず” response” の接尾辞がつきます）は文字列がデフォルトとなる数値要素を定義します。異なるタイプの戻り値のルール プロジェクトを作成する場合は、ILOG ルール プロジェクトをインポートする前に .XSD をマニュアルでダウンロードして編集し、返されるタイプの数値要素が要求通りに定義されるようにします。以下に例を挙げます：

```
<xsd:complexType name="claim_dataresponse">
  <xsd:annotation>
    <xsd:appinfo>
      <dmname>claim_dataresponse</dmname>
    </xsd:appinfo>
  </xsd:annotation>
  <xsd:sequence>
    <xsd:element name="value" type="xsd:string"/>
  </xsd:sequence>
</xsd:complexType>
```

```
I    <xsd:complexType name="claim_dataresponse">
      <xsd:annotation>
        <xsd:appinfo>
          <dmname>claim_dataresponse</dmname>
        </xsd:appinfo>
      </xsd:annotation>
      <xsd:sequence>
        <xsd:element name="value" type="xsd:boolean"/>
      </xsd:sequence>
    </xsd:complexType>

I    <xsd:complexType name="claim_dataresponse">
      <xsd:annotation>
        <xsd:appinfo>
          <dmname>claim_dataresponse</dmname>
        </xsd:appinfo>
      </xsd:annotation>
      <xsd:sequence>
        <xsd:element name="value" type="xsd:長整数型"/>
      </xsd:sequence>
    </xsd:complexType>
```

IBM Analytical Decision Management に使用する外部ルールの作成に関する情報に関しては、Application Designers Guide をご覧ください。

# ヘルプの利用

3 つのレベルのユーザー支援が提供されます。

- IBM® Analytical Decision Management でパッケージ化されている各アプリケーションでは、カスタム ヘルプが用意されています。ヘルプを起動するには、任意の画面またはダイアログでヘルプ アイコンをクリックします。



- コーチ テキスト: 小さなポップアップ ウィンドウ内に表示され、現在のアプリケーションまたは業務上の課題に関する情報を示します。コーチテキストは、状況依存のカスタム ヘルプの追加レイヤーとして、特定のサイトやアプリケーションに簡単に設定することができます。コーチ テキストは、ユーザー インタフェースの任意の時点で、小さなコーチ テキスト アイコンをクリックすることにより表示することができます。



- 標準のツール ヒント: ページまたはダイアログ内の個々のコントロールに用意されています。ツール ヒントを表示するには、ボタン、リンクその他のコントロールにカーソルを合わせます。

## 利用可能なドキュメント

それぞれの製品またはアプリケーションのインストールディスクにドキュメントが含まれているか、電子ダウンロードファイルによって提供されます。また、ドキュメントはIBM® SPSS® Collaboration and Deployment Services インストール ディレクトリにもインストールされています (たとえば、C:\Program Files\IBM\SPSS\Deployment\5.0\help\en\DecisionManagement)。

PDF 度ドキュメントはウェブ上の

<http://www-01.ibm.com/support/docview.wss?uid=swg27024118>

でも利用可能で、リリースノートは

<http://www-01.ibm.com/support/docview.wss?uid=swg27024117>で利

用可能です

全文書のオンラインで、集中型の集合については、[IBM Analytical Decision Management 7 Information Center](#) を参照してください。次の表には、Information Center の各提出物へのリンクが記載されています。



テーブル 9-1  
IBM Analytical Decision Management のドキュメント

ドキュメント	説明
Windows 向けインストール ガイド UNIX 向けインストール ガイド	製品のインストール、起動、アンインストールに関する指示
アプリケーション デザイナーズ ガイド	アプリケーションの設計と設定に関する指示。完全な XML スキーマ リファレンス、およびユーザー インターフェイスのカスタマイズに関する詳細を含む。
アプリケーション ユーザーズ ガイド	ビジネス ユーザー向けのユーザー インターフェイスに関する一般情報が含まれています。
IBM SPSS Modeler Advantage	IBM® SPSS® Modeler Advantage ビジネス ユーザーに予測モデル作成という武器を与えます。予測モデルを使用すると過去に発生した事象からパターンを特定できるので、それらのパターンを使用すれば将来発生する可能性がある事象を予測することができます。
IBM Analytical Decision Management for Customer Interactions	IBM® Analytical Decision Management for Customer Interactions は、企業に連絡してきた顧客に対してオファーするプロモーションを判断し、提案をコールセンター、Web サイト、または店内にリアルタイムで伝えます。アプリケーションは、ビジネス ルールのロジックを予測モデルを通じて取得する洞察と組み合わせることで、各顧客について、最も収益性の高い判断を特定します。
IBM Analytical Decision Management for Campaign Optimization	IBM® Analytical Decision Management for Campaign Optimization は、IBM Analytical Decision Management for Customer Interactions に類似しています。しかし、IBM Analytical Decision Management for Campaign Optimization は、最適化を使用し、予算やチャンネル容量および連絡先ポリシーなどの制約内で、個々の顧客に最適なオファーを決定することで、キャンペーン投資へのリターンを最大化する付加価値があります。

ドキュメント	説明
IBM Analytical Decision Management for Claims	IBM® Analytical Decision Management for Claims ビジネス ルールのロジックを予測モデルから得られる洞察と組み合わせることで、クレームをリアルタイムで処理し、最善のアクションを特定します。たとえば、クレームに対して、迅速な支払いのための「迅速処理」、通常処理のための「標準」、または専門調査部門への問い合わせのための「問い合わせ」を設定します。
IBM SPSS Rules Management	IBM® SPSS® Rules Management は、共有ルールの作成および編集のための中心的なツールを提供します。このツールは、アプリケーションでレコードの選択および処理を行うため、および判断を自動化するために使用できます。

# ストリームの共有

IBM® Analytical Decision Managementで作成されたモデルとプロジェクトは、ストリーム ファイルとして保存され、必要であればIBM® SPSS® Modelerで変更と使用ができます。例えば、IBM® SPSS® Modeler Advantageまたは他のIBM Analytical Decision Managementアプリケーションを使用してモデルまたはプロジェクトを作成し、IBM Analytical Decision Managementを通して展開する前にSPSS Modelerの中でストリームをカスタマイズしたいとします。代わりに、SPSS Modelerで作成されたストリームは、展開の設定が以下に記述されたように指定されていれば、IBM Analytical Decision Managementで開くことができます。

## IBM Analytical Decision Managementでストリームを開く

IBM Analytical Decision ManagementまたはIBM SPSS Modeler Advantageでストリームを開くときは：

- 展開タイプが指定されていない場合（なし）は、ストリームを開くことができますが使えるのは[データ]タブだけです。
- 展開タイプがスコアリングのみの場合は、[スコア]タブも使用できるかも知れませんが、スコアリング枝がIBM SPSS Modeler AdvantageまたはIBM Analytical Decision Managementにサポートされたタイプの1つのデータ ソース ノードから始まるのが条件となります（マージされたデータ セットがないこと）。
- 展開タイプがモデルのリフレッシュの場合は、[モデリング]タブも使用できるかも知れませんが、スコアリング枝がIBM SPSS Modeler AdvantageまたはIBM Analytical Decision Managementにサポートされたタイプの1つのデータ ソース ノードから始まり、選択されたモデル作成ノードが自動モデル作成ノード（予測モデル製作用にサポートされます）またはディシジョン リスト ノード（予測ルールモデル製作用にサポートされます）のどちらかであることが条件となります。さらに、ターゲット フィールドやモデリングに使用される他の情報を定義する、データ型ノードが存在しなければなりません。

## IBM SPSS Modelerに展開の設定を指定

SPSS Modelerでは、展開設定は[ストリームのプロパティ]ダイアログ ボックスに、以下のように指定します：

- ▶ SPSS Modeler メニューから次の項目を選択します。  
[ツール] > [ストリームのプロパティ] > 展開

- ▶ 展開のタイプリストから、スコアリングのみまたはモデルリフレッシュを適宜選択します。
- ▶ 必要に応じてスコアリングノード、モデリングノード、およびモデルノードを選択します。

[チェック]を選択して必要なオプションがすべて指定されていることを確認するか、または、[保存]を選択してストリームをIBM® SPSS® Collaboration and Deployment Services Repositoryに保存します。

IBM SPSS Modeler AdvantageまたはIBM Analytical Decision Managementに作成されたストリームには、展開情報は自動的に設定されます。

### IBM SPSS ModelerストリームをIBM Analytical Decision Managementで使用するためのヒント

- IBM SPSS Modeler AdvantageまたはIBM Analytical Decision Managementを使用してストリームを作成しその後SPSS Modelerで変更する方が、SPSS Modelerを使用して最初から構築するより簡単かも知れません。この方法で作成されたストリームは、IBM Analytical Decision Managementで完全に機能することが可能になる方法で構造化されます。
- IBM SPSS Modeler Advantageで作成されたすべてのストリームで、スーパーノードが枝の中の特定の場所に含まれており、エキスパートユーザーにはカスタマイズすることができます。ストリームのカスタマイズまたは追加はいずれも、これらのスーパーノードの中に追加し、ストリーム中に保存されIBM SPSS Modeler Advantageに認識されることを保証することを強くお勧めします。
- 各入力ノードのすぐ下流にある式ルールを含むローカルルールノードは、そのフィールドが式マネージャに含まれていなければなりません。ローカルルールは、SPSS Modelerによってではなく、IBM Analytical Decision ManagementまたはIBM SPSS Modeler Advantage中だけで作成できます。
- データ区分ノードがモデリングノードの上流にある（そして、データ型ノードの下流にある）場合は検証およびテスト用のデータソース作成についてのモデルの検証を有効にするための自動的データ区分がIBM SPSS Modeler Advantageで可能となり、チェックボックスの状態は、データ区分ノードが可能であるかどうかによって設定されます。
- 自動データ準備（ADP）ノードがデータ区分ノードの下流にある場合は、信頼できるモデル作成のための自動的データクリーンアップ準備が可能となり、チェックボックスの状態は、データ区分ノードが可能であるかどうかによって設定されます。

- ルール ノードまたはセレクトノードがモデル作成ノードの上流に存在する場合は、これらは使用する選択を指定セクションに書き込むために使用されます。
- ストリーム パラメータが定義されている場合（[ストリームのプロパティ]ダイアログ ボックスの[パラメータ]タブ）は、[パラメータ]リンクがIBM SPSS Modeler Advantageの[モデル] タブと[スコア] タブに表示されます。ユーザーはこのリンクをクリックして、これらの値を指定することができます。

### **IBM SPSS ModelerからのIBM SPSS Modeler Advantageの起動**

SPSS Modelerスプラッシュ画面から、IBM SPSS Modeler Advantageによりモデルの作成を選択できます。

# アクセス機能

アクセス機能は、運動や視覚の制限など、情報技術製品をうまく活用できないユーザーを支援します。IBM は、年齢や能力に関係なく、誰もが利用できる製品を提供するよう努めています。本製品では、標準的なブラウザ ナビゲーション キーを採用しています。

IBM® Analytical Decision Management はブラウザ ベースのアプリケーションのため、たとえば、視覚障害者のような障害者向けのアクセシビリティ オプションは、ブラウザの設定で制御します。キーボード ショートカット、スクリーン リーダー等のユーザー インタフェースのアクションについての詳細は、お使いのウェブ ブラウザの文書を参照してください。

## ベンダー ソフトウェア

IBM Analytical Decision Management には、IBM 使用許諾書で保証されていない一部のベンダー ソフトウェアの使用が必要になる場合があります。IBM では、これらの製品のアクセス機能については明言しません。ベンダーの製品に関するアクセス機能情報については、ベンダー にお問い合わせください。

## IBMとアクセス機能

IBM Noアクセス機能に対するコミットメントの詳細については、[IBM Human Ability and Accessibility Center](#) にお問い合わせください。

# ヘルプ アクセス機能

画面読み上げソフトウェアの追加サポートを提供するために、代替バージョンのユーザーヘルプシステムが製品にインストールされています。このバージョンのヘルプに切り替えるには、バッチ ファイルを実行して既存のヘルプ システムを、画面読み上げソフトウェアが効率的に読み上げられる形式に変換する必要があります。代替バージョンでは、目次は右側にあり、ヘルプ トピックは左側に表示されます。

画面読み上げソフトウェアのサポートを有効にするには

- ▶ サーバー上のヘルプのルート ディレクトリ（たとえば、C:\Program Files\IBM\SPSS\Deployment\5.0\help\en\DecisionManagement\ClaimsManagement\userhelp）に移動します。

- ▶ 画面読み上げソフトウェアに適した形式に切り替えるには、バッチ ファイル `format-for-screen-readers.bat` をダブルクリックして実行します。

デフォルトの形式に戻すには、バッチ ファイル `restore-default-format.bat` を実行します。

注：IBM Analytical Decision Management 7 Information Center は、アクセス機能が有効です。

# 表示

この情報は世界中に提供される製品とサービスのために作成されました。

IBM は本文書内で考察された製品、サービス、または、機能を他の国では提供しないかもしれません。各地域で現在入手可能な製品やサービスに関する情報は現地の IBM 担当者にお問い合わせください。IBM 製品、プログラム、または、サービスに言及することは、IBM 製品、プログラム、または、サービスしか使用できないことを述べることを、あるいは、暗示することを意図していません。IBM の知的所有権に抵触することなく、機能的に同等な製品、プログラム、または、サービスを使用することができます。ただし、IBM 以外の製品、プログラム、または、サービス運用の評価と確認は使用者の責任とします。

IBM は、本書に記載されている内容に関して特許権（特許出願中のものを含む）を保有している場合があります。本書の提供は、お客様にこれらの特許権について実施権を許諾することを意味するものではありません。実施権についてのお問い合わせは、書面にて下記宛先にお送りください。

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

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.

**以下の条項は、イギリスまたはこのような条項が法律に反する他の国では適用されません。** International Business Machines は、明示的または黙示的に関わらず、第三者の権利の侵害しない、商品性または特定の目的に対する適合性の暗黙の保証を含むがこれに限定されない、いかなる保証なく、本出版物を「そのまま」提供します一部の州では、特定の取引の明示的または暗示的な保証の免責を許可していないため、この文が適用されない場合があります。

この情報には、技術的に不適切な記述や誤植を含む場合があります。情報については変更が定期的に行われます。これらの変更は本書の新版に追加されます。IBM は、本書に記載されている製品およびプログラムについて、事前の告知なくいつでも改善および変更を行う場合があります。



IBM 以外の Web サイトに対するこの情報内のすべての参照は、便宜上提供されているものであり、決してそれらの Web サイトを推奨するものではありません。これらの Web サイトの資料はこの IBM 製品の資料に含まれるものではなく、これらの Web サイトの使用はお客様の責任によるものとします。

IBM はお客様に対する一切の義務を負うことなく、自ら適切と考える方法で、情報を使用または配布することができるものとします。

以下のことを可能にする目的でこのプログラムの情報を望むライセンス所有者は、(i) 独自に作成したプログラムと他のプログラム（このプログラムを含む）の間の情報の交換、および、(ii) 交換された情報の相互使用、下記にご連絡下さい：

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

このような情報は適切な条件の下で、場合によっては有償で、使用可能となります。

本書に記述されたライセンス対象プログラムとそれに使用できるすべてのライセンス対象資料は、BM Customer Agreement、IBM International Program License Agreement、または、我々の間の相当する契約の条件に基づいて提供されます。

本書に記述された性能は、管理された環境下で判定されたものです。従って、異なる使用環境で取得された結果は大きく異なる可能性があります。一部の測定は開発段階のシステムで行われた場合もあり、これらの測定結果が一般的に入手可能なシステムと同じであることは保証されません。また、一部の測定値は外挿法により推定されています。実際の結果は異なる場合もあります。本書のユーザーは個々の環境で入手したデータを検証する必要があります。

IBM 以外の製品に関する情報は、その製品の供給者、出版物、もしくはその他の公に利用可能なソースから入手したものです。IBM は、それらの製品のテストは行っておりません。したがって、他社製品に関する実行性、互換性、またはその他の要求については確認できません。IBM 以外の製品の性能に関する質問は、それらの製品の供給者にお願いします。

IBMの将来的な方向性や意図に関する記述は予告なしに変更または撤回されることがあり、単に目標または目的を表すものです。

この情報には、日常の業務処理で用いられるデータや報告書の例が含まれています。できる限り詳細に説明するため、例には、個人、企業、ブランド、製品などの名前が使用されています。これらの名称はすべて架空のものであり、実際の企業で使用される名称および住所とは一切関係ありません。

電子的に複製されたこの情報を表示する場合には、写真やカラーのイラストは表示されない場合があります。

## 商標

IBM、IBM ロゴ、および `ibm.com`、SPSS は、世界の多くの国で登録された IBM Corporation の商標です。IBM の商標の現在のリストは、<http://www.ibm.com/legal/copytrade.shtml> を参照してください。

Adobe、Adobe のロゴ、PostScript、および、PostScript のロゴは、Adobe Systems Incorporated のアメリカ合衆国、その他の国家、または両方における登録商標、または、商標です。

Linux は、米国およびその他の国における Linus Torvalds の登録商標です。

Microsoft、Windows、Windows NT、および Windows のロゴは、米国およびその他の国における Microsoft 社の商標です。

UNIX は、米国およびその他の国における The Open Group の登録商標です。

Java およびすべての Java ベースの商標およびロゴは、米国およびその他の国の Sun Microsystems, Inc. の商標です。

SAS は SAS Institute Inc. のアメリカ合衆国、その他の国家、または両方における登録商標です。

その他の製品名およびサービス名等は、IBM または他の会社の商標です。

# 索引

小数点記号, 36  
非監視学習, 60  
分布図, 57, 70  
制御  
  一般的なオプション, 90  
収益  
  最大化, 5  
  シミュレート, 11  
商標, 106  
概要  
  管理, 89  
注釈  
  ルールへの追加, 53  
管理  
  概要, 89  
  アプリケーションのショートカット, 91  
  スコアリング宛先のオプション, 93  
  データ オプション, 92  
  一般的なオプション, 90  
  ホームページ, 91  
設定, 24  
真理値 (真偽) 表データ, 63  
  
bank\_customer\_data.txt, 8  
bank\_response\_data.txt, 8  
clustering, 60  
Cognos BI ソース, 41  
Excel データ ファイル, 38  
IBM Analytical Decision Management と IBM  
  SPSS Modeler の間のストリームの共有, 99  
IBM Cognos BI ソース, 41  
IBM Cognos BI ソース ノード  
  アイコン, 43  
IBM SPSS Collaboration and Deployment  
  Services Enterprise View データ ソース,  
  40  
IBM SPSS Collaboration and Deployment  
  Services Repository  
  保存先, 25-27  
IBM SPSS Modeler Advantage  
  例, 5  
IBM SPSS Statistics データ ファイル, 38  
ILOG ルール  
  アプリケーションで 사용되는, 53  
  プロジェクト メタデータのダウンロード,  
  93  
k-means モデル, 60  
keywords, 25  
Kohonen モデル, 60  
ROI グラフ, 74  
TwoStep クラスタ モデル, 60

URL  
  アプリケーション, 21  
  
アイコン, IBM Cognos BI, 43  
アクセス機能, 102  
アソシエーション モデル, 62, 66  
  テーブル形式データ, 63  
  テーブルフォーマットデータとトランザクション  
  フォーマット データ, 63  
  トランザクション形式データ, 63  
アソシエーション ルール モデル, 62  
アソシエーションモデリングの例, 17  
アプリケーション  
  起動, 7  
  テスト, 78  
アプリケーション URL, 21  
アプリケーション ショートカットの制御, 91  
アプリケーションのショートカット, 21, 91  
アプリケーションのテスト, 78  
アプリケーションのホームページ, 23  
アプリケーション起動ページ, 21, 91  
  アプリケーションの追加, 22  
  カスタマイズ, 22  
  プロジェクト設定ダイアログ, 22  
  
インタラクティブ リスト モデル, 58, 65  
  
式エディタ, 45  
  
オンライン ヘルプ, 96  
  アクセス機能, 102  
  代替システム, 102  
  画面読み上げソフトウェア, 102  
  
カテゴリ データ, 34  
カンマ, 36  
  
ギャラリー  
  未完了の作業, 23  
  実行時間の長いジョブ, 23  
  
手動クラスタ  
  clustering, 61  
  クラスタモデリングの例, 15  
  クラスタリング モデル, 60, 65  
  
ゲイン グラフ, 11, 70  
  
コーチ テキスト, 96

## 索引

- サンプル アプリケーション, 6
- サンプル アプリケーション用の zip ファイル, 6
- サンプル アプリケーション用のアイテム, 6
- サンプル アプリケーション用のデータ, 6
- サンプル アプリケーション用のファイル, 6
- サンプル データ, 6
- サンプル ファイル, 6
  
- シミュレーション
  - デフォルトの日付, 25
- シミュレーション実施日, 25
- ショートカット
  - アプリケーション, 21
- シーケンス検出, 62
  
- スクリーン リーダー, 102
- スコアリング モデル, 14, 81, 88
  - Cognos BI サーバーの宛先, 84
  - スコアリングの宛先の選択, 81
  - データの選択, 87
  - データベースの宛先, 82
  - ファイルの宛先, 82
  - 出力フィールドの選択, 86
  - フィールドをマッピング, 85
- スコアリング宛先オプションのロック, 93
- スコアリング宛先のオプション, 93
- スコアリング・フィールドをマッピング, 85
- ストリーム
  - アップロード, 25
- ストリームのアップロード, 25
- ストリームファイルのアップロード, 25
  
- セグメント ルール
  - モデルを使用, 47
  
- タイプレス データ, 35
- ダイレクト メール の例, 5
  
- ツールヒント, 96
  
- テキスト ベースのデータ ファイル, 36
- デモ用アイテム, 6
- データ型, 34
- 名義データ, 34
- 序数データ, 34
- 応答データ, 8
- 連続データ, 34
- 顧客データ, 8
- データ オプション, 92
- データ オプションのロック, 92
- データ ファイル, 6
  
- データソース
  - 定義, 8, 30, 35
  - Excel, 38
  - IBM Cognos BI, 41
  - IBM SPSS Collaboration and Deployment Services Enterprise View, 40
  - IBM SPSS Statistics, 38
  - 式エディタ, 45
  - テキストベース, 36
  - データベース, 39
  - ファイル ソース, 36
  - 入力フィールド, 45
  - 追加フィールドの追加, 43
  - フィールドの関連付け, 46
  - プレビュー, 34
  - 測定レベル, 34
  - データのプレビュー, 34
  - データベース ソース, 39
  - テーブル形式データ, 63
  
- トランザクション形式データ, 63
  
- 法的な表示, 104
- 一般的なオプション, 90
- 利用可能なドキュメント, 96
  
- 未完了の項目
  - ギャラリ, 23
  - 管理の概要, 89
  - 予測の重要度グラフ, 57
  - 変数の重要度のグラフ, 57
  - 日付のフォーマット, 24
  - 時間のフォーマット, 24
  - 通貨のフォーマット, 24
  
- バスケット データ, 63
- バージョン ラベル, 27
- バージョンのラベル付け, 27
  
- ビジネス ルール, 47
  
- ファイルの保存, 25-27
- ファイルの開封, 25
- フィールド
  - 作成, 45
  - 定義, 8
  - target, 8
  - マップ済み, 46
  - 区切り文字, 36
- 入力フィールド, 8
  - 定義, 45
  - マップ済み, 46

- 対象フィールド, 8
  - モデルの構築, 10, 55
- 追加フィールドの編集, 43
- 追加フィールドの追加, 43
- 追加フィールドの追加/編集, 43
- フィールドの関連付け, 46
- フィールドをマッピング, 85
- フラグ型データ, 34
- プロジェクト データ ソース, 30
- プロジェクト データ モデル, 8, 30
- プロジェクト設定ダイアログ, 22
- プロジェクトのロック, 28
- プロフィット グラフ, 11, 73
  
- ヘルプ, 96
  - アクセス機能, 102
  - 代替システム, 102
  - 画面読み上げソフトウェア, 102
- ヘルプの利用, 96
  - 利用可能なドキュメント, 96
- 起動ページ, 7, 21, 91
  - アプリケーションの追加, 22
  - カスタマイズ, 22
  - プロジェクト設定ダイアログ, 22
- ペーパー ロール データ, 63
  
- ホーム ページを非表示にする, 91
- ホームページ, 23, 91
  
- マイ データ ソース, 30
- 式マネージャ, 45
  
- モデル
  - 予測, 55-56
  - 構築, 10, 55-56
  - 結果, 57
  - 評価, 11, 68
  - association, 62
  - clustering, 60
  - Cognos BI サーバーの宛先選択, 84
  - インタラクティブ リスト, 58
  - オプションの設定, 64-66
  - グラフ, 57
  - スコア宛先の選択, 81
  - スコアリング, 14, 81, 88
  - スコアリング・フィールドをマッピング, 85
  - テスト, 13, 78
  - データ選択, 87
  - データベース宛先の選択, 82
  - 作成時間の制限, 25
  - 環境設定の設定, 25
  - ファイル宛先の選択, 82
  - 対象フィールド, 10, 55
  - 出力フィールドの選択, 86
  - 予測ルール, 58
    - ルールで使用, 47
    - レコード選択, 87
    - 利益をシミュレート, 75
  - 未精製モデル, 62
  - 予測モデル, 55-56
  - モデルにおけるデータ選択, 87
  - モデルにおけるレコード選択, 87
  - モデルの構築
    - 対象フィールド, 10
  - モデルの評価, 11, 68
    - インタラクティブ リスト モデル, 12
    - クラスタリング モデル, 15, 17
    - 予測ルールモデル, 12
  - モデルのCognos BI サーバーの宛先, 84
  - モデル内の利益シミュレーション, 75
  - モデル内のスコアの宛先, 81
  - モデル内のチャート, 57
  - モデルのテスト, 13, 78
  - モデル内のデータベースの宛先, 82
  - モデル内のファイルの宛先, 82
  - モデルの出力フィールド, 86
- ユーザー環境設定, 24
  
- 区切り文字, 36
- リフト グラフ, 71
  
- ルール, 47
  - 再利用, 51, 53, 93
  - 作成, 47
  - 共有, 51, 53, 93
  - 外部, 53, 93
  - 注釈, 53
  - 選択, 50
  - ILOG, 53, 93
  - エクスポート, 51
  - セグメント, 47
  - モデルを使用, 47
  - 選択ロジック, 50
  - 共有ルール, 51
  - 包含ルール, 50
  - 外部ルール
    - アプリケーションで使用される, 53
    - プロジェクト メタデータのダウンロード, 93
  - 選択ルール, 50
  - 除外ルール, 50
  - ルールのエクスポート, 51
  - 予測ルールモデル, 58, 65

索引

レスポンス グラフ, 72  
測定レベル, 34