

IBM® SPSS® Modeler Advantage 7



Hinweis: Lesen Sie zunächst die allgemeinen Informationen unter Hinweise auf S. 95, bevor Sie dieses Informationsmaterial sowie das zugehörige Produkt verwenden.

Diese Ausgabe bezieht sich auf IBM Analytical Decision Management 7 und alle nachfolgenden Versionen sowie Anpassungen, sofern dies in neuen Ausgaben nicht anders angegeben ist.

Screenshots von Adobe-Produkten nachgedruckt mit Genehmigung durch Adobe Systems Incorporated.

Screenshots von Microsoft-Produkten werden mit Genehmigung der Microsoft Corporation abgedruckt.

Lizenziertes Material - Eigentum von IBM

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

Eingeschränkte Rechte für Benutzer der US-Regierung: Verwendung, Vervielfältigung und Veröffentlichung eingeschränkt durch GSA ADP Schedule Contract mit der IBM Corp.

Informationen zu IBM Business Analytics

Die Software IBM Business Analytics liefert umfassende, einheitliche und korrekte Informationen, mit denen Entscheidungsträger die Unternehmensleistung verbessern können. Ein umfassendes Portfolio aus [Business Intelligence](#), [Vorhersageanalyse](#), [Finanz- und Strategiemangement](#) sowie [Analyseanwendungen](#) bietet Ihnen sofort klare und umsetzbare Einblicke in die aktuelle Leistung und gibt Ihnen die Möglichkeit, zukünftige Ergebnisse vorherzusagen. Durch umfassende Branchenlösungen, bewährte Vorgehensweisen und professionellen Service können Unternehmen jeder Größe die Produktivität maximieren, Entscheidungen automatisieren und bessere Ergebnisse erzielen.

Als Teil dieses Portfolios unterstützt IBM SPSS Predictive Analytics-Software Unternehmen dabei, zukünftige Ereignisse vorherzusagen und proaktiv Maßnahmen zu ergreifen, um bessere Geschäftsergebnisse zu erzielen. Kunden aus Wirtschaft, öffentlichem Dienst und dem Bildungsbereich weltweit nutzen IBM SPSS-Technologie als Wettbewerbsvorteil für Kundengewinnung, Kundenbindung und Erhöhung der Kundenumsätze bei gleichzeitiger Eindämmung der Betrugsmöglichkeiten und Minderung von Risiken. Durch die Einbindung von IBM SPSS-Software in ihre täglichen Operationen wandeln sich Organisationen zu “Predictive Enterprises” – die Entscheidungen auf Geschäftsziele ausrichten und automatisieren und einen messbaren Wettbewerbsvorteil erzielen können. Wenn Sie weitere Informationen wünschen oder Kontakt zu einem Mitarbeiter aufnehmen möchten, besuchen Sie die Seite <http://www.ibm.com/spss>.

Technischer Support

Kunden mit Wartungsvertrag können den technischen Support in Anspruch nehmen. Kunden können sich an den technischen Support wenden, wenn sie Hilfe bei der Arbeit mit den Produkten von IBM Corp. oder bei der Installation in einer der unterstützten Hardware-Umgebungen benötigen. Zur Kontaktaufnahme mit dem technischen Support besuchen Sie die Website von IBM Corp. unter <http://www.ibm.com/support>. Sie müssen bei der Kontaktaufnahme Ihren Namen, Ihre Organisation und Ihre Supportvereinbarung angeben.

Inhalt

1	Informationen zu IBM Analytical Decision Management	1
	Neuerungen	1
	Informationen zu IBM Analytical Decision Management.	3
2	Beispiel für die Modellierung der Kundenreaktion	5
	IBM SPSS Modeler Advantage	5
	Profit für Direktwerbungskampagnen maximieren.	5
	Beispieldateien.	6
	Daten	6
	Beispiele	6
	Die Anwendung starten	7
	Datenquellen einrichten	8
	Erstellen des Modells	10
	Auswertung des Modells	10
	Vorhersageregelnmodelle	12
	Testen des Modells.	13
	Scoren des Modells	13
	Beispiel für Cluster-Modellierung	14
	Beispiel für Assoziationsmodellierung	16
3	Navigation im Arbeitsbereich	20
	Anwendungen starten	20
	Projektkonfiguration	21
	Die Hauptseite	22
	Galerie	22
	Voreinstellungen anpassen.	23
	Öffnen und Speichern von Projekten, Modellen und Regeln	24
	Speichern im IBM SPSS Collaboration and Deployment Services Repository	25
	Arbeiten mit Beschriftungen	26
	Projekte sperren	26
	Objekteigenschaften.	27

4 Verwaltung von Datenquellen 28

Vorschau von Daten	31
Messniveaus	31
Erstellung einer neuen Datenquelle	32
Dateiquelle	33
Datenbankquelle	36
Enterprise-Ansichts-Quelle	36
IBM Cognos BI-Quelle	37
Hinzufügen von Zusatzfeldern	39
Auswählen von Eingabefeldern	40
Ausdruckseditor	41
Abbildung von Feldern	41

5 Regeln verwenden 43

Regeln erstellen	43
Segmentregeln definieren	43
Selektionsregeln definieren	45
Regeln exportieren und wiederverwenden	46
Anmerkungen hinzufügen	48
Verwenden von Regeln aus dem ILOG Business Rule Management System	48

6 Vorhersagemodelle erstellen 50

Informationen zu Vorhersagemodellen	50
Erstellen von Modellen	50
Erstellen eines Vorhersagemodells	51
Ergebnisse der Vorhersagemodellierung	52
Erstellen eines Vorhersageregelmodells (interaktiv)	53
Erstellen eines Clustermodells	54
Erstellen eines Assoziationsmodells	56
Optionale Modelleinstellungen	58
Auswertung von Modellen	62
Maximalen Profit simulieren	69
Testen des Modells	72
Verwenden von Modell-Scores in Anwendungen	73

7 *Scoren von Modellen* **75**

Scoring von Modellen in eine Datenbanktabelle, eine Datei oder einen Cognos BI-Server	75
Scoring-Ziel auswählen	75
Ausgabefelder auswählen	79
Daten und Datensatzgruppen auswählen	80
Datensätze für Scoring auswählen	81
Scoring von Modellen	81

8 *Anwendungen verwalten* **83**

Allgemeine Optionen	84
Vorkonfiguration von Anwendungs-Verknüpfungen	85
Ausblenden der Hauptseite	85
Sperrung der Datenoptionen	86
Sperrung der Optionen zum Scoring-Ziel	86
Herunterladen von Projekt-Metadaten zur Verwendung beim Erstellen externer Regeln	86

9 *Hilfe aufrufen* **89**

Verfügbare Dokumentation	89
------------------------------------	----

Anhänge

A *Freigabe von Streams* **91**

B *Eingabehilfen* **93**

Zugriffsmöglichkeiten auf die Hilfe	93
---	----

<i>C Hinweise</i>	<i>95</i>
<i>Index</i>	<i>98</i>

Informationen zu IBM Analytical Decision Management

Neuerungen

In IBM® Analytical Decision Management 7 sind folgende neue Funktionen enthalten:

Optimierung. Das Ziel der Optimierung besteht darin, die Lösung zu ermitteln, die am besten für ein bestimmtes Ziel geeignet ist, beispielsweise die Maximierung des Umsatzes aus einer Marketing-Kampagne bzw. die Minimierung des Betrugs- bzw. Abwanderungsrisikos. Das Ergebnis der Optimierung ist eine Lösung, die diese Frage für alle möglichen Kombinationen innerhalb ihrer Population beantwortet, und zwar auf die Weise, die den zu optimierenden Wert (z. B. Profit) am besten maximiert.

Clustermodelle. Neu in IBM® SPSS® Modeler Advantage - mit Methoden zur Cluster-Modellierung können Gruppen von Datensätzen mit ähnlichen Werten oder Mustern leichter erkannt werden. Diese Verfahren werden beim Marketing (Kundensegmentierung) und in anderen Geschäftsanwendungen verwendet (Datensätze in Clustern, die nur aus einem einzigen Datensatz bestehen, enthalten möglicherweise Fehler oder es handelt sich dabei um Betrugsfälle). Clustering-Modelle werden häufig als nicht überwachte Lernmodelle bezeichnet, da sie kein Zielfeld benutzen und keine bestimmte Vorhersage liefern, die als „wahr“ oder „falsch“ ausgewertet werden kann. Mit IBM Analytical Decision Management können Sie automatisch Clustermodelle erstellen. Sie können auch Cluster kombinieren, umbenennen, bearbeiten oder löschen oder benutzerdefinierte Cluster auf der Grundlage bereits bestehenden Geschäftswissens hinzufügen.

Assoziationsmodelle. Neu in IBM SPSS Modeler Advantage - Assoziationsmodelle ermöglichen eine Vorhersage darüber, welche Elemente höchstwahrscheinlich nah beieinander liegen und wie stark die Beziehung zwischen den Elementen ist. Assoziationsregeln ordnen eine bestimmte Schlussfolgerung (z. B. den Kauf eines bestimmten Produkts) einer Menge von Bedingungen (z. B. dem Kauf mehrerer anderer Produkte) zu.

Daten. Die Registerkarte „Daten“ ermöglicht jetzt mehrere Dateneingaben, das Definieren des Projektdatenmodells in einer Application-Ansicht, das Referenzieren von Modellausgaben, als wären es Daten, das Hochladen von lokalen Dateien sowie eine Vorschau und das Bestimmen der Menge von Daten.

Verbesserung der Integration von IBM Cognos-Daten. Wie auch IBM® SPSS® Modeler 15 unterstützt IBM Analytical Decision Management jetzt den Import von in Cognos im Vorab erstellten Berichten, Cognos-Filtern und -Berechnungen sowie die Steuerung der automatischen Disaggregation.

Mehrere Elemente. Anwendungen können ab jetzt mehr als eine Dimensionshierarchie umfassen. Beim Starten einer Anwendung, werden Sie dazu aufgefordert, mindestens eines oder mehrere der verfügbaren Elemente (Hierarchien) auszuwählen, die in der Anwendung verwendet werden sollen, wenn die ausgewählte Anwendung für mehrere Elemente konfiguriert wurde. Der Dimensionsbaum auf der Registerkarte „Definieren“ gibt den Bereich der möglichen Entscheidungen oder Empfehlungen an, die zurückgegeben werden können. Wenn eine Anwendung mehrere Dimensionshierarchien aufweist, sind diese auf separaten Registerkarten des Dimensionsbaumbereichs auf der Registerkarte „Definieren“ enthalten.

Dynamische Zuordnung. In den Vorgängerversionen musste ein Objekt im Dimensionsbaum vorhanden sein, um es zuzuordnen zu können. Jetzt können Dimensionselemente ohne untergeordnete Elemente von einer Tabelle abgerufen werden, die durch einen Schlüssel im Projektdatenmodell verknüpft ist. Beispielsweise liegt das Dimensionsobjekt *Autos* vor, das keine untergeordneten Elemente aufweist. Sie können dynamische Zuordnungen für *Autos* einrichten, um zum Beispiel Felder in der Tabelle *BMW*, *Audi* und *Ford* zu referenzieren.

Mehrfachzuordnungen. Die neue Option Mehrfachzuordnung ermöglicht die Zuordnung von Datensätzen zu mehreren Dimensionselementen (z. B. mehrere Angebote) in den einzelnen Segmenten. Mehrfachzuordnungen können für jede Regel anstelle von einzelnen Zuordnungen pro Regel durchgeführt werden.

Verbesserung der Benutzeroberfläche. Viele der Änderungen und Neuerungen an der IBM Analytical Decision Management-Benutzeroberfläche wurden eingeführt, um die Benutzerfreundlichkeit und Skalierbarkeit aller Anwendungen zu verbessern.

Erweiterungen der Galerie. Die Galerie wurde um zahlreiche Features erweitert, wie die Option Galerie löschen, mit der mehrere Objekte auf einmal aus der Galerie gelöscht werden können.

Erweiterungen des Regeleditors. Je nach dem, von wo der Regeleditor gestartet wird, können Benutzer jetzt den Regeltyp für Selektionsregeln festlegen (einschließend oder ausschließend), sie können Aggregationsregeln Risikopunkte zuweisen, eine Dimension für Zuordnungsregeln auswählen sowie die Operatoren *ODER* und *Aufteilen ODER* verwenden, um mehrere Ausdrücke innerhalb einer Regel zu kombinieren und so einen *ODER*-Ausdruck zu erstellen. Beim Erstellen von Regeln können Sie jetzt auch eine Dimension aus einem Hierarchiebaum auswählen (für Anwendungen mit mehr als einem Hierarchiebaum) oder den Ausdruckseditor starten, um einen Ausdruck zu erstellen. Diese Option stand bisher im Dialogfeld des Regeleditors noch nicht zur Verfügung.

Erweiterte Integration mit dem ILOG Business Rule Management System. In IBM® SPSS® Collaboration and Deployment Services Deployment Manager können Sie das Dialogfeld „Neue externe Regel hinzufügen“ dazu verwenden, Verweise zu ILOG-Regeln zu definieren, die dann wiederum in IBM Analytical Decision Management verwendet werden können. Bisher waren diese Regeln nur per Webdienst zugänglich. Mit dieser Version können Sie nun auf lokale Regeln in der ILOG-Datenbank verweisen. Bei der Option „Lokal“ wird die Regel über den IBM® SPSS® Modeler Server ausgeführt (der Server, der in Deployment Manager für die Verwendung durch IBM Analytical Decision Management definiert wurde).

WhatIf? -Erweiterungen. Das Dialogfeld WhatIf? wurde um zahlreiche neue Features erweitert. So ist es ab sofort möglich als Simulationsdatum einen Datumsbereich anzugeben, innerhalb dessen die WhatIf?-Auswertung ausgeführt werden soll.

Hochladen von Stream-Dateien. Sie können ab sofort SPSS Modeler-Stream-Dateien aus dem lokalen Dateisystem öffnen und bei Bedarf im Repository speichern.

Mehrere SPSS Modeler Server-Zugangsberechtigungen. Wenn in Deployment Manager mehrere SPSS Modeler Server-Zugangsberechtigungen definiert sind, können Sie nun auch mehrere zur Verwendung mit IBM Analytical Decision Management angeben. Wenn sich dann der IBM Analytical Decision Management-Benutzer anmeldet, wird vom Server die erste Zugangsberechtigung in der Liste verwendet, um zu prüfen, ob der Benutzer berechtigt ist. Dieser Vorgang wird so lange fortgesetzt, bis in der Liste eine Zugangsberechtigung mit der entsprechenden Berechtigung gefunden wird. Wenn der Benutzer keine Berechtigung für eine Zugangsberechtigung in der geordneten Liste aufweist, kann dieser IBM Analytical Decision Management nicht verwenden. Somit kann genau geregelt werden, welche Benutzer Zugang zu welchen Teilen des SPSS Modeler Server-Dateisystems haben.

In der vorhergehenden Version, IBM SPSS Decision Management 6.2, wurden die folgenden Funktionen hinzugefügt:

Unterstützung für Cognos-Daten. Bei Verwendung mit SPSS Modeler Server 14.2 oder höher bietet IBM Analytical Decision Management die Möglichkeit, beim Definieren von Datenquellen Daten aus einem IBM Cognos Business Intelligence-Server zu lesen und die Scoring-Ergebnisse wieder auf einen IBM Cognos BI-Server zu schreiben. Nach der Definition einer Verbindung mit einem Cognos BI-Server, können Sie Pakete, Dimensionen, Abfragen und Berichte durchsuchen, um die gewünschten Felder auszuwählen. [Für weitere Informationen siehe Thema IBM Cognos BI-Quelle in Kapitel 4 auf S. 37.](#)

Integration mit dem ILOG Business Rule Management System. In einem Business Rules Management System wie ILOG erstellte Regeln können in IBM Analytical Decision Management-Anwendungen referenziert und verwendet werden, vorausgesetzt, sie wurden so entwickelt, dass sie dasselbe Datenmodell unterstützen, das auch im aktuellen IBM Analytical Decision Management-Projekt verwendet wird. Diese externen Regeln können ausgewählt und auf dieselbe Weise verwendet werden wie andere Modelle und Regeln in der IBM Analytical Decision Management-Anwendung. [Für weitere Informationen siehe Thema Verwenden von Regeln aus dem ILOG Business Rule Management System in Kapitel 5 auf S. 48.](#)

Unterstützung für Linux auf System Z. SUSE Enterprise Server 10 wird auf 64-Bit IBM System z10 unterstützt.

Informationen zu IBM Analytical Decision Management

IBM® Analytical Decision Management macht prädiktive Analytik für tatsächliche betriebswirtschaftliche Probleme nutzbar und ermöglicht es Ihnen, speziell auf Ihre Kunden oder Ihre Branche zugeschnittene Anwendungen zu erstellen. Während Anwendungen typischerweise

für die Behebung spezieller Probleme konfiguriert sind, basieren alle auf einem gemeinsamen Funktionssatz:

- Entscheidungen mithilfe von Geschäftsregeln automatisieren.
- Erkenntnisse aus Vorhersagemodellen gewinnen.
- Priorisierung, Optimierung oder Simulation einsetzen, um optimale Entscheidungen zu treffen.

Es sind verschiedene Anwendungspakete erhältlich, die auf bestimmte Geschäftsprobleme zugeschnitten sind. Weitere Informationen erhalten Sie bei Ihrem Händler.

Beispiel für die Modellierung der Kundenreaktion

IBM SPSS Modeler Advantage

IBM® SPSS® Modeler Advantage ist eine benutzerfreundliche Anwendung, die Geschäftsnutzern das ganze Potential der prädiktiven Modellierung zugänglich macht. Mithilfe von Vorhersagemodellen können Sie Muster in Vergangenheitsdaten aufdecken und die Wahrscheinlichkeit künftiger Ereignisse beurteilen.

Sie können beispielsweise ein Modell verwenden, um auf der Grundlage von Merkmalen wie Einkommen, Alter und Mitgliedschaft in bestimmten Organisationen vorherzusagen, welche Kunden mit der geringsten Wahrscheinlichkeit kündigen oder mit der höchsten Wahrscheinlichkeit auf ein bestimmtes Angebot reagieren. Die daraus hervorgehenden Vorhersagen können zur Erstellung von Listen von Zielkunden oder relevanten Fällen benutzt, als Ausgangsmaterial für die strategische Planung oder im Kontext einer Prognoseanwendung zusammen mit Regeln eingesetzt werden.

Profit für Direktwerbungskampagnen maximieren

In diesem Beispiel eines Vorhersagemodells vermarkten Sie Finanzprodukte und -dienstleistungen an die Kunden einer Bank. In einer Testkampagne wurde kürzlich einer Auswahl von Kunden eine von mehreren neuen Dienstleistungen angeboten und ihre Reaktion erfasst. Die Rücklaufquote war jedoch niedriger als erhofft, und die Kampagne verfehlte die Gewinnschwelle.

Ihr Ziel ist es, diesen Verlust mit IBM® SPSS® Modeler Advantage in einen Gewinn umzuwandeln, indem Sie unter Berücksichtigung von erwarteten Kosten, Umsatz und anderen Faktoren die aussichtsreichsten verfügbaren Kunden ansprechen.

Die Ergebnisse der Testkampagne werden im Folgenden zusammengefasst.

Tabelle 2-1
Ergebnisse der Testkampagne

Ergebnisse der Testkampagne	
Anzahl der Anschreiben	3000
Anzahl der Antworten	111
Antwortquote	3.7%
Kosten	6750
Umsatz	6660
Gesamtprofit (Verlust)	(90)

Dieses Beispiel verwendet die Datendateien *bank_customer_data.txt* und *bank_response_data.txt*, die mit IBM® Analytical Decision Management bereitgestellt und installiert werden. Eine fertig gestellte Version des Beispiels steht mit der Datei *bank_response_model.str* zur Verfügung. Genaueres zu Installation und Beispieldateien können Sie ggf. von Ihrem Administrator erfahren. [Für weitere Informationen siehe Thema Beispieldateien auf S. 6.](#)

Beispieldateien

In IBM® Analytical Decision Management sind Daten und andere Beispieldateien enthalten, mit denen Sie die in den Anleitungen dargestellten Beispiele rekonstruieren können.

Daten

Bei der Installation wird eine Reihe von Beispieldatendateien im Verzeichnis *\Demos* von IBM® SPSS® Modeler Server installiert (z. B. *C:\Program Files\IBM\SPSS\ModelerServer\15\Demos*). Diese Dateien können zusammen mit vorgefertigten Beispielanwendungen verwendet werden. Es handelt sich dabei um folgende Dateien:

- *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*

Beispieldateien für Anwendungen, Modelle und Regeln, die diese Datendateien verwenden, gehen davon aus, dass das oben angegebene Standardverzeichnis verwendet wird. Wenn Datendateien auf einem anderen Laufwerk oder in einem anderen Verzeichnis installiert wurden, müssen die Datenquellenpfade entsprechend aktualisiert werden.

Beispiele

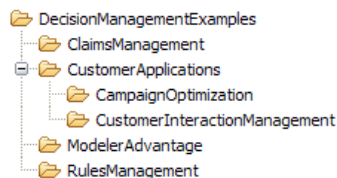
Es wird eine Reihe von Beispielanwendungsdateien, Modellen und Regeln für die Verwendung mit den vorgefertigten IBM® Analytical Decision Management-Anwendungen bereitgestellt. Die Beispiele befinden sich in gepackter Form im Verzeichnis *\Demo* auf dem Installationsdatenträger oder stehen in einer Datei zum Herunterladen zur Verfügung. Gehen Sie wie folgt vor, um die Beispieldateien für eine oder mehrere vorgefertigte Anwendungen in IBM® SPSS® Collaboration and Deployment Services Repository zu importieren.

- ▶ **DVD.** Kopieren Sie die Datei *DecisionManagementExamples.pes* aus dem Ordner *\Demo* auf der DVD an einen temporären Speicherort.

oder

- ▶ **Heruntergeladene Datei.** Laden Sie die Datei *SPSS_Decision_Mgmt_70_demo.zip* an einen temporären Speicherort herunter und extrahieren Sie sie.
- ▶ Starten Sie IBM® SPSS® Collaboration and Deployment Services Deployment Manager und melden sich bei Ihrem Content-Server an.
- ▶ Klicken Sie mit der rechten Maustaste auf den Ordner für das Inhalts-Repository und wählen Sie die Option Importieren aus.
- ▶ Klicken Sie im Dialogfeld „Importieren“ auf Durchsuchen und wählen Sie die Datei *DecisionManagementExamples.pes* aus, die Sie zuvor an einen temporären Speicherplatz kopiert haben. Klicken Sie auf OK.
- ▶ Klicken Sie im Dialogfeld zu den Importkonflikten auf OK. Nach Abschluss des Importvorgangs werden folgende neue Ordner angezeigt, die Beispielprojekte enthalten, die in IBM Analytical Decision Management geöffnet werden können.

Abbildung 2-1
Ordner für Beispielobjekte



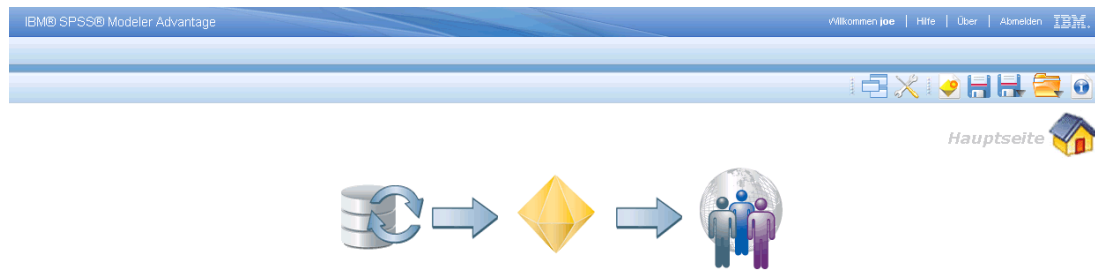
Beachten Sie, dass alle Beispieldateien für Anwendungen, Modelle und Regeln voraussetzen, dass im IBM® SPSS® Modeler Server-Verzeichnis *Demos* Datendateien abgelegt sind (z. B. *C:\Program Files\IBM\SPSS\ModelerServer\15\Demos*). Wenn die Datendateien auf einem anderen Laufwerk oder in einem anderen Verzeichnis installiert wurden, müssen die Datenquellenpfade für die lokale Konfiguration entsprechend aktualisiert werden.

Die Anwendung starten

Der IBM® Analytical Decision Management-Startbereich enthält Verknüpfungen zu den auf Ihrer Site verfügbaren Anwendungen. Diese Verknüpfungen werden von Ihrem Administrator konfiguriert. Sie können evtl. aus weiteren Anwendungen wählen, indem Sie auf den Link *Anwendung hinzufügen* in der rechten oberen Ecke klicken.

- ▶ Öffnen Sie den IBM Analytical Decision Management-Startbereich, um mit der Bearbeitung des Beispiels zu beginnen. Sie können in Ihrem Browser ein Lesezeichen auf diese Seite setzen oder eine Verknüpfung auf Ihrem Desktop erstellen, damit Sie später leichter darauf zugreifen können.
- ▶ Zur Erstellung eines neuen Modells wählen Sie aus dem Dropdown-Menü der Anwendung in der Gruppe IBM® SPSS® Modeler Advantage die Option *Neu* und klicken Sie auf *Wechseln*. Die IBM SPSS Modeler Advantage-Hauptseite wird angezeigt.

Abbildung 2-2
IBM SPSS Modeler Advantage Hauptseite



Jedes Symbol auf der Homepage stellt einen anderen Arbeitsschritt des Modellierungsprozesses dar, angefangen von der Definition der Daten bis zur Erstellung und Auswertung des Modells. Klicken Sie auf das Datensymbol links, um anzufangen.

Datenquellen einrichten

Auf der Registerkarte „Daten“ können Sie Datenquellen für die Beispiel-Datendateien aus dem Beispiel konfigurieren. Daten-Sets können von Ihrem Administrator konfiguriert oder von Ihnen selbst hinzugefügt werden.

Das Projektdatenmodell gibt die Felder vor, die in der Anwendung verfügbar sind, und wird evtl. zur Auswertung von Regeln und Modellen benötigt. Alle anderen Datenquellen müssen auf das Datenmodell abgebildet werden, um in der Anwendung benutzt werden zu können.

Dieses Beispiel verwendet zwei Datendateien, die beide mit der Anwendung geliefert werden.

- **bank_customer_data.txt** enthält Informationen über Bestandskunden, z. B. Kunden-Nr., Zahlungsverlauf, Anzahl erworbener Produkte oder Länge der Kundenbeziehung in Monaten, sowie demografische Daten wie Alter oder Einkommen. Das Projektdatenmodell basiert auf dieser Datei.
- **bank_response_data.txt** enthält Informationen über Kunden, denen in früheren Kampagnen eine oder mehrere Leistungen angeboten wurden, sowie Angaben zur Reaktion der Kunden. Durch Analyse dieser Daten kann die Anwendung ermitteln, welche Kunden bei künftigen Kampagnen mit der höchsten Wahrscheinlichkeit auf ähnliche Angebote eingehen werden.

Standardmäßig werden die Beispiel-Datendateien im Verzeichnis IBM® SPSS® Modeler Server *Demos* abgelegt, auf das Sie über Ihre Anwendung zugreifen können (z. B. *C:\Program Files\IBM\SPSS\ModelerServer\15\Demos*). Wenden Sie sich an den Administrator, falls Sie Hilfe beim Installieren dieser Dateien benötigen.

So definieren Sie das Projektdatenmodell:

- ▶ Klicken Sie im Bereich Projektdatenquellen auf Datenquelle hinzufügen.
- ▶ Geben Sie *bank customer data* als Datenquellennamen ein.
- ▶ Wählen Sie Datei und dann *bank_customer_data.txt*.

- ▶ Achten Sie darauf, dass das Messniveau des Felds *Schlechtes Zahlungsverhalten* unter Eingabefelder festlegen auf Flag eingestellt ist. Die Einstellung „Flag“ bedeutet, dass das Feld zwei mögliche Werte haben kann, die für *wahr* und *falsch* stehen (oft als *T* und *F* oder in numerischen Feldern als 0 und 1 kodiert) und das Vorliegen bzw. Nichtvorliegen des jeweiligen Attributs ausdrücken.
- ▶ Achten Sie außerdem darauf, dass das Messniveau des Felds *Geschlecht* unter Eingabefelder festlegen auf Nominal eingestellt ist. (Die Werte dieses Felds geben einfach an, ob das Geschlecht „männlich“ oder „weiblich“ ist, und nicht *wahr* oder *falsch*, wie es beim Messniveau Flag der Fall wäre.)
- ▶ Klicken Sie auf Speichern.

Da dies die erste für das Projekt definierte Datenquelle ist, wird sie automatisch als Projektdatenmodell ausgewählt und zur Definition der für die Anwendung erforderlichen Felder benutzt. Alle anderen Daten-Sets werden relativ zu dieser Quelle abgebildet.

So definieren Sie die Antwort-Datenquelle:

Das Antwort-Daten-Set enthält weitgehend dieselben Felder, die auch im Projektdatenmodell erscheinen, aber zwei zusätzliche Felder mit der Bezeichnung *Kampagnen-ID* und *Antwort*. Diese Felder geben an, welches Angebot dem jeweiligen Kunden unterbreitet wurde und ob es angenommen wurde. Sie können zur Erstellung eines Modells benutzt werden, das vorhersagt, welche Kunden in Zukunft mit der höchsten Wahrscheinlichkeit reagieren.

- ▶ Wählen Sie im Bereich Projektdatenquellen Datenquelle hinzufügen.
- ▶ Wählen Sie *bank response data* als Datenquellennamen.
- ▶ Wählen Sie Datei und dann *bank_response_data.txt*.
- ▶ Achten Sie darauf, dass das Messniveau des Felds *Antwort* unter Eingabefelder festlegen auf Flag eingestellt ist.
- ▶ Deaktivieren Sie unter Eingabefelder festlegen das Kontrollkästchen neben dem Feld *Kampagnen-ID*, um anzugeben, dass dieses Feld bei der Modellierung nicht als Eingabe verwendet werden soll. (Mit diesem Feld kann angegeben werden, für welche Kampagnen ein Kunde in der Vergangenheit ausgewählt wurde. Es ist jedoch nicht brauchbar zur Vorhersage zukünftiger Antworten.)
- ▶ Klicken Sie auf Speichern.

Da einige der Felder des Datenmodells im Antwort-Daten-Set fehlen, wird für diese Datenquelle in der Spalte Kompatibel ein orangenes Symbol angezeigt, das darauf hinweist, dass manche, aber nicht alle Felder übereinstimmen. Das Daten-Set ist zur Modellierung geeignet, aber nicht zur Simulation, zum Scoring oder Bereitstellen. (Dazu müssten alle Felder des Datenmodells vorliegen.)

Erstellen des Modells

Auf der Registerkarte „Modellierung“ können Sie Datenquelle, Zielfeld und andere Optionen zur Erstellung und Auswertung des Modells festlegen.

- ▶ Wählen Sie *bank response data* als Datenquelle und *Antwort* als Zielfeld.
- ▶ Wenn die Meldung Datenscan erforderlich angezeigt wird, klicken Sie auf das Symbol, um einen Datenscan durchzuführen. Klicken Sie anschließend auf Modell erstellen.

Wenn das Modell fertig ist, wird ein Verteilungsdiagramm angezeigt. Das Diagramm zeigt die beobachteten Werte auf der horizontalen Achse, wobei der Wert 1 für eine positive Antwort steht.

- Die Höhe der Balken drückt die Gesamtrücklaufquote aus. Von den 3000 Kunden des Daten-Sets haben 111 geantwortet, was einer Gesamttrefferquote von 3,8 % entspricht.
 - Die farbigen Bereiche jedes Balkens stehen für die vom Modell vorhergesagten Antworten, wobei Rot einer positiven Vorhersage und Blau einer negativen Antwort entspricht, wie von der Legende rechts oben abzulesen. Aufgrund der niedrigen Gesamtrücklaufquote gibt es relativ wenige vorhergesagte Antworten (rote Bereiche). Der Balken für den Wert „falsch“ (*Antwort = 0*) ist jedoch überwiegend blau, was bedeutet, dass der größte Teil derjenigen, für die keine Antwort vorhergesagt wurde, auch tatsächlich nicht antworteten.
- ▶ Um einen besseren Eindruck davon zu bekommen, wie das Modell funktioniert, fahren Sie mit dem nächsten Abschnitt ([Auswertung des Modells](#)) fort.

Auswertung des Modells

Klicken Sie auf der Registerkarte „Modell“ auf Auswerten, um zu untersuchen, welche Resultate das Modell für Ihre Kundendaten bringt und wie es sich auf Ihren Profit auswirken kann.

- ▶ Wählen Sie *bank response data* als Datenquelle.
- ▶ Wählen Sie als zu erstellende Diagrammtypen Gewinnndiagramm und Profitdiagramm.
- ▶ Wählen Sie Maximalen Profit simulieren und Sie legen als Populationsgröße 3000 fest. Dies entspricht der Zahl aus der Testkampagne.
- ▶ Geben Sie 2.25 für die Kosten und 60 für den Umsatz ein. Diese Werte geben Ihre Kosten für jedes Anschreiben (z. B. für Herstellung, Kuvertierung und Versand) sowie das zu erwartende Umsatzvolumen pro positiver Antwort an.
- ▶ Wählen Sie als Gesuchte Antwort den Wert 1 aus.
- ▶ Klicken Sie auf Ausführen.

Das Gewinnndiagramm zeigt die Verbesserung (den „Lift“), die durch Einsatz des Modells erreicht werden kann. Da bei der Erstellung des Modells die Option Daten automatisch partitionieren gewählt wurde (unter „Optionale Einstellungen“), werden für Trainings- und Testpartition getrennte Diagramme angezeigt. Die Diagramme zeigen den Anteil der in jedem Perzentil enthaltenen Treffer (positiven Antworten) im Verhältnis zur Gesamtzahl der Treffer in der Stichprobe. Die gerade diagonale Linie bildet die ohne Modellverwendung erwartete Antwort ab

– um die Hälfte der Treffer zu realisieren, müssten Sie auch die Hälfte der Personen kontaktieren. Die gekrümmte Linie zeigt die durch das Modell zu erreichende Verbesserung – indem Sie die 60 % der Kunden mit der höchsten Antwortwahrscheinlichkeit ansprechen, decken Sie fast alle Treffer ab, kommen aber mit viel weniger Anschreiben aus.

Das Profitdiagramm zeigt den Gewinn in Form des Gesamtprofits an, je nach den angegebenen Kosten- und Umsatzwerten. Bei der zum Trainieren des Modells verwendeten Stichprobe scheint der Profit direkt unterhalb des 30. Perzentils sein Maximum zu erreichen.

Ein konkretes Beispiel erhalten Sie, indem Sie auf der Registerkarte „Simulieren“ den erwarteten maximalen Profit für die von Ihnen festgelegte Populationsgröße abrufen. Der erwartete Profit wird mit der folgenden Formel ermittelt:

$$\text{erwarteter Profit} = (\text{Antwortwahrscheinlichkeit} * \text{Umsatz}) - \text{Kosten}$$

Hierbei wird die Antwortwahrscheinlichkeit für jeden Kunden anhand des Modells ermittelt. Indem Sie sich auf diejenigen konzentrieren, die laut Modell am wahrscheinlichsten antworten, können Sie Ihren Profit pro Kunde steigern. Je mehr Kunden Sie aufnehmen, desto mehr steigt der Gesamtprofit – bis zu einem bestimmten Punkt –, selbst wenn der Ertrag pro Kunde allmählich zurückgeht, wenn Kunden mit geringeren Antwortquoten aufgenommen werden. Die Simulation erleichtert Ihnen die Entscheidung, wo die Grenze gezogen werden sollte, um den höchstmöglichen Profit zu realisieren.

Der Profitvergleich fasst dieses Ergebnis ausführlicher zusammen.

- **Zielgruppe 28 %.** In dieser Spalte wird das Ergebnis aus der gezielten Ansprache der 28 % von Kunden mit der vom Modell vorhergesagten höchsten Antwortquote angezeigt. Durch selektive Aufnahme der 840 Kunden mit der höchsten Antwortwahrscheinlichkeit können Sie Ihren Ertrag maximieren und einen erwarteten Profit von 3.450,00 US-Dollar erzielen.
- **Stichprobe 28 %.** Zeigt das Ergebnis an, wenn Sie eine zufällige Stichprobe derselben Größe ziehen würden. In diesem Fall würden Sie dieselbe Anzahl an Angeboten (840) versenden, doch die erwartete Antwortquote wäre geringer aus (entsprechend der Quote aus der Testkampagne), da Sie nicht von den zusätzlich vom Modell bereitgestellten Informationen profitieren. Dies führt zu einem Verlust, da die Gesamtantwortquote nicht hoch genug ist, um die Versandkosten für jedes Angebot wieder einzubringen.
- **Jeder.** Zeigt das Ergebnis an, wenn Sie 3.000 Kunden in die Stichprobe aufnehmen, was genau dem Ergebnis aus der Testkampagne entspricht. Auch dies ergibt einen Verlust, da Sie nicht die Vorteile des Modells nutzen.

An den Ergebnissen können Sie erkennen, wie durch Einsatz des Modells die Rücklaufquote verbessert und eine verlustträchtige Kampagne in eine rentable umgewandelt werden kann. Die Differenz zwischen einem Verlust von 90,00 US-Dollar und einem Profit von 3.450,00 US-Dollar stellt den durch Vorhersagemodelle gewonnenen Vorteil dar. Dieser Ertrag beruht auf den 3.000 Kunden aus der Testkampagne, doch durch Vergrößerung der Zielgruppe steigen auch die Vorteile entsprechend.

- ▶ Wählen Sie nach der Überprüfung der Simulationsergebnisse unter dem Diagramm „Vorhergesagter Profit“ die Option Zum Scoring verwenden aus, um die Registerkarte „Scoring“

auf der Grundlage des angegebenen Prozentsatzes (in diesem Fall 28 %) zu aktualisieren. [Für weitere Informationen siehe Thema Scores des Modells auf S. 13.](#)

Vorhersageregelm Modelle

Vorhersagemodelle bieten zwar meist die gewünschte Vorhersagekraft, doch in manchen Fällen sind möglicherweise etwas mehr Kontrolle und Einblick in die zugrunde liegende Logik des Modells vonnöten. In diesem Fall können Vorhersageregelm Modelle (interaktiv) verwendet werden. Mit Vorhersageregelm Modellen (interaktiv) können Sie manuelle Regeln zur Segmentierung Ihrer Daten erstellen oder automatisch Segmente mit hoher oder niedriger Konzentration des Werts Ihres ausgewählten Felds finden. Beispielsweise könnten Sie nach Kunden suchen, die eine positive Antwort auf Ihre Kampagne gegeben haben, und anschließend Segmente mit einer höheren Antwortwahrscheinlichkeit ermitteln.

- ▶ Kehren Sie zur Registerkarte „Modellierung“ zurück, klicken Sie auf Modell ändern und wählen Sie die Option Vorhersageregelm Modell aus.
- ▶ Wählen Sie *bank response data* als Datenquelle und *Antwort* als Zielfeld.
- ▶ Legen Sie 1 als die gesuchte Antwort fest, d. h., Sie suchen Kunden, die eine positive Antwort gegeben haben.
- ▶ Wählen Sie Erstellung beginnen.
- ▶ Wählen Sie Modell erweitern, um Segmente mit einer höheren Antwortwahrscheinlichkeit zu finden. Die drei besten Regeln zur Ermittlung der Segmente mit der höchsten Antwortwahrscheinlichkeit werden erstellt und angezeigt.
- ▶ Klicken Sie auf das Einschließen- bzw. Ausschließen-Symbol, um festzulegen ob die von den einzelnen Regeln ausgewählten Kunden eingeschlossen werden sollen oder nicht. Das zweite und dritte Regelsegment beispielsweise zeigen eine Antwortquote von etwas mehr als 8 %, was zwar höher ist als die Gesamtquote der Testkampagne von 3,7 %, jedoch nicht annähernd so hoch wie die Quote von 15,65 % für das erste Segment. Um zu ermitteln, wie sich das Einschließen oder Ausschließen dieser beiden Regeln auf den Gesamtprofit auswirken, können Sie die Profitsimulation in beide Richtungen laufen lassen und das Ergebnis vergleichen. Gehen Sie dazu wie folgt vor:
 - ▶ Setzen Sie alle drei Regelsegmente auf Einschließen und klicken Sie auf Auswerten.
 - ▶ Wählen Sie als zu erstellende Diagrammtypen Gewinndiagramm und Profitdiagramm.
 - ▶ Wählen Sie Maximalen Profit simulieren und legen Sie als Populationsgröße 3000, als Kosten 2.25 und als Umsatz 60 fest. Diese Angaben werden auch bei der Auswertung des Vorhersagemodells verwendet. Die gesuchte Antwort wird automatisch eingetragen.
 - ▶ Klicken Sie auf Ausführen und öffnen Sie die Registerkarte „Simulieren“, wenn die Diagramme erstellt wurden.

Der maximale simulierte Profit für dieses Modell beträgt 3.172,00 US-Dollar, etwas weniger als für das automatische Modell in diesem Fall. Wahlweise können Sie das Dialogfeld „Auswerten“ schließen, die Einstellung für das zweite und dritte Regelsegment auf Ausschließen ändern und

die Profitsimulation wiederholen, um herauszufinden, wie sich dadurch das Ergebnis ändert. Der wahre Vorteil von Vorhersageregelnmodellen (interaktiv) sind ihre erhöhte Kontrolle und die tieferen Einblicke, die sie bieten.

Testen des Modells

Nach der Modellbildung können Sie eine Auswahl der Kundendatensätze testen und prüfen, welche davon mit hoher Wahrscheinlichkeit auf Ihre Kampagne reagieren würden.

- ▶ Wählen Sie Testen auf der Registerkarte „Modellierung“, um eine Vorschau aufzurufen, die zeigt, wie bestimmte Datensätze beim Modell-Scoring verarbeitet werden. Das Dialogfeld „Testen“ öffnet sich.
- ▶ Erhöhen Sie im Bereich Datensätze die Rückgabezahl auf 100.
- ▶ Wählen Sie im Bereich Anzuzeigende Felder auswählen die Option *Kunden-ID*. Stellen Sie außerdem sicher, dass die Felder *\$XF-Antwort* und *\$XFRP-Antwort* ausgewählt sind, da diese die vorhergesagte Antwort (in diesem Fall 0 für nein, 1 für ja) und den Neigungs-Score für jeden Kunden angeben.

Sie können wahlweise weitere Felder einschließen, wie *Age* (Alter), *Wochen seit letztem Angebot* und *Has Bad Payment Record* (Hat schlechtes Zahlungsverhalten), die dabei helfen zu verstehen, wie oder warum ein bestimmter Kunde vom Modell auf eine bestimmte Weise gescort wird.

Scoren des Modells

Ein Modell zu scoren heißt, es auf die Datenquelle anzuwenden, für die Sie sich interessieren. Wenn Sie beispielsweise ein Modell generiert haben, das vorhersagt, welche Kunden mit der höchsten Wahrscheinlichkeit antworten, können Sie nun potenzielle Kunden bewerten (scoren), um zu ermitteln, welche davon in die Kampagne aufgenommen werden sollen. Im Dialogfeld „Auswerten“ haben Sie den Prozentsatz der Datensätze ermittelt, die im Hinblick auf Profitmaximierung gescort werden sollten, und den Prozentsatz angegeben, den Sie unter Berücksichtigung dieses Ergebnisses scoren wollten.

- ▶ Wählen Sie die Registerkarte „Scoren“.
- ▶ Achten Sie darauf, dass im Bereich Ausgabefelder angeben das Feld *Kunden-ID* ausgewählt ist.
- ▶ Geben Sie ein Ziel für die Scoring-Ergebnisse an. Dabei kann es sich entweder um eine Datenbanktabelle oder um eine Datei handeln. (Wenn Sie lediglich eine Vorschau der Scores anzeigen möchten, können Sie diesen Schritt überspringen.)
- ▶ Unter Scoren sollten die Einstellungen bereits so konfiguriert sein, dass die ersten 28 % der Datensätze auf der Grundlage der bei der Evaluierung des Modells ausgeführten Profitsimulation gescort werden. (Wenn Sie bei der Anzeige der Simulationsergebnisse nicht die Schaltfläche Zum Scoren verwenden gewählt haben, können Sie die Einstellungen hier aktualisieren.)
- ▶ Klicken Sie unten im Dialogfeld auf Vorschau, um ein Muster der Scoring-Ergebnisse anzuzeigen.

Da Sie sich für ein Scoring nach Neigung entschieden haben, werden nur Kunden mit der höchsten Neigung, zu antworten, mit aufgenommen. Durch die Verknüpfung der Kunden-ID mit Ihrer Kundendatenbank könnten Sie eine Mailingliste generieren, die auf diesem Ergebnis beruht.

- ▶ Klicken Sie auf Scoren. Das Modell wird auf die Kundendaten angewandt, und unten im Dialogfeld erscheint eine Bestätigungsmitteilung.

Beispiel für Cluster-Modellierung

Mit Clustermodellen können Sie Datensätze in Gruppen mit ähnlichen Merkmalen kategorisieren. So können Sie natürliche Gruppen in Ihren Daten auffinden. Beispielsweise können Sie Kunden nach demografischen Merkmalen oder Kaufverhalten segmentieren.

Für dieses Beispiel wird die Datendatei *bank_customer_data.txt* verwendet, die mit dem Produkt bereitgestellt und installiert wird. Eine fertig gestellte Version des Beispiels steht mit der Datei *bank_cluster_model.str* zur Verfügung. Genaueres zu Installation und Beispieldateien können Sie ggf. von Ihrem Administrator erfahren. [Für weitere Informationen siehe Thema Beispieldateien auf S. 6.](#)

- ▶ Kehren Sie zum Anwendungs-Startbereich zurück und erstellen Sie ein neues IBM® SPSS® Modeler Advantage-Projekt.
- ▶ Fügen Sie auf der Registerkarte „Daten“ im Feld „Projektdatenquellen“ eine neue Datenquelle mit der Bezeichnung *bank customer data* (Bankkundendaten) hinzu, wie in [Datenquellen einrichten auf S. 8](#) beschrieben.
- ▶ Klicken Sie auf der Registerkarte „Modellierung“ auf Modell ändern, wählen Sie die Option Clustermodell und klicken Sie auf Speichern.
- ▶ Wählen Sie als Datenquelle *bank customer data*.
- ▶ Erweitern Sie den Abschnitt Optionale Einstellungen. Bei Clustermodellen stehen zusätzlich optionale Einstellungen unter Optionen für automatische Cluster zur Verfügung, um ein Evaluationsfeld auszuwählen oder einen gewünschten Bereich für die zu findenden Cluster festzulegen.

Beim Aufbau eines Clustermodells werden im Hintergrund tatsächlich mehrere verschiedene Clustermodelle erstellt. Diese Modelle werden dann verglichen und anhand eines Qualitätsmaßstabs bewertet und das beste Modell wird für das automatische Cluster-Modell ausgewählt. Diese beiden Einstellungen für „Optionen für automatische Cluster“ bieten also gewisse Einflussmöglichkeiten darauf, welches der Modelle tatsächlich ausgewählt wird. Wenn hier ein Evaluationsfeld angegeben wird, wird das Modell ausgewählt, das die Werte des Evaluationsfelds am besten differenziert. Und wenn der gewünschte Bereich festgelegt wird, werden alle Modelle, die Cluster außerhalb des angegebenen Bereichs finden, verworfen.

- ▶ Wählen Sie für dieses Beispiel die Option Gewünschten Bereich für die Anzahl der gefundenen Cluster einstellen und verwenden Sie den Standard-Mindestwert 3 und den Höchstwert 15. Dadurch wird ein Cluster-Modell gewährleistet, das nicht eine zu hohe oder zu geringe Anzahl an Clustern aufweist.

- ▶ Beim Rest der optionalen Einstellungen gibt es keine Abweichungen von den anderen Modellierungstypen. Heben Sie die Auswahl der Option Daten automatisch bereinigen und zur zuverlässigen Modellerstellung aufbereiten auf, um das Beispiel für das automatische Clustering leichter verständlich zu machen.
- ▶ Minimieren Sie das Feld Optionale Einstellungen und erweitern Sie das Feld Cluster.

Im Feld „Cluster“ gibt es einen Abschnitt für manuelle Cluster und einen Abschnitt für automatische Cluster. Bei manuellen Clustern können Sie Cluster auf der Grundlage Ihrer Kenntnisse über die Daten definieren. In diesem Beispiel definieren wir ein manuelles Cluster für wertvolle Kunden (*high value customers*) und ein manuelles Cluster für junge, alleinstehende Kunden (*young single customers*). Gehen Sie dazu wie folgt vor:

- ▶ Klicken Sie auf das Symbol Neue Regel erstellen. Geben Sie der Regel den Namen *high value customers* (wertvolle Kunden), fügen Sie die folgenden beiden Ausdrücke hinzu und klicken Sie auf OK.

Months as a Customer > 12 (Länge der Kundenbeziehung in Monaten)

Income > 65000 (Einkommen)

- ▶ Erstellen Sie eine weitere neue Regel namens *young single customers* (junge, alleinstehende Kunden) mit folgenden Ausdrücken.

Age < 35 (Alter)

Marital Status = U (Familienstand)

- ▶ Die beiden neuen Regeln werden unter dem Abschnitt Manuelle Cluster aufgeführt. Um zu sehen, wie viele Kunden von diesen manuellen Clustern erfasst werden, klicken Sie auf Datensatzzahl in Clustern. Erweitern Sie anschließend die manuellen Cluster erneut und betrachten Sie die Spalte „Anzahl“, um zu sehen, wie viele Kunden jeweils für die einzelnen Kunden gefunden wurden. Sie sehen, dass 67 Kunden vom Typ „high value customers“ (wertvolle Kunden) und 33 Kunden vom Typ „young single customers“ (junge, alleinstehende Kunden) gefunden wurden.
- ▶ Klicken Sie dann auf Automatische Cluster suchen, um automatisch etwaige weitere Cluster zu finden, die interessant sein könnten. Erweitern Sie abschließend den Abschnitt Automatische Cluster. Wie Sie sehen, wurden drei automatische Cluster gefunden.
- ▶ Klicken Sie auf Automatische Cluster anzeigen, um weitere Details anzuzeigen. Der Ergebnis-Viewer für automatische Cluster wird angezeigt.

Auf der Seite *Modellübersicht* finden Sie grundlegende Informationen zum verwendeten Modellalgorithmus für die automatischen Cluster (im vorliegenden Beispiel *TwoStep*), die Anzahl der gefundenen automatischen Cluster sowie grundlegende Angaben zur Clusterqualität.

Außerdem stehen weitere Diagramme zur Verfügung. Sie können die Maus über verschiedene Bereiche im Diagramm bewegen, um weitere Details anzuzeigen und mit einigen der Diagramme zu interagieren (z. B. mit dem Diagramm *Prädiktoreinfluss*).

Mithilfe der Ansicht *Cluster* (letztes Diagramm) können Sie außerdem Cluster vergleichen, um nach interessanten Mustern zu suchen. In der Praxis würden Sie die automatischen Cluster untersuchen und ihnen dann aussagekräftigere Namen geben (indem Sie in der Ansicht *Cluster* auf die einzelnen Clusterbeschriftungen klicken). Nach der Aktualisierung der Namen der Clusterbeschriftungen werden die neuen Namen in allen anderen Diagrammen und auch auf der Registerkarte „Modellierung“ verwendet. Anschließend fahren Sie mit [Evaluation](#) und [Test](#) des Modells fort, und wenn Sie damit zufrieden sind, führen Sie abschließend ein Scoring durch.

Weitere Details zur Cluster-Modellierung finden Sie unter [Erstellen eines Clustermodells auf S. 54](#).

Beispiel für Assoziationsmodellierung

Assoziationsmodellierung ist eine wichtige Technik in der prädiktiven Analytik. Assoziationsmodelle ermöglichen eine Vorhersage darüber, welche Elemente höchstwahrscheinlich gemeinsam vorkommen und wie stark die Beziehung zwischen den Elementen ist. Beispielsweise können Sie auf der Grundlage des aktuellen Warenkorbs eines Kunden seinen nächsten wahrscheinlichen Einkauf ermitteln.

Für dieses Beispiel wird die Datendatei *retail_purchase_data.txt* verwendet, die mit dem Produkt bereitgestellt und installiert wird. Die Daten bestehen aus Kaufdaten zu Elektronikartikeln wie Fernsehern, Computern und Smartphones.

Eine fertig gestellte Version des Beispiels steht mit der Datei *retail_association_model.str* zur Verfügung. Genaueres zu Installation von Beispieldateien können Sie ggf. von Ihrem Administrator erfahren. [Für weitere Informationen siehe Thema Beispieldateien auf S. 6](#).

- ▶ Kehren Sie zum Anwendungs-Startbereich zurück und erstellen Sie ein neues IBM® SPSS® Modeler Advantage-Projekt.
- ▶ Fügen Sie auf der Registerkarte „Daten“ im Feld „Projektdatenquellen“ eine neue Datenquelle mit der Bezeichnung *retail_purchase_data* (Einzelhandelsdaten) hinzu. Klicken Sie auf Datei und wählen Sie *retail_purchase_data.txt*.
- ▶ Klicken Sie auf der Registerkarte „Modellierung“ auf Modell ändern, wählen Sie die Option Assoziationsmodell und klicken Sie auf Speichern.
- ▶ Wählen Sie als Datenquelle *retail_purchase_data* (Einzelhandelsdaten).
- ▶ Wählen Sie als Datenformat die Option Tabellenform. In einem Transaktionsdaten-Set würden nur Elemente angezeigt, die eingeschlossen wurden, und alle Elemente würden in einer einzigen Spalte angezeigt. In diesem Daten-Set dagegen befindet sich jedes Element in einer separaten Spalte mit dem Flag *true* (wahr) bzw. *false* (falsch), um anzuzeigen, ob es gekauft wurde oder nicht.
- ▶ Klicken Sie auf Modell erstellen. Sie erhalten eine Fehlermeldung, die besagt, dass die Erstellung gescheitert ist und dass möglicherweise die Schwellenwerte zu hoch sind. Schließen Sie die Fehlermeldung.

Bei der Suche nach Assoziationsregeln ist es normalerweise eine gute Idee, nach Regeln zu suchen, die generell anwendbar, zuverlässig und einfach sind.

- ▶ Erweitern Sie den Abschnitt „Optionale Einstellungen“. Ändern Sie unter „Optionen für Assoziationsregeln“ den ersten Schwellenwert von 10 zu 1 und senken Sie den zweiten Schwellenwert von 80 auf 1. Normalerweise würden Sie mit verschiedenen Schwellenwerten experimentieren und jeweils das Modell erneut erstellen, um einen ausgewogenen Mittelwert zwischen zu vielen und zu wenigen Regeln zu finden.

- ▶ Klicken Sie auf Modell erstellen.

In den Modellergebnissen steht jede Zeile für eine separate Assoziationsregel. Die erste Regel in der Liste besagt, dass Kunden, die ein Smartphone (*Smart Phone*) kaufen, mit hoher Wahrscheinlichkeit auch einen Großbildfernseher (*Big Screen TV*) kaufen. Die Spalte „Unterstützung“ zeigt an, dass ca. 15 % der Kunden (also etwa 1 von 6) das Smartphone (*Smart Phone*) gekauft haben, und die Spalte „Konfidenz“ zeigt, dass diese Kunden mit ca. 30 % Wahrscheinlichkeit auch einen Großbildfernseher (*Big Screen TV*) kaufen werden.

Selbst Sets mit nur wenigen Regeln (239 in diesem Fall) können also möglicherweise im geschäftlichen Kontext verwendet werden, um vorherzusagen oder sogar zu empfehlen, was die Kunden als Nächstes kaufen könnten.

- ▶ Klicken Sie als Nächstes auf Testen und anschließend auf Ausführen, um zu sehen, wie die Vorhersagen bzw. Empfehlungen erstellt werden.

In diesem Fall suchte das Modell in den Regeln nach den höchsten Konfidenzwerten, ermittelte, was der Benutzer bereits gekauft hatte, und erstellte auf der Basis dieser beiden Daten Empfehlungen dazu, was die Kunden vermutlich kaufen würden.

- ▶ Klicken Sie beispielsweise auf das Vorschausymbol neben dem ersten Datensatz, um zu sehen, dass der erste Kunde einen Großbildfernseher (*Big Screen TV*) und Lautsprecher (*Speakers*) erworben hat, jedoch keine weiteren Produkte. Schließen Sie die Vorschau.

Die Empfehlung für diesen Kunden lautet, ihm einen Standardfernseher (*Standard TV*) anzubieten, wie in der zweiten Spalte zu sehen ist. Der Wert der dritten Spalte, .29, bedeutet, dass eine 29%ige Wahrscheinlichkeit besteht, dass Kunden, die bereits einen Großbildfernseher (*Big Screen TV*) und Lautsprecher (*Speakers*) erworben haben, auch einen Standardfernseher (*Standard TV*) kaufen.

Wie können wir jedoch wissen, ob es sich hierbei um eine Anomalie handelt oder um einen vorhergesagten Trend, der allgemein anwendbar ist? Notieren Sie sich die Regelnummer (in diesem Beispiel 115), damit wir zurückgehen und uns die Regel genauer ansehen können.

- ▶ Schließen Sie das Dialogfeld „Testen“.
- ▶ Blättern Sie durch die Modellergebnisse, bis Sie Regel Nummer 115 sehen. Beachten Sie, dass Sie bei großen Modellen, die wesentlich mehr Ergebnisseiten aufweisen, auch auf Regel anhand der ID suchen klicken und die Regelnummer eingeben können.
- ▶ Klicken Sie auf den Pfeil neben einer der Ergebnisspaltenüberschriften und wählen Sie folgende Optionen:
Spalten > Instanzen

Die Spalte „Instanzen“ wird angezeigt. Angesichts der Anzahl der Instanzen ist es klar, dass es sich nicht um eine Anomalie handelt, denn Regel 115 beruht auf 58 Kunden, die alle drei Produkte gekauft haben. Wir können somit getrost die Regel als Grundlage für zukünftige Empfehlungen verwenden.

Wir haben eine einzelne Regel betrachtet, um zu sehen, wie gut sich diese als Empfehlung eignet. IBM® Analytical Decision Management bietet jedoch auch leistungsstarke Optionen zur Anpassung des Modells, um es für unsere jeweiligen konkreten Zwecke relevanter zu machen. Beispielsweise möchten wir vorhersagen, welche Kunden am Kauf eines Großbildfernsehers (*Big Screen TV*) interessiert sind:

- ▶ Klicken Sie auf Filter anwenden. Dadurch wird ein Dialogfeld angezeigt, in dem Sie Filter definieren können, sodass nur Regeln, die den Filtern entsprechen, im Modell angezeigt werden.
- ▶ Wählen Sie die Option Filter aktivieren und wählen Sie *Big Screen TV* (Großbildfernseher) als Element aus.
- ▶ Klicken Sie auf Speichern und dann noch einmal auf Speichern, um zu den Modellergebnissen zurückzukehren.

Die Anzahl der Regeln ist von 239 auf 34 gesunken.

Sie können auch sicherstellen, dass nur Regeln verwendet werden, denen eine angemessene Anzahl an Kunden zugrunde liegt. Dazu könnten Sie beispielsweise Regeln ausschließen, die auf dem Verhalten von weniger als 10 Kunden beruhen:

- ▶ Klicken Sie auf den Pfeil neben der Spaltenüberschrift „Instanzen“ und wählen Sie die Option Sortieren: Aufsteigend aus.
- ▶ Wählen Sie in der Spalte „Ausschließen“ alle Regeln mit weniger als 10 Instanzen aus.
- ▶ Beachten Sie: Nun wurde zwar ein Großteil der Regeln gefiltert und ausgeschlossen, diese Regeln befinden sich derzeit jedoch noch immer im Modell. Um sie dauerhaft zu entfernen, sodass sie nicht verwendet werden, wenn das Modell gescort wird, klicken Sie auf Neues Modell erstellen – ausgeschlossene und gefilterte Regeln löschen und dann auf OK.

Die Anzahl der ausgegebenen Regeln wurde auf 20 reduziert, die jeweils auf mindestens 10 Instanzen beruhen und alle *Big Screen TV* (Großbildfernseher) vorhersagen.

- ▶ Um zu sehen, wie viele dieser Regeln angewendet werden, klicken Sie auf Testen und dann auf Ausführen. *Big Screen TV* (Großbildfernseher) wird nun für mehrere weitere Kunden empfohlen und einige weisen ein höheres Konfidenzniveau auf als andere.
- ▶ Schließen Sie das Dialogfeld „Testen“.
- ▶ Erweitern Sie den Abschnitt „Optionale Einstellungen“, sodass die speziell für Assoziationsmodelle geltenden Scoring-Optionen angezeigt werden.

Standardmäßig ist die maximale Anzahl an Vorhersagen 3. Dies beruht auf den Werten mit der höchsten Konfidenz. Alternativ können Sie das Regelkriterium ändern, um Regeln auf der Grundlage der höchsten Anzahl an Instanzen oder anhand des höchsten Werts für *Lift* (Anstieg der Wahrscheinlichkeit, dass ein bestimmter Artikel gekauft wird) auszuwählen. Außerdem ist es sinnvoll, sicherzustellen, dass Kunden, die bereits einen Großbildfernseher (*Big Screen*

TV) gekauft haben, kein weiterer angeboten ist. Wählen Sie dazu die Option Sicherstellen, dass Vorhersagen nicht vorhanden aus.

Die Standardoptionen für das Scoring eignen sich für die meisten Situationen. In einigen Fällen kann jedoch die zusätzliche Flexibilität, die diese Optionen bieten, von Vorteil sein.

Der nächste Schritt im Vorgang bestünde darin, das Modell mithilfe der Registerkarte „Scoring“ zu scoren. [Für weitere Informationen siehe Thema Scoring von Modellen in eine Datenbanktabelle, eine Datei oder einen Cognos BI-Server in Kapitel 7 auf S. 75.](#)

Weitere Informationen zur Assoziationsmodellierung finden Sie unter Erstellen eines Assoziationsmodells auf S. 56.

Navigation im Arbeitsbereich

Anwendungen starten

Im Anwendungs-Startbereich werden konfigurierbare Verknüpfungen zu Ihren installierten Anwendungen angezeigt. Jedes Feld repräsentiert eine installierte Anwendung. Klicken Sie auf das Dropdown-Menü, um aus den zuletzt gespeicherten Projekten, Modellen oder Regeln zu wählen. Außerdem können Sie eine oder mehrere eigene Verknüpfungen hinzufügen, die auf bestimmte beschriftete Versionen verweisen.

- ▶ Die URL des Anwendungs-Startbereichs lautet *http://hostname:port/DM*, wobei *hostname* die IP-Adresse oder der Name des Repository-Rechners und *port* die Portnummer des Anwendungsservers ist. Falls Sie diese Informationen nicht kennen, wenden Sie sich an Ihren Administrator. Sie können auch in Ihrem Browser ein Lesezeichen auf diese Seite setzen oder eine Verknüpfung auf Ihrem Desktop einrichten.
- ▶ Zur Erstellung eines neuen Projekts, eines neuen Modells bzw. einer neuen Regel wählen Sie aus der Dropdown-Liste für die gewünschte Anwendung die Option Neu aus und klicken Sie dann auf Wechseln.
- ▶ Um ein neues Projekt, ein neues Modell bzw. eine neue Regel zu öffnen, wählen Sie eine kürzlich verwendete Version bzw. Verknüpfung aus der Dropdown-Liste im Anwendungsfeld aus, oder wählen Sie Durchsuchen in der Dropdown-Liste aus, um ein anderes Objekt bzw. eine andere Version auszuwählen. Um rasch den Repository-Standort eines Objekts in der Liste anzuzeigen, bewegen Sie den Mauszeiger im Dropdown-Menü auf das gewünschte Objekt.
- ▶ Um eine neue Verknüpfung hinzuzufügen, die immer in der Dropdown-Liste verfügbar ist, klicken Sie auf das dreieckige Pfeilsymbol rechts oben im Anwendungsfeld. Sie können die Anzahl der anzuzeigenden Verknüpfungen festlegen.

Andere Methoden zum Öffnen einer Anwendung

- Kopieren oder Schreiben Sie die URL in die Adresszeile Ihres Browsers oder fügen Sie sie der Liste der gespeicherten Lesezeichen bzw. Favoriten hinzu.
- Erstellen Sie eine Desktop-Verknüpfung, die auf die URL verweist.
- Setzen Sie auf der Intranetsite oder dem Portal Ihres Unternehmens einen Link auf die Anwendungs-URL.

Den Startbereich anpassen

Mit den Steuerelementen rechts oben im Startbereich können Sie Anwendungen hinzufügen und zwischen Feld- und Listenansicht umschalten. Sie können die Felder durch Klicken und Ziehen umordnen.

- ▶ Zum Hinzufügen eines neuen Anwendungsfelds wählen Sie in der rechten oberen Ecke des Anwendungs-Startbereichs die Option Anwendung hinzufügen. (Falls keine zusätzlichen Anwendungen verfügbar sind, ist dieser Link nicht verfügbar.)

Eine Liste verfügbarer Anwendungen wird angezeigt. Optional können Sie Verknüpfungen zu einer oder mehreren beschrifteten Versionen erstellen.
- ▶ Sie können die Anordnung der Anwendungsfelder im Startbereich ändern, indem Sie auf die Titelleiste eines Felds klicken, die Maustaste gedrückt halten und das Feld an seine neue Position ziehen.
- ▶ Zum Entfernen einer Anwendung klicken Sie auf das Lösch-Symbol (rotes X) in der rechten oberen Ecke des Anwendungsfelds. Alle Verknüpfungen, die auf diese Anwendung verweisen, werden ebenfalls gelöscht. (Das Lösch-Symbol ist nicht verfügbar, wenn das Anwendungsfeld auf der Seite vom Administrator gesperrt wurde.)
- ▶ Ungespeicherte Projekte, Modelle und Regeln werden in der Galerie des Startbereichs aufgeführt. Diese Galerie enthält Modelle, die sich noch im Aufbau befinden, sowie fertige, aber noch nicht gespeicherte Modelle. [Für weitere Informationen siehe Thema Galerie auf S. 22.](#)

Projektkonfiguration

Beim Starten einer Anwendung, die eine Registerkarte „Definieren“ mit einem Dimensionsbaum aufweist, werden Sie dazu aufgefordert, mindestens eines der verfügbaren Elemente (der Hierarchien) auszuwählen, das in der Anwendung verwendet werden soll, wenn die ausgewählte Anwendung so konfiguriert wurde, dass sie mehrere Dimensionshierarchieelemente aufweist. Der Dimensionsbaum auf der Registerkarte „Definieren“ gibt den Bereich der möglichen Entscheidungen oder Empfehlungen an, die zurückgegeben werden können. Wenn eine Anwendung mehrere Dimensionshierarchien aufweist, sind diese auf separaten Registerkarten des Dimensionsbaumbereichs auf der Registerkarte „Definieren“ enthalten. Weitere Informationen zur Registerkarte „Definieren“ und deren Dimensionsbaum finden Sie unter .

- ▶ Im Dialogfeld für die Projektkonfiguration werden alle verfügbaren Dimensionshierarchien angezeigt, die in der Anwendungsvorlage definiert sind. Wählen Sie mindestens eine Dimensionshierarchie (ein Element) aus, die Sie zum Definieren des Projekts verwenden möchten. Eine Beschreibung für jedes Element ist in Abhängigkeit davon verfügbar, ob diese vom Anwendungsentwickler in der Anwendungsvorlage definiert wurde.

Beim Namen des Elements handelt es sich um den Namen der obersten Dimension. Dimensionen können nur einem Element angehören.

- ▶ Wenn Sie mehrere Elemente auswählen, können Sie die Reihenfolge angeben, in der diese in der Anwendung angezeigt werden. Ferner können Sie angeben, welche Elemente voneinander abhängig sein können.

Die Reihenfolge, die Sie hier angeben, ist wichtig. Die Elemente werden auf geordneten Registerkarten angezeigt. Dabei kann auf Hierarchiemitglieder vorangegangener Registerkarten auf nachfolgenden Registerkarten Bezug genommen werden. Ein Verweis auf nachfolgende Registerkarten ist allerdings nicht möglich. Beispielsweise können sich Regeln auf einer zweiten und dritten Registerkarte auf Hierarchiemitglieder auf der ersten Registerkarte beziehen. Regeln

auf der ersten Registerkarte können sich jedoch nicht auf Mitglieder auf der zweiten oder dritten Registerkarte beziehen. Ebenso ist es nicht möglich, dass sich Regeln auf der zweiten Registerkarte auf Mitglieder auf der dritten Registerkarte beziehen.

- ▶ Klicken Sie auf OK, wenn Sie die Arbeit im Dialogfeld für die Projektkonfiguration beendet haben.

Die Hauptseite

Die Anwendung stellt auf der Hauptseite die verschiedenen Schritte des Arbeitsablaufs in Form von Symbolen dar. Klicken Sie auf ein Symbol, um zu dem betreffenden Schritt zu gelangen.

Daten. Die von der Anwendung zu verwendenden Daten-Sets auswählen.

Modell. Erstellen Sie anhand analytischer Daten Modelle und bewerten Sie diese mittels Prädiktoreinfluss-, Verteilungs- und Gains-Diagrammen. Dadurch können Sie in den Daten verborgene Muster aufdecken und so Ihre Entscheidungsprozesse weiter verbessern.

Scoren. Scoren Sie das Modell.

Galerie

Die Galerie wird im Anwendungs-Startbereich angezeigt. Sie führt ungespeicherte Arbeitsvorgänge auf, wie etwa ein Modell, das über Nacht erstellt werden soll: Sie können den Vorgang gegen Büroschluss starten, den Browser schließen und am nächsten Morgen das Ergebnis abrufen. Sie können die Galerie auch benutzen, um Ihre Arbeit wiederherzustellen, wenn der Browser unbeabsichtigt geschlossen wurde. Zu den Objekten in der Galerie können Projekte, Regeln, Modelle, Datenscans, Auswertungsvorgänge und je nach Anwendung andere langlaufende Aufgaben gehören.

- ▶ Falls der Name des Objekts unterstrichen ist, können Sie zum Öffnen darauf klicken. Wenn z. B. ein Datenscan noch nicht abgeschlossen wurde, können Sie auf dessen Titel in der Galerie klicken, um die Registerkarte „Daten“ zu öffnen.
- ▶ Wenn Sie ein in der Galerie angezeigtes Objekt nicht mehr benötigen, klicken Sie zum Entfernen auf das Symbol „Löschen“. Es wird dann eine Bestätigungsaufforderung angezeigt, in der das dauerhafte Löschen des Objekts bestätigt werden muss. Zum Löschen mehrerer Objekte aus der Galerie klicken Sie auf Galerie löschen. Daraufhin wird das Dialogfeld „Galerie löschen“ angezeigt, in dem Sie die Registerkarten auswählen können, aus denen alle Objekte gelöscht werden sollen.
- ▶ Auf den Registerkarten für Projekte und Simulationen in der Galerie können Sie über das Dropdownmenü Erstellt in die auf der Registerkarte für jede Anwendung angezeigten Objekte filtern. Wenn Sie z. B. in einer bestimmten Anwendung arbeiten, möchten Sie möglicherweise nicht, dass für andere Anwendungen vorgesehene Objekte ebenfalls angezeigt werden.

Voreinstellungen anpassen

Es gibt eine Reihe von Voreinstellungen, mit denen Sie IBM® Analytical Decision Management an Ihre Bedürfnisse anpassen können. Diese Anpassung besteht hauptsächlich aus der Festlegung eigener Standardeinstellungen, damit bei Erstellung, Test und Speicherung Ihrer Arbeit nicht jedes Mal dieselben Vorgänge abgewickelt werden müssen. Die von Ihnen festgelegten Voreinstellungen werden in alle Anwendungen übernommen, die Sie benutzen, gelten aber nicht für andere Personen.

Um auf die Benutzervoreinstellungen zuzugreifen, klicken Sie in einer beliebigen Anwendung in der Symbolleiste auf das Symbol Voreinstellungen konfigurieren.

Abbildung 3-1
Symbol „Voreinstellungen konfigurieren“



Sie können für die verschiedenen Bereiche wie in den folgenden Abschnitten beschrieben Voreinstellungen vornehmen.

Allgemeine Einstellungen

Standardanzahl der kürzlich geöffneten Objekte in den Anwendungs-Verknüpfungen. Wählen Sie die Anzahl der Objekte, die im Dropdown-Menü des Anwendungs-Startbereichs verfügbar sein sollen.

Datumsformat. Legen Sie das Format fest, das zur Anzeige und Aufzeichnung von Datumsangaben verwendet wird.

Zeitformat. Legen Sie das Format fest, das zur Anzeige und Aufzeichnung von Zeitangaben verwendet wird.

Währung. Wählen Sie die Standardwährung für neue Projekte. Beachten Sie, dass dadurch bestehende Daten mit anderer Währung nicht überschrieben werden.

Dezimalstellen. Wählen Sie die gewünschte Anzahl an Dezimalstellen aus.

Anzahl der in der Vorschau darzustellenden Zeilen. Legen Sie die Höchstzahl der Datensätze fest, die in einer Vorschau angezeigt werden, z. B. bei der Überprüfung einer Datenquelle nach der Importierung.

Modelleinstellungen

Standardmäßige Modellerstellung. Legen Sie fest, ob die Anwendung das Modell automatisch anhand der verfügbaren Daten erstellen soll (Vorhersagemodell), ob ein Modell verwendet werden soll, das Segmente ausfindig macht, die in Bezug auf ein bestimmtes Resultat eine höhere oder geringere Wahrscheinlichkeit aufweisen (Vorhersageregelmotell (interaktiv)), ob ein Modell verwendet werden soll, das Datensätze in eine bestimmte Anzahl von Clustern unterteilt (Clustermodell), oder ob ein Modell verwendet werden soll, das vorhersagt, welche

Elemente höchstwahrscheinlich nah beieinander liegen und wie stark die Beziehung zwischen den Elementen ist (Assoziationsmodell). Im Dropdownmenü sind nur Modelltypen verfügbar, die für die Anzeige in der Anwendung konfiguriert wurden.

Daten automatisch bereinigen und zur zuverlässigen Modellerstellung aufbereiten. Wenn Sie diese Option wählen, können Sie einen Großteil der Datenvorbereitung von der Anwendung übernehmen lassen. Z. B. Datenprobleme beheben, weniger nützliche Daten herausfiltern oder neue Attribute ableiten.

Sie können wählen, ob für den Bereinigungsprozess Geschwindigkeit und Genauigkeit gleichermaßen berücksichtigt oder einem der beiden Priorität gegeben werden soll, z. B., wenn Sie mit sehr großen Daten-Sets arbeiten oder ein schnelles Ergebnis benötigen.

Beschränken Sie die Zeit zur Erstellung eines einzelnen Modells. Legen Sie mit dieser Option die Standardzeit fest, die zur Erstellung eines Modells aufgebracht wird. Dies kann insbesondere bei der Verarbeitung sehr großer Daten-Sets nützlich sein und sich deutlich auf die zur Modellerstellung benötigte Zeit auswirken.

Test- und Simulationseinstellungen

Standardanzahl zu testender Datensätze. Legen Sie die Standardanzahl der für Testläufe heranzuziehenden Datensätze fest.

Öffnen und Speichern von Projekten, Modellen und Regeln

Sie können Ihre Arbeit im IBM® SPSS® Collaboration and Deployment Services Repository öffnen bzw. dort speichern oder in das lokale Dateisystem herunterladen. Sie können auch IBM® SPSS® Modeler-Stream-Dateien aus dem lokalen Dateisystem hochladen und im IBM SPSS Collaboration and Deployment Services Repository speichern. Die Speicherung im IBM SPSS Collaboration and Deployment Services Repository bringt wichtige Vorteile wie Unterstützung mehrerer Versionen, Beschriftung und Sicherheit mit sich. Abhängig von der lokalen Konfiguration werden möglicherweise nicht alle Optionen unterstützt.



Erstellt ein neues Modell.



Speichert das aktuelle Projekt, das Modell oder die Regel am letzten bekannten Ort im IBM SPSS Collaboration and Deployment Services Repository. Optional können Sie beim Speichern eine neue Beschriftung übernehmen.



Speichert und beschriftet das aktuelle Projekt, das Modell oder die Regel, oder speichert diese unter einem neuen Namen. Gegebenenfalls können Sie zwischen der Speicherung im IBM SPSS Collaboration and Deployment Services Repository und dem Herunterladen ins lokale Dateisystem wählen.



Öffnet ein Projekt, ein Modell oder eine Regel aus dem IBM SPSS Collaboration and Deployment Services Repository und/oder dem lokalen Dateisystem.



Lädt ein Projekt, ein Modell oder eine Regel aus dem lokalen Dateisystem hoch. Der Zugriff auf das Dateisystem wird nicht für alle Anwendungen unterstützt. Wenn Sie einen SPSS Modeler-Stream hochladen, können Sie auswählen, ob die Datei geöffnet oder im Repository gespeichert werden soll.

Speichern im IBM SPSS Collaboration and Deployment Services Repository

Das IBM® SPSS® Collaboration and Deployment Services Repository bringt wichtige Vorteile wie Unterstützung mehrerer Versionen, Beschriftung und Sicherheit mit sich. Jedes Mal, wenn Sie speichern, erstellen Sie eine neue Version, und Sie können jederzeit zu früheren Versionen zurückkehren. Außerdem können Sie Versionen durch Beschriftungen verwalten und mithilfe von Sicherheitseinstellungen festlegen, wer ein Objekt aufrufen, anzeigen oder entfernen kann.

In der Baumansicht im linken Fensterbereich sind die Verzeichnisse des IBM SPSS Collaboration and Deployment Services Repository aufgeführt, zu deren Anzeige Sie berechtigt sind. Zum Suchen können Sie den Verzeichnisbaum oder die Suchfunktion nutzen. Sie können auch die Dateiliste nach den im linken unteren Bereich verfügbaren Beschriftungen filtern. Zum Beispiel können Sie auf eine Beschriftung namens *Produktion* klicken, um nur Dateien mit dieser Beschriftung anzuzeigen.

Klicken Sie neben einer Datei ggf. auf das Vorschau- oder Informationssymbol, um weitere Details zu dieser anzuzeigen.

Ordner. Zeigt die verfügbaren Verzeichnisse an und ermöglicht es Ihnen, die Verzeichnisstruktur zu durchsuchen. Wählen Sie den Ordner zum Speichern aus. Sie können auch einen neuen Ordner erstellen oder einen vorhandenen Ordner umbenennen. Verwenden Sie hierzu die Symbole über der Ordnerliste.

Nach Beschriftung filtern. Beschränkt die Anzeige auf Objekte mit der gewählten Beschriftung. Beim Durchsuchen von Verzeichnissen werden nur Objekte mit dieser Beschriftung angezeigt.

Beim ist darüber hinaus das Dropdown-Menü Speichern als verfügbar. Wählen Sie im Dropdown-Menü die Option Einzelner Regelsatz oder Einzelregeln aus, um anzugeben, wie die Regel gespeichert werden soll. Wenn nur referenzierte Regeln verfügbar sind, ist das Dropdown-Menü deaktiviert und die Regel wird als einzelner Regelsatz gespeichert, da referenzierte Regeln nicht einzeln gespeichert werden können.

Wenn Sie die Option für einzelne Regeln auswählen, wird eine Registerkarte angezeigt, in der Sie Details für jede einzelne Regel eingeben können, die Sie speichern möchten.

Name. Geben Sie einen eindeutigen Namen für das Objekt ein, das Sie speichern.

Schlüsselwörter. Geben Sie Schlüsselwörter zur Erleichterung der Suche nach dem gespeicherten Objekt ein. Sie könnten z. B. „Absatz“ und „1-2009“ wählen.

Beschreibung. Geben Sie eine ausführlichere Beschreibung des gespeicherten Objekts ein. Diese könnte z. B. „Analyse des Einzelhandel-Absatzes; 1. Quartal 2009“ lauten.

Beschriftung(en). Falls Sie mehrere Versionen desselben gespeicherten Objekts erstellen, können Sie jeder eine eigene, eindeutige Beschriftung geben. Wählen Sie entweder eine vorhandene Beschreibung aus oder erstellen eine neue.

Arbeiten mit Beschriftungen

Beschriftungen werden dazu verwendet, eine bestimmte Version eines Projekts, eines Modells oder einer Regel zu kennzeichnen. Wenn z. B. eine Regel in mehreren Projekten zum Einsatz kommt, können Sie festlegen, dass die Version mit der Beschriftung *Produktion* benutzt werden soll. Wenn eine neuere Version der Regel erstellt wird, ändert sich die in Ihrem Projekt benutzte nicht, sofern nicht außerdem die Beschriftung *Produktion* zu dieser Version verschoben wird.

Bei der Referenzierung gemeinsam genutzter Objekte wird die Verwendung einer bestimmten beschrifteten Version sehr empfohlen – insbesondere in einer Produktionsumgebung –, da spätere Änderungen an der neuesten Version sonst zu unerwarteten Konsequenzen führen können. In bestimmten Fällen, z. B. bei der Referenzierung einer Regel innerhalb einer Regel, muss eine beschriftete Version verwendet werden. [Für weitere Informationen siehe Thema Regeln exportieren und wiederverwenden in Kapitel 5 auf S. 46.](#)

Wenn keine Version festgelegt wird, wird die jeweils neueste verwendet. Das bedeutet, dass spätere Änderungen an der Regel automatisch übernommen werden. (Stellen Sie sich *NEUESTE* als eine spezielle Beschriftung vor, die bei der Erstellung einer neuen Version automatisch auf diese übertragen wird.)

Jedes Mal, wenn Sie ein neues Objekt oder eine Version speichern, können Sie die Beschriftungen für diese Version auswählen. [Für weitere Informationen siehe Thema Öffnen und Speichern von Projekten, Modellen und Regeln auf S. 24.](#)

Einschränkungen für Beschriftungsnamen Bei der Definition von Beschriftungen für IBM® Analytical Decision Management dürfen die Zeichen **?** **#** **&** und **:** (Doppelpunkt) nicht verwendet werden, da die Anwendung sonst Referenzen auf Objekte nicht korrekt interpretieren könnte, die diese Beschriftungen verwenden. Bei den folgenden Beispiel handelt es sich um ungültige Beschriftungsnamen:

Testen & Bereitstellen

BEREITSTELLEN:PHASE1

Projekte sperren

Wenn Sie ein Projekt, ein Modell oder eine Regel zum Bearbeiten öffnen, wird das Objekt im IBM® SPSS® Collaboration and Deployment Services Repository gesperrt, so dass andere Benutzer es nicht bearbeiten können. Ein Kontrollkästchen oben rechts im Arbeitsbereich weist

darauf hin. Sie können die Markierung dieses Kontrollkästchens entfernen, um anderen die Bearbeitung des Objekts zu erlauben.

- Wenn ein Projekt gesperrt wurde, können es andere Benutzer zwar noch öffnen, aber keine Änderungen speichern.
- Wenn Sie ein gesperrtes Projekt schließen (z. B. durch Schließen des Browsers), wird die Sperrung aufgehoben und andere Benutzer können das Objekt bearbeiten.
- Wenn Sie ein gesperrtes Projekt wieder freigeben, ohne es zu schließen (indem Sie die Markierung des Kontrollkästchens entfernen), werden Sie ggf. zum Speichern der Änderungen aufgefordert. Ungespeicherte Änderungen gehen verloren, wenn das Objekt freigegeben wird.
- Nach der Freigabe eines Projekts können Sie den Arbeitsbereich weiterhin bedienen, aber die Änderungen nicht mehr speichern.

Objekteigenschaften

Wenn Sie einen Regelsatz, ein Modell oder Projekt bearbeiten, können Sie diesem Schlüsselwörter und eine Beschreibung hinzufügen, um das Wiederauffinden zu einem späteren Zeitpunkt zu erleichtern. Zur Anzeige des Dialogfelds „Eigenschaften bearbeiten“ klicken Sie auf das Eigenschaften-Symbol der Anwendung.

Abbildung 3-2
Eigenschaften-Symbol



Schlüsselwörter. Geben Sie Schlüsselwörter zur Erleichterung der Suche nach dem gespeicherten Objekt ein. Sie könnten z. B. „Absatz“ und „Q1-2012“ wählen.

Beschreibung. Geben Sie eine ausführlichere Beschreibung des gespeicherten Objekts ein. Diese könnte z. B. „Analyse des Einzelhandel-Absatzes; erstes Quartal 2012“ lauten.

Verwaltung von Datenquellen

Auf der Registerkarte „Daten“ können Sie Datenquellen für Analysen, Simulationen und Tests sowie fürs Scoring und andere Verwendungszwecke definieren. Daten-Sets können von Ihrem Administrator konfiguriert oder von Ihnen selbst hinzugefügt werden.

Die Registerkarte ist in drei große Bereiche aufgeteilt:

- **Projektdatenmodell.** Definiert die von der Anwendung benötigten Felder. Alle anderen Daten-Sets werden relativ zu dieser Quelle abgebildet. Im Abschnitt Datenquellenfelder werden die Eingabefelder und -typen, die Ihr Projekt verwenden wird, aufgeführt. Bei Bedarf können Sie im Abschnitt Zusatzfelder Ausdrücke oder Modellausgaben aus anderen Datenquellen hinzufügen, um Ihr Projektdatenmodell zu erweitern.
- **Projektdatenquellen.** Listet die Datenquellen auf, die für die Benutzung mit dem aktuellen Projekt, dem Modell oder der Regel gespeichert wurden.
- **Eigene Datenquellen.** Listet die von Ihnen definierten oder gespeicherten Datenquellen auf. Daten-Sets aus dieser Liste können in andere Projekte oder aus diesen kopiert werden, sodass Daten anwendungsübergreifend gemeinsam genutzt werden können.

Projektdatenmodell

Das Projektdatenmodell fungiert als Vorlage für die Eingabefelder und -typen, die Ihr Projekt verwenden wird.

Datenquelle. Zeigt die für die Anwendung ausgewählte Datenquelle an.

Schlüssel. Wenn es sich bei der für die Anwendung ausgewählten Datenquelle um eine Datenquelle vom Typ „Application-Ansicht“ handelt und Application-Ansichts-Schlüssel in der Datenquellentabelle definiert wurden, bietet dieses Dropdown-Feld eine Liste mit Schlüsseln, aus denen Sie auswählen können. Application-Ansichts-Schlüssel sind Teil der in IBM® SPSS® Collaboration and Deployment Services Deployment Manager erstellten IBM® SPSS® Collaboration and Deployment Services Enterprise View. Wenn keine Schlüssel für die Application-Ansichts-Tabelle definiert wurden oder es sich bei der für die Anwendung ausgewählten Datenquelle nicht um eine Datenquelle vom Typ „Application-Ansicht“ handelt, wird stattdessen im Dropdown-Feld Schlüssel eine Liste aller Felder für operative Daten aus der Datenquelle zur Auswahl bereitgestellt.

Das ausgewählte Schlüsselfeld wird verwendet, um etwaige zusätzliche Datenquellentabellen, die im Abschnitt Zusatzfelder definiert wurden, zuzuordnen. Während der Laufzeit dienen die Schlüssel zur Verbindung der primären Projektdatenmodell-Tabelle mit etwaigen sekundären Tabellen. Es wird davon ausgegangen, dass das Schlüsselfeld im Projektdatenmodell eindeutig ist (z. B. eine Kunden-ID, eine Schadensmeldungs-ID usw.). Jede sekundäre Tabelle muss eine übereinstimmende Spalte enthalten.

Das Schlüsselfeld ist optional, wenn keine zusätzlichen Tabellen für die Datenquelle definiert wurden.

Operativ. Die Markierung dieses Kästchens weist darauf hin, dass die Daten im Feld für die Verwendung in Regeln und Ausdrücken sowie für Vorhersagen verfügbar sind. Entfernen Sie die Markierung, falls das Feld für die Analyse als Zielfeld benutzt werden soll. Wenn Sie z. B. eine Anfrage durchführen, um zu beurteilen, wie gut Kunden auf eine Direktwerbungskampagne ansprechen würden, dann wäre das Ergebnisfeld analytisch, da Ihnen die Antwortwerte nicht im Voraus bekannt sind.

Anmerkung: Bei Auswahl einer IBM SPSS Collaboration and Deployment Services Enterprise View-Datenquelle als Projektdatenmodell sind die Kästchen dieser Spalte deaktiviert.

Feldname. Zeigt den Namen jedes Felds wie in der Datenquelle aufgeführt an sowie ein Symbol, das auf den Speichertyp hinweist.

Messniveau. Zeigt das Messniveau jedes Felds der Datenquelle an. Falls Sie eines der Messniveaus ändern, müssen Sie den Datenscan aktualisieren, damit die Datenkompatibilität gewährleistet bleibt. [Für weitere Informationen siehe Thema Messniveaus auf S. 31.](#)

Werte. Zeigt die Werte für die Felder der Datenquelle an. Z. B. Minimal- und Maximalwerte in einem bestimmten Bereich. Wenn Sie bei Flag-Feldern den Mauszeiger über den Wert halten, wird der für dieses Feld definierte „wahre“ und „falsche“ Wert in einem Tool Tip angezeigt.

Zusatzfelder hinzufügen/bearbeiten. In diesem Abschnitt können Sie bei Bedarf Ausdrücke oder Modellausgaben aus anderen Datenquellen hinzufügen, um Ihr Projektdatenmodell zu erweitern, indem Sie auf Zusatzfelder hinzufügen/bearbeiten klicken.

Projektdatenquellen

Projektdatenquellen können sowohl vom Administrator vordefinierte als auch von Benutzern hinzugefügte Datenquellen umfassen. Der Administrator hat auch die Möglichkeit, eine oder mehrere Datenquellen zu sperren, so dass Benutzer sie nicht verändern oder entfernen können, oder alle Datenoptionen zu sperren, so dass Benutzer keine neuen Datenquellen erstellen können.

Falls die Eingabedaten einer Datenquelle denen des Projektdatenmodells nicht direkt entsprechen, können Sie erstere auf letztere abbilden und die Diskrepanz beheben. Wenn beispielsweise das Projektdatenmodell ein Feld namens *Kauf* mit den Werten *Ja* und *Nein* enthält (Messniveau „Flag“), müssen alle benutzten Datenquellen über ein vergleichbares Feld verfügen, das entsprechend abgebildet werden kann.

Name. Gibt den Namen der Datenquelle an und zeigt ein Symbol, das auf den Dateityp hinweist.

Anzahl der Datensätze. Klicken Sie auf das Symbol in dieser Spalte, um die Datensatzzahl für jede beliebige Datenquelle in der Tabelle anzuzeigen. Die Anzahl wird neben dem Symbol angezeigt.

Vorschau. Ermöglicht Ihnen die Anzeige einer Stichprobe der in der Quelle enthaltenen Daten. [Für weitere Informationen siehe Thema Vorschau von Daten auf S. 31.](#)

Kompatibel. Zeigt entweder einen Hinweis an, dass die Datenquelle als Projektdatenmodell benutzt wird, oder andernfalls einen grünen, orangenen oder roten Kreis, der die Kompatibilität der Datenquelle mit dem Projektdatenmodell angibt.

- Ein grüner Kreis bedeutet, dass die Datenquelle mit der Datenquelle des Projektdatenmodells operativ kompatibel ist. Operativ kompatible Datenquellen müssen alle operativen Felder des Projektdatenmodells enthalten, können aber auch zusätzliche aufweisen. Solche Datenquellen sind für Regeln, Scoring, Simulation und Tests geeignet.
- Ein orangener Kreis bedeutet, dass die Datenquelle über wenigstens ein mit dem Projektdatenmodell kompatibles Feld verfügt, dessen Name und Typ identisch sind. Solche Datenquellen können ebenfalls zusätzliche Felder enthalten und sind zur Erstellung und Auswertung von Modellen geeignet.
- Ein roter Kreis bedeutet, dass die Datenquelle mit der Datenquelle des Projektdatenmodells inkompatibel ist und die Felder vor der Benutzung in der Anwendung abgebildet werden müssen. Inkompatible Datenquellen sind solche, die wenigstens ein Feld aufweisen, dessen Typ mit dem entsprechenden Typ des Projektdatenmodells inkompatibel ist.

Felder abbilden. Mit dieser Option können Sie die Felder der Datenquelle mit denen des Projektdatenmodells vergleichen und alle kompatiblen Felder abbilden, so dass sie den vom Projektdatenmodell verlangten entsprechen, sowie Abbildungen aufheben. [Für weitere Informationen siehe Thema Abbildung von Feldern auf S. 41.](#)

Kopieren. Erlaubt das Kopieren der Datenquelle in den Bereich Eigene Datenquellen.

Eigene Datenquellen

Datenquellen in dieser Liste werden unter Ihrem Benutzerkonto gespeichert, sodass sie Ihnen bei jeder Anmeldung zur Verfügung stehen, und können in beliebige Projekte, Modelle oder Regeln, die Sie öffnen, kopiert werden (vorausgesetzt, Sie haben vom Administrator die entsprechenden Berechtigungen erhalten). Die Felder in diesem Bereich der Registerkarte funktionieren genauso wie die des Bereichs Projektdatenquellen, mit dem Unterschied, dass es keine Kompatibilitäts-Spalte gibt.

Die Spalte Kopieren ermöglicht es Ihnen, die Datenquelle in den Bereich Eigene Datenquellen zu kopieren.

Arbeiten mit Datenquellen

- ▶ Um eine neue Datenquelle entweder zu den Projektdatenquellen oder Eigenen Datenquellen hinzuzufügen, wählen Sie Datenquelle hinzufügen. [Für weitere Informationen siehe Thema Erstellung einer neuen Datenquelle auf S. 32.](#)
- ▶ Zum Ändern des Messniveaus oder Typs eines Felds in der Datenquelle im Projektdatenmodell wählen Sie das entsprechende Niveau. [Für weitere Informationen siehe Thema Messniveaus auf S. 31.](#)
- ▶ Um Datenquellen in die oder aus der Liste der Projektdatenquellen zu kopieren, klicken Sie auf den entsprechenden Pfeil in der Spalte „Kopieren“.

- ▶ Um Feldnamen für eine Datenquelle auf das Projektdatenmodell abzubilden, klicken Sie auf den entsprechenden Link in der Spalte „Kompatibel“. (Nachdem die Feldnamen abgebildet wurden, wird der Link nicht mehr angezeigt.) [Für weitere Informationen siehe Thema Abbildung von Feldern auf S. 41.](#)
- ▶ Für eine Vorschau der Datenquelle klicken Sie auf das Vorschau-Symbol. [Für weitere Informationen siehe Thema Vorschau von Daten auf S. 31.](#)
- ▶ Um weitere Felder (Ausdrücke oder Modellausgaben aus anderen Datenquellen) hinzuzufügen, um Ihr Projektdatenmodell zu erweitern, klicken Sie auf Zusatzfelder hinzufügen/bearbeiten. [Für weitere Informationen siehe Thema Hinzufügen von Zusatzfeldern auf S. 39.](#)

Vorschau von Daten

Um zu überprüfen, ob Sie die richtige Datenquelle ausgewählt haben, klicken Sie auf das Vorschau-Symbol. Ein Muster der in der Quelle enthaltenen Daten wird angezeigt. Die Vorschau enthält alle etwaigen Zusatzfelder, die im Abschnitt Zusatzfelder der Registerkarte „Daten“ aufgeführt sind.

Messniveaus

Mit dem Messniveau können die von jedem Datenfeld dargestellten Informationen charakterisiert werden, und es kann festlegen, wie ein bestimmtes Feld in Regeln, bei der Modellierung oder in anderen Anwendungen verwendet wird. Das Messniveau kann im Projektdatenmodell einer Datenquelle festgelegt werden. Beispielsweise können Sie das Messniveau für ein ganzzahliges Feld mit den Werten 0 und 1 auf *Flag* setzen, um anzugeben, dass 0 = *Falsch* und 1 = *Wahr*. Alternativ können Sie das Messniveau im Dialogfeld Datenquellen-Editor ändern, wenn Sie die zu verwendenden Eingabefelder festlegen. [Für weitere Informationen siehe Thema Auswählen von Eingabefeldern auf S. 40.](#)

Die folgenden Messniveaus stehen zur Verfügung:

- **Stetig.** Wird zur Beschreibung numerischer Werte verwendet, beispielsweise des Bereichs 0–100 oder 0,75–1,25. Stetige Werte können ganze Zahlen, reelle Zahlen oder Daten/Uhrzeiten sein.
- **Kategorial.** Wird für Zeichenfolgewerte benutzt, wenn die genaue Anzahl unterschiedlicher Werte unbekannt ist. Dies ist ein **uninstantiierter** Datentyp, was bedeutet, dass noch nicht alle möglichen Informationen über die Speicherung und Verwendung der Daten bekannt sind. Sobald Daten gelesen worden sind, wird das Messniveau *Flag*, *Nominal* oder *Typlos*.
- **Flag.** Wird für Daten mit zwei unterschiedlichen Werten verwendet, die das Vorliegen oder Nichtvorliegen eines Merkmals ausdrücken, wie z. B. *true* (wahr) und *false* (falsch), *Yes* (Ja) und *No* (Nein) bzw. 0 und 1. In IBM® Analytical Decision Management wird der erste Wert aus der Kategorienliste als der “falsche” Wert und der zweite als der “wahre” Wert verwendet. Hierzu wird ein entsprechender Tool Tip angezeigt, wenn Sie den Mauszeiger auf die Werte in der Registerkarte “Daten” bewegen. Die Daten können als Text, ganze Zahlen, reelle Zahlen, Datum, Uhrzeit oder Zeitstempel dargestellt werden.

- **Nominal.** Wird zur Beschreibung von Daten mit mehreren verschiedenen Werten benutzt, die jeweils als Element einer Menge behandelt werden, z. B. *small/medium/large* (klein/mittel/groß). Nominale Daten können einen beliebigen Speichertyp aufweisen - "Numerisch", "Zeichenkette" oder "Datum/Uhrzeit". Beachten Sie, dass die Werte durch die Festlegung des Messniveaus *Nominal* nicht automatisch in den Speichertyp "Zeichenkette" umgewandelt werden.
- **Ordinal.** Wird für Daten mit mehreren unterschiedlichen Werten benutzt, die über eine implizite Reihenfolge verfügen. Beispielsweise können Gehaltskategorien oder Zufriedenheitsranglisten als ordinale Daten behandelt werden. Die Reihenfolge wird durch die natürliche Sortierreihenfolge der Datenelemente definiert. Zum Beispiel ist 1, 3, 5 die natürliche Sortierreihenfolge für eine Menge ganzer Zahlen, während HIGH, LOW, NORMAL (HOCH, NIEDRIG, NORMAL) (alphabetisch aufsteigend) die Reihenfolge einer Menge von Zeichenfolgen ist. Das Messniveau "Ordinal" ermöglicht es Ihnen, eine Menge kategorialer Daten als Ordinaldaten zur Visualisierung, zur Modellbildung und zum Export in andere Anwendungen wie z. B. IBM® SPSS® Statistics zu definieren, die Ordinaldaten als eigenen Datentyp behandeln. Sie können Ordinalfelder überall dort benutzen, wo Nominalfelder benutzt werden können. Außerdem können Felder beliebigen Speichertyps (reell, ganze Zahl, Zeichenkette, Datum, Uhrzeit usw.) als Ordinalfelder definiert werden.
- **Typlos.** Wird für Daten benutzt, die keinem der obigen Typen entsprechen, für Felder mit einem einzigen Wert sowie für nominale Daten, deren Menge mehr als die zulässige Anzahl von Elementen enthält. Dieser Typ ist außerdem in Fällen nützlich, wo das Messniveau andernfalls eine Menge mit sehr vielen Elementen wäre (z. B. Kontonummern).

Erstellung einer neuen Datenquelle

- ▶ Zum Öffnen des Datenquellen-Editors wählen Sie entweder im Bereich „Projektdatenquellen“ oder im Bereich „Eigene Datenquellen“ den Befehl Datenquelle hinzuzufügen.
- ▶ Geben Sie einen Namen für die Datenquelle ein.
- ▶ Wählen Sie den Typ der Quelle:
 - **Datei.** Es werden zahlreiche Dateitypen unterstützt, z. B. Dateien im Textformat (*.txt, *.csv, etc.), im Excel-Format, im von IBM® SPSS® Statistics verwendeten *.sav-Format und im SAS-Format. [Für weitere Informationen siehe Thema Dateiquelle auf S. 33.](#)
 - **Datenbank.** Alle Datenbanken, die ODBC unterstützen, z. B. Microsoft SQL Server, DB2, Oracle u. a. [Für weitere Informationen siehe Thema Datenbankquelle auf S. 36.](#)
 - **Enterprise-Ansicht.** . Erstellt eine Verbindung zu einer IBM® SPSS® Collaboration and Deployment Services Enterprise View, wie in IBM® SPSS® Collaboration and Deployment Services definiert. [Für weitere Informationen siehe Thema Enterprise-Ansichts-Quelle auf S. 36.](#)
 - **IBM Cognos BI-Server** Jede Datenbank, die IBM Cognos BI-Server unterstützt. [Für weitere Informationen siehe Thema IBM Cognos BI-Quelle auf S. 37.](#)
- ▶ Legen Sie für den Datenquellentyp geeignete Optionen fest.

Sie können auf das Symbol Anzahl der Datensätze klicken, um die Anzahl an Datensätzen in der aktuellen Datenquelle anzuzeigen, oder auf Vorschau, um die ersten zehn Datensätze der Datenquelle anzuzeigen.

Dateiquelle

Die Anwendung unterstützt verschiedene Datendateitypen. Wenn Sie den Dateinamen eingeben oder eine Datei suchen, erkennt die Anwendung den Dateityp automatisch und blendet im Datenquellen-Editor zusätzliche Felder speziell für den betreffenden Typ ein. Falls der Dateityp nicht korrekt erkannt wird oder eine bestimmte Programmversion erfordert (z. B. Microsoft Excel 1997-2003), können Sie den richtigen Dateityp von Hand einstellen.

Sie können auf Durchsuchen klicken, um eine Datei aus dem Repository auszuwählen, oder auf Lokale Datei hochladen, um eine Datei aus Ihrem lokalen Dateisystem auszuwählen (falls aktiviert).

Wenn Sie eine lokale Datei hochladen, wird diese in den IBM® SPSS® Modeler Server hochgeladen. Navigieren Sie im Dialogfeld "Hochladen" zu der lokalen Datei, die hochgeladen werden soll, und wählen Sie sie aus. Navigieren Sie anschließend zu dem SPSS Modeler Server-Ziel zum Hochladen und wählen Sie es aus. Bei der Auswahl des Ziels können Sie den Dateinamen bei Bedarf ändern.

Die Option Lokale Datei hochladen ist standardmäßig deaktiviert. Administratoren können das browserbasierte IBM® SPSS® Collaboration and Deployment Services Deployment Manager verwenden, um die Option zu aktivieren bzw. zu deaktivieren. Ferner können Sie damit die maximal zulässige Dateigröße erhöhen oder verringern.

Textbasierte Datendateien

Wenn Sie eine textbasierte Datenquelle auswählen, werden Sie aufgefordert, weitere Details einzugeben.

Erste Zeile enthält Spaltennamen. Wählen Sie diese Einstellung, wenn die Namen der einzelnen Spalten in der ersten Zeile der Datenquelle angegeben sind.

Kodierung. Gibt die verwendete Textkodierungsmethode an. Sie haben die Wahl zwischen der System-StandardEinstellung und UTF-8.

Dezimaltrennzeichen. Hier können Sie das Trennzeichen für die Dezimalstellen in den Daten festlegen.

- **Punkt (.)**. Als Dezimaltrennzeichen wird ein Punkt verwendet.
- **Komma (,)**. Als Dezimaltrennzeichen wird ein Komma verwendet.

Trennzeichen. Mit den für dieses Steuerelement aufgelisteten Kontrollkästchen können Sie angeben, welche Zeichen, z. B. das Komma (,), die Feldbegrenzungen in der Datei definieren. Außerdem können Sie mehr als ein Trennzeichen angeben, z. B. „|” für Datensätze mit mehreren Trennzeichen. Das Standardtrennzeichen ist das Komma.

Anmerkung: Wenn das Komma auch als Dezimaltrennzeichen definiert ist, funktionieren die Standardeinstellungen in diesem Fall nicht. Wenn das Komma sowohl als Feldtrennzeichen als auch als Dezimaltrennzeichen festgelegt ist, wählen Sie in der Liste “Trennzeichen” Andere aus. Geben Sie dann manuell ein Komma in das Eingabefeld ein.

Wählen Sie Mehrere leere Trennzeichen zulassen aus, um mehrere nebeneinander liegende leere Trennzeichen als ein einzelnes Trennzeichen zu behandeln. Beispiel: Wenn auf einen Datenwert vier Leerzeichen und ein weiterer Datenwert folgen, wird diese Gruppe als zwei statt fünf Felder betrachtet.

Erweiterte Optionen

EOL-Kommentarzeichen Geben Sie Zeichen wie # oder ! ein, um auf Anmerkungen in den Daten hinzuweisen. Wenn ein solches Zeichen in der Datendatei angezeigt wird, werden alle Daten bis zu diesem Zeichen, jedoch nicht einschließlich des nächsten Zeilenumbruchs, ignoriert.

Eingabefelder festlegen. Geben Sie die Anzahl der aus jedem Datensatz zu verwendenden Eingabefelder an.

Führende Zeichen überspringen. Legen Sie fest, wie viele Zeichen am Anfang des ersten Datensatzes ignoriert werden sollen.

Nach Typ zu durchsuchende Zeilen. Legen Sie fest, wie viele Zeilen nach angegebenen Datentypen durchsucht werden sollen.

Führende und abschließende Leerzeichen löschen. Wählen Sie diese Option aus, um führende und abschließende Leerzeichen in Zeichenketten beim Importieren zu entfernen. Sie können wahlweise führende, abschließende, führende und abschließende oder keine Leerzeichen entfernen.

Ungültige Zeichen. Wählen Sie Verwerfen, um ungültige Zeichen aus der Datenquelle zu entfernen. Wählen Sie Ersetzen mit, um ungültige Zeichen durch das angegebene Symbol (nur ein Zeichen) zu ersetzen. Ungültige Zeichen sind Null-Zeichen (0) bzw. alle Zeichen, die in der angegebenen Kodierungsmethode nicht vorhanden sind.

Anführungszeichen. Mit den Dropdown-Listen können Sie angeben, wie einfache und doppelte Anführungszeichen beim Importieren zu behandeln sind. Sie können alle Anführungszeichen Verwerfen, Als Text einschließen, d.h. in den Feldwert einschließen, oder Paaren und verwerfen, um Anführungszeichenpaare zu finden und zu löschen. Kann einem Anführungszeichen kein zweites Anführungszeichen zugeordnet werden, wird eine Fehlermeldung ausgegeben. Sowohl die Option Verwerfen als auch Paaren und verwerfen speichert den Feldwert (ohne Anführungszeichen) als Zeichenkette.

Eingabefelder festlegen. Für weitere Informationen siehe [Thema Auswählen von Eingabefeldern auf S. 40](#).

Excel-Datendateien

Wenn Sie eine Excel-Datenquelle auswählen, werden Sie aufgefordert, weitere Details einzugeben:

Erste Zeile enthält Spaltennamen. Wählen Sie diese Einstellung, wenn die Namen der einzelnen Spalten in der ersten Zeile der Datenquelle angegeben sind.

Benannter Bereich. Ermöglicht die Auswahl eines benannten Zellenbereichs, der im Excel-Arbeitsblatt definiert ist. Wenn Sie einen benannten Bereich verwenden, sind andere Einstellungen für Arbeitsblatt und Datenbereich nicht mehr anwendbar und werden daher deaktiviert.

Arbeitsblatt wählen. Ermöglicht die Auswahl des zu importierenden Arbeitsblatts nach Namen oder Index.

- **Nach Namen.** Wählen Sie den Namen des zu importierenden Arbeitsblattes aus.
- **Nach Index.** Geben Sie den Indexwert des zu importierenden Arbeitsblatts an – 0 für das erste Arbeitsblatt, 1 für das zweite usw.

Bereich im Arbeitsblatt. Sie können Daten ab der ersten nichtleeren Zeile oder ab einem explizit angegebenen Zellenbereich importieren.

- **Bereich beginnt bei erster nicht leerer Zeile.** Sucht die erste nichtleere Zelle und verwendet diese als linke obere Ecke des Datenbereichs.
- **Expliziter Zellenbereich.** Ermöglicht die Angabe eines expliziten Bereichs nach Zeile und Spalte. Um beispielsweise den Excel-Bereich A1:D5 anzugeben, können Sie in das erste Feld A1 und in das zweite Feld D5 eintragen (oder alternativ R1C1 und R5C4). Alle Zeilen im angegebenen Bereich werden ausgegeben, einschließlich der Leerzeilen.

Bei leeren Zeilen. Wenn eine leere Zeile gefunden wird, können Sie angeben, ob die Zeile übersprungen oder ignoriert werden soll, oder mit Leere Zeilen zurückgeben festlegen, dass alle Daten, einschließlich der Leerzeilen, bis zum Ende des Arbeitsblattes gelesen werden sollen.

Eingabefelder festlegen. [Für weitere Informationen siehe Thema Auswählen von Eingabefeldern auf S. 40.](#)

IBM SPSS Statistics Datendateien

Wenn Sie eine IBM® SPSS® Statistics-Datenquelle auswählen, werden Sie aufgefordert, weitere Details einzugeben.

Variablenamen. Wählen Sie eine Methode zur Behandlung von Variablenamen und -beschriftungen beim Importieren aus einer SPSS Statistics-*SAV*-Datei.

- **Namen und Beschriftungen lesen.** Wählen Sie diese Option, um sowohl Variablenamen als auch Variablenbeschriftungen einzulesen; dies ist die Standardeinstellung. Beschriftungen können in Diagrammen, Modellbrowsern und anderen Ausgabetypen angezeigt werden.
- **Beschriftungen als Namen lesen.** Wählen Sie diese Option, um anstelle der kurzen Feldnamen die beschreibenden Variablenbeschriftungen aus der SPSS Statistics-*SAV*-Datei einzulesen und diese Beschriftungen als Variablenamen zu verwenden.

Werte. Wählen Sie eine Methode zur Behandlung von Werten und Beschriftungen beim Importieren aus einer SPSS Statistics-*SAV*-Datei.

- **Daten und Beschriftungen lesen.** Wählen Sie diese Option, um sowohl die Werte selbst als auch die Wertebeschriftungen einzulesen; dies ist die Standardeinstellung.
- **Beschriftungen als Daten lesen.** Wählen Sie diese Option, wenn Sie anstelle der numerischen oder symbolischen Codes, mit denen die Werte dargestellt werden, die Wertebeschriftungen aus der *SAV*-Datei verwenden möchten. Wenn Sie diese Option beispielsweise für Daten mit

dem Feld „Geschlecht“ wählen, dessen Werte 1 und 2 für *männlich* und *weiblich* stehen, wird das Feld in eine Zeichenkette konvertiert, und *männlich* und *weiblich* werden als tatsächliche Werte importiert.

Vor dem Auswählen dieser Option sollten Sie Ihre SPSS Statistics-Daten unbedingt auf fehlende Werte überprüfen. Wenn ein numerisches Feld beispielsweise Beschriftungen nur für fehlende Werte verwendet (0 = *No Answer*, -99 = *Unknown*), werden bei Auswahl der obigen Option nur die Wertelabels *No Answer* und *Unknown* importiert und das Feld in eine Zeichenkette konvertiert. In diesem Fall sollten Sie die Werte selbst importieren.

Eingabefelder festlegen. Für weitere Informationen siehe [Thema Auswählen von Eingabefeldern auf S. 40](#).

Datenbankquelle

Wenn Sie eine Datenbankquelle auswählen, werden Sie aufgefordert, weitere Details einzugeben:

Verbindung testen. Geben Sie Ihren Datenbank-Benutzernamen und das Passwort ein und klicken Sie auf diese Schaltfläche, um zu prüfen, ob Sie auf die Datenbank zugreifen können.

Tabellenname. Wählen Sie die Tabelle aus, deren Daten Sie verwenden möchten.

Führende und abschließende Leerzeichen löschen. Wählen Sie diese Option aus, um führende und abschließende Leerzeichen in Zeichenketten beim Importieren zu entfernen. Sie können wahlweise führende, abschließende, führende und abschließende oder keine Leerzeichen entfernen.

Tabellen- und Spaltennamen in Anführungszeichen. Legen Sie fest, ob die Tabellen- und Spaltennamen in Anführungszeichen eingeschlossen werden sollen, wenn Abfragen an die Datenbank gesendet werden (wenn sie z.B. Leerzeichen oder Satzzeichen enthalten).

- Bei Auswahl der Option Nach Bedarf werden Tabellen- und Feldnamen *nur* in Anführungszeichen gesetzt, wenn sie Nichtstandardzeichen enthalten. Nichtstandardzeichen sind Nicht-ASCII-Zeichen, Leerzeichen und alle nicht alphanumerischen Zeichen außer einem Punkt (.
- Wählen Sie Nie, wenn Tabellen- und Feldnamen *nie* in Anführungszeichen gesetzt werden sollen.
- Wählen Sie Immer, wenn *alle* Tabellen- und Feldnamen in Anführungszeichen gesetzt werden sollen.

Eingabefelder festlegen. Für weitere Informationen siehe [Thema Auswählen von Eingabefeldern auf S. 40](#).

Enterprise-Ansichts-Quelle

Wenn Sie eine Enterprise-Ansichts-Quelle auswählen, werden Sie aufgefordert, weitere Details einzugeben:

Application-Ansicht. In der Dropdown-Liste wird ggfs. die ausgewählte Application-Ansicht angezeigt. Wenn in der aktuellen Sitzung Verbindungen zu anderen Application-Ansichten hergestellt wurden, werden diese ebenfalls in der Dropdown-Liste angezeigt. Klicken Sie auf die Schaltfläche Durchsuchen, um nach anderen Application-Ansichten im Repository zu suchen.

Tabelle. Wählen Sie die Application-Ansichts-Tabelle aus, deren Daten Sie verwenden möchten.

Datenanbieter. In der Dropdown-Liste werden die Namen von Datenanbieterdefinitionen für die ausgewählte Application-Ansicht aufgeführt. Wählen Sie eine Datenanbieterdefinition aus oder wählen Sie die Option -Keine Datenquelle- aus. Die Auswahl der Option -Keine Datenquelle- bedeutet, dass die Datenquelle in der Anwendung praktisch unbrauchbar ist und nicht für „Test“, „Simulieren“, die „Was, wenn“-Analyse usw. zur Verfügung steht.

Umgebung. Der Umgebungswert, den Sie in IBM® SPSS® Collaboration and Deployment Services festlegen, bestimmt, welche Datenanbieterdefinitionen verfügbar sind.

Eingabefelder festlegen. Für weitere Informationen siehe Thema [Auswählen von Eingabefeldern auf S. 40](#).

Nähere Informationen zu IBM® SPSS® Collaboration and Deployment Services Enterprise View und vollständige Einzelheiten darüber, wie diese Daten konfiguriert werden, finden Sie in der Begleitdokumentation von IBM SPSS Collaboration and Deployment Services.

IBM Cognos BI-Quelle

Es können Daten aus der IBM Cognos BI-Datenbank oder einzelne Listenberichte für die Verwendung in IBM® Analytical Decision Management importiert werden.

Anmerkung: Wenn der mit IBM Analytical Decision Management verwendete IBM® SPSS® Modeler Server den IBM Cognos BI-Server nicht unterstützt, steht diese Option nicht zur Verfügung.

Wenn Sie IBM Cognos BI-Server auswählen und auf Verbinden klicken, werden Sie aufgefordert, die Details für die Serververbindung einzugeben.

- ▶ Geben Sie die Server-URL des IBM Cognos-Servers ein, den Sie für den Import bzw. Export von Daten verwenden möchten. Wenden Sie sich an Ihren Cognos-Systemadministrator, wenn Sie sich nicht sicher sind, welche URL Sie verwenden müssen.
- ▶ Wählen Sie den Modus für die Verbindung aus. Wenn Sie sich als ein bestimmter Benutzer (z. B. als Administrator) anmelden möchten, wählen Sie die Option Anmeldedaten festlegen aus und geben Sie den Namespace, den Benutzernamen und das Passwort für IBM Cognos ein:
 - Geben Sie den Namespace des IBM Cognos-Sicherheitsauthentifizierungsanbieters ein, über den die Anmeldung beim Server erfolgt. Der Authentifizierungsanbieter dient dazu, Benutzer, Gruppen und Rollen zu definieren und zu verwalten und den Authentifizierungsprozess zu steuern.
 - Geben Sie den Benutzernamen für IBM Cognos ein, der für die Anmeldung beim Server verwendet wird.
 - Geben Sie das Passwort (Passwort) ein, das zum angegebenen Benutzernamen gehört.
- ▶ Um sich stattdessen ohne Anmeldeinformationen anzumelden (bei dieser Variante können die Felder für Namespace, Benutzernamen und Passwort nicht ausgefüllt werden), verwenden Sie die Option Anonyme Verbindung verwenden.

Anmerkung: Bei einigen Serververbindungen sind anonyme Verbindungen nicht zulässig.

Auswählen von Cognos-Objektdetails

Wählen Sie nach dem Einrichten der IBM Cognos-Serververbindung den Modus aus. Der Modus Daten ermöglicht das Navigieren zu einem Paket, aus dem die Daten importiert werden sollen. Der Modus Listenbericht ermöglicht das Navigieren zu einem vordefinierten Listenbericht, aus dem die Daten importiert werden sollen.

Ein Paket enthält ein IBM Cognos-Modell und sämtliche Ordner, Abfragen, Berichte, Ansichten, Verknüpfungen, URLs und Aufgabendefinitionen, die dem betreffenden Modell zugeordnet sind. Ein IBM Cognos-Modell definiert Geschäftsregeln, Datenbeschreibungen, Datenbeziehungen, Geschäftsdimensionen und -hierarchien sowie andere administrative Aufgaben.

- ▶ Wählen Sie den Modus Daten aus, und klicken Sie auf Durchsuchen, um ein Paket auszuwählen. Daraufhin wird ein Dialogfeld mit allen verfügbaren Paketen angezeigt, aus denen Inhalte importiert werden können.
- ▶ Oder wählen Sie den Modus Listenbericht aus, und klicken Sie auf Durchsuchen, um einen einzelnen Listenbericht auszuwählen. Daraufhin wird ein Dialogfeld mit Paketen und Ordnern angezeigt, aus denen ein Bericht ausgewählt werden kann. *Anmerkung:* Es können nur einzelne Listenberichte importiert werden. Der Import von mehreren Listenberichten wird nicht unterstützt.
- ▶ Sie können die Liste bei Bedarf erweitern, um das erforderliche Paket oder den erforderlichen Bericht zu finden. Wählen Sie das gewünschte Element aus, und klicken Sie auf OK. Die Paket- oder Berichtsdetails werden im Datenquellen-Editor angezeigt.
- ▶ Nachdem Sie ein Paket ausgewählt haben, können Sie die zu importierenden Objekte auswählen. D. h., Sie können z. B. die Abfragesubjekte (stehen für Datenbanktabellen) für Paketobjekte oder Abfrageelemente (stehen für Tabellenspalten) für einzelne Felder auswählen, die aus einem bestimmten Paket verwendet werden sollen. *Anmerkung:* Die Daten müssen im Format UTF-8 vorliegen.
- ▶ Wählen Sie die Option Daten vor Import aggregieren aus, wenn aggregierte Daten anstelle von Rohdaten importiert werden sollen. Aggregierte Daten umfassen in der Regel deutlich weniger Datensätze als Rohdaten.

Objekte auswählen. Mit dieser Option werden Objekte (auch Namespaces genannt) angezeigt, die aus Paketen oder Listenberichten ausgewählt werden können. Erweitern Sie ein oder mehrere Objekte, um die Eingabefelder anzuzeigen, die importiert werden können. Durch die Auswahl eines Objekts werden automatisch alle zugehörigen Eingabefelder ausgewählt.

Parameter überprüfen. Wenn Parameter für ein Datenobjekt oder einen Bericht definiert wurden, können vor dem Importieren des Objekts oder des Berichts Werte für die jeweiligen Parameter festgelegt werden. In dieser Tabelle ist die Parameterliste enthalten, die für Filter und Berechnungen erforderlich ist. Möglicherweise gibt es Standardwerte für die Parameter. Wenn ein oder mehrere Parameter keinen Standardwert aufweisen, wird eine Warnung ausgegeben. Geben Sie die gewünschten Parameterwerte in der Spalte Wert ein. Mithilfe von Parametern können vor dem Importieren von Daten Einstellungen vorgenommen werden (z. B. zum Ausführen einer parametrisierten Berechnung).













Filter überprüfen. Wenn für ein Paket Filter definiert wurden, können ein oder mehrere Filter importiert werden. Wenn ein zu importierender Filter Importdaten zugeordnet ist, wird der Filter vor dem Datenimport angewendet. Filter sind möglicherweise parametrisiert.

Eingabefelder überprüfen. Listet die Eingabefelder (auch als Datenbankobjekte bezeichnet) auf, die zur Verarbeitung in IBM® Analytical Decision Management importiert werden. Wenn ein bestimmtes Feld nicht benötigt wird, deaktivieren Sie das Kontrollkästchen neben dem Feldnamen.

Cognos-Objektsymbole

Die verschiedenen Objekttypen, die von einer IBM Cognos BI-Datenbank importiert werden können, werden in Form von Symbolen dargestellt (siehe Tabelle unten).

Tabelle 4-1
Cognos-Objektsymbole

Symbol	Objekt
	Paket
	Namespace
	Abfragesubjekt
	Abfrageelement
	Kennzahldimension
	Kennzahl
	Dimension
	Ebenenhierarchie
	Ebene
	Filter
	Berichte
	Standalone-Berechnung

Hinzufügen von Zusatzfeldern

In den meisten IBM® Analytical Decision Management-Anwendungen können Sie zu Ihrem Projektdatenmodell Zusatzfelder wie Ausdrücke oder Modellausgaben hinzufügen. Im Abschnitt Zusatzfelder der Registerkarte „Daten“ unter „Projektdatenmodell“ werden alle Ausdrücke oder Modellausgaben angezeigt, die als Teil der Daten aufgenommen werden sollen. Beachten Sie, dass der Abschnitt Zusatzfelder in der Anwendung IBM® SPSS® Rules Management nicht enthalten ist.

- Klicken Sie je nach Anwendung im Abschnitt „Projektdatenmodell“ auf der rechten Bildschirmseite auf Zusatzfelder hinzufügen/bearbeiten.

So fügen Sie einen Ausdruck hinzu:

- ▶ Klicken Sie auf Ausdruck hinzufügen. Der Ausdrucksektor wird angezeigt. [Für weitere Informationen siehe Thema Ausdrucksektor auf S. 41.](#)

So fügen Sie eine Modellausgabe hinzu:

- ▶ Klicken Sie auf Modellausgabe hinzufügen und navigieren Sie dann zu einem Modell und wählen Sie es aus. Das Dialogfeld Eingaben und Ausgaben für das ausgewählte Modell wird angezeigt.
- ▶ Wenn die Modelleingabefelder kompatibel sind, wird der Abschnitt Modelleingaben-Zuordnung des Dialogfelds minimiert. Wenn die Modelleingabefelder nicht kompatibel sind, wird der Abschnitt erweitert und Sie müssen die verfügbaren Tabelleneingaben passenden Eingaben des ausgewählten Modells zuordnen.
- ▶ Im Abschnitt Modellausgabe werden alle verfügbaren Modellausgaben für das ausgewählte Modell aufgelistet. Wählen Sie die Modellausgabefelder aus, die Sie zu Ihren Daten hinzufügen möchten. In der Spalte Modellausgabenname wird der vom Modell generierte Name des Ausgabefelds angezeigt.

Alle Felder der Modellausgaben können ausgewählt werden. Modellausgabefelder sind alle Ausgabefelder des Streams, die nicht gleichzeitig Eingaben für den Stream sind (alle vom Stream hinzugefügten Felder). Die Eingabefelder umfassen möglicherweise nicht alle Elemente im Projektdatenmodell. Ein Modellausgabefeld kann also denselben Namen tragen wie ein Projektdatenmodell-Feld. Daher können Sie auch die Namen der Datenfelder hier ändern, wenn Sie möchten.

In der Spalte Messniveau wird der Messniveautyp der Modellausgabe angezeigt, dem das entsprechende Typsymbol vorangestellt ist, und in der Spalte Werte werden die Werte der Modellausgabe angezeigt. Beachten Sie, dass in diesen Spalten möglicherweise Ergebnisse für einige der Felder angezeigt werden, für andere jedoch nicht. Typischerweise sind die Messniveautypen und die Werte für tatsächliche Modellausgaben (vorhergesagter Wert, Wahrscheinlichkeit usw.) ohne Datenscan bekannt. Bekannte Daten werden angezeigt. Für andere Ausgabefelder aus einem Stream (z. B. einem abgeleiteten Feld) sind die Messniveautypen bzw. die Werte jedoch möglicherweise nicht ohne Datenscan bekannt.

- ▶ Unten im Dialogfeld wird der Abschnitt Modellparameter aufgenommen, wenn der ausgewählte Stream Scoring-Parameter enthält. Geben Sie Werte für die Modellparameter ein und klicken Sie auf Speichern, wenn Sie die Arbeit in diesem Dialogfeld abgeschlossen haben. Alle ausgewählten Modellausgaben werden dem Abschnitt Zusatzfelder der Registerkarte „Daten“ hinzugefügt.

Auswählen von Eingabefeldern

Wenn Sie über große Datenmengen verfügen, können Sie die aus Ihrer Quelle importierte Datenmenge reduzieren oder modifizieren und beispielsweise nur die für Ihr Interessengebiet relevanten Felder verwenden.

- ▶ Klicken Sie im unteren Bereich des Dialogfelds „Datenquellen-Editor“ auf Eingabefelder festlegen; im Dialogfeld werden alle Datenfelder Ihrer Quelle eingeblendet. Außerdem werden

alle Ausdrücke angezeigt, die Sie definiert haben. [Für weitere Informationen siehe Thema Ausdruckseditor auf S. 41.](#)

- ▶ Wählen Sie die Felder und Ausdrücke aus, die Sie verwenden möchten.

Außer der Auswahl der Daten können Sie auch Änderungen des Messniveaus vornehmen, falls erforderlich. [Für weitere Informationen siehe Thema Messniveaus auf S. 31.](#)

Ausdruckseditor

Mit dem Ausdrucksmanager können Sie zusätzliche Felder oder Attribute für die Benutzung mit der Anwendung ableiten. Wenn Sie beispielsweise Bankdaten verwenden, können Sie einen Ausdruck erstellen, der das Verhältnis zwischen dem Einkommen eines Kunden und der Anzahl seiner Darlehenskonto zeigt. Ausdrücke sind immer numerisch mit stetigem Messniveau – dies kann nicht geändert werden.

Anmerkung: Ausdrücke sind genauso wie operative Daten an verschiedenen Stellen der Anwendung verfügbar, z. B. in Regeln und Modellprädiktoren.

- ▶ Um vorhandene Ausdrücke anzuzeigen, klicken Sie auf der Registerkarte „Daten“ auf Zusatzfelder hinzufügen/bearbeiten und dann auf Ausdruck hinzufügen.

Erstellen eines Ausdrucks

- ▶ Geben Sie im Ausdruckseditor einen eindeutigen Namen für den Ausdruck ein.
- ▶ Alle verfügbaren Eingaben aus der Datenquelle werden angezeigt. Sie können Ihren Ausdruck zusammenstellen, indem Sie die gewünschten Eingaben und Rechenzeichen doppelklicken oder in das Feld Ausdrucksgleichung ziehen.
- ▶ Geben Sie eine Beschreibung ein, aus der hervorgeht, wofür der Ausdruck verwendet wird.
- ▶ Klicken Sie auf Speichern.

Abbildung von Feldern

Beim Hinzufügen einer neuen Datenquelle müssen Sie eventuell alle ihrer Felder abbilden, um die Kompatibilität mit dem Projektdatenmodell sicherzustellen. Falls beispielsweise das Projektdatenmodell ein Feld namens *Kauf* mit den Werten *Ja* und *Nein* enthält (Messniveau Flag), müssen alle verwendeten Datenquellen über ein kompatibles Feld verfügen. Falls die Feldnamen nicht übereinstimmen, können sie entsprechend abgebildet werden. Beachten Sie, dass die Eingabefelder und die entsprechenden abgebildeten Felder denselben Datentyp aufweisen müssen.

Wenn das Projektdatenmodell geändert wird, kann es sein, dass danach viele Datenquellen nicht mehr korrekt darauf abgebildet werden. In diesem Fall ändert sich bei jeder der entsprechenden Datenquellen im Bereich Projektdatenquellen die Farbe des die Kompatibilität anzeigenden Kreises – Sie müssen diese dann auf das neue Projektdatenmodell abbilden.

- ▶ Klicken Sie im Bereich Projektdatenquellen auf Felder abbilden. Ein Dialogfeld wird geöffnet, und das System versucht zu ermitteln, welche Abbildung für Sie am besten geeignet ist. Sie

können jederzeit zu der vom System vorgeschlagenen Abbildung zurückkehren, indem Sie auf Zurücksetzen klicken.

- ▶ Wählen Sie ein Feld in der Spalte Verfügbare Eingaben sowie das entsprechende Feld in der Spalte Pflichteingabefelder und klicken auf Abbilden, um die beiden Felder abzubilden. Wiederholen Sie diesen Vorgang, bis alle Felder der Spalte Verfügbare Eingaben abgebildet wurden.
- ▶ Wenn Sie fertig sind, klicken Sie auf OK.

Regeln verwenden

Regeln ermöglichen die Automatisierung von Entscheidungen auf Grundlage von Geschäftslogik oder den Ergebnissen von Vorhersagemodellen. Je nach Anwendung können Sie mit Regeln z. B. Kunden aufgrund ihres Zahlungsverlaufs ausschließen, verdächtige Schadensmeldungen an eine spezielle Prüfungsabteilung übergeben oder diejenigen Kunden ansprechen, die einem Vorhersagemodell zufolge mit der höchsten Wahrscheinlichkeit kündigen, kaufen oder Betrug begehen. Während Regeln die Grundlage der Automatisierung sind, liefern Modelle die Vorhersagekraft. Die Möglichkeit, beides zu kombinieren, ist eines der Hauptmerkmale, durch die sich prädiktive Anwendungen von anderen Analyseprogrammen unterscheiden.

Regeln können an verschiedenen Stellen einer Anwendung erstellt, angewandt und gespeichert werden.

Regeln erstellen

- ▶ Klicken Sie an der Stelle in der Anwendung, an der Sie die Regel hinzufügen möchten, auf das Symbol Neue Regel erstellen.
- ▶ Alternativ können Sie auf das Symbol Vorhandene Regel suchen klicken, um nach Regeln zu suchen, die zur aktuellen Anwendung hinzugefügt oder wiederverwendet werden können. [Für weitere Informationen siehe Thema Regeln exportieren und wiederverwenden auf S. 46.](#)
- ▶ Zur Erstellung einer gemeinsam genutzten Regel starten Sie IBM® SPSS® Rules Management über den Anwendungs-Startbereich.

Segmentregeln definieren






Segmentregeln kommen in Anwendungen an vielen Orten zum Einsatz, um Datensätze je nach Bedarf in anderen Anwendungen auszuwählen und zuzuordnen. Segmente werden durch einen oder mehrere Ausdrücke definiert, die den Wert „wahr“ oder „falsch“ zurückgeben. Beispielsweise kann eine Regel mit einem einzelnen Segment existieren, das den Inhalt „wahr“ ergibt, wenn der Wert des Felds „Alter“ unter 18 liegt. Segmente können anhand von in den Daten definierten Feldern oder Ausdrücken definiert werden oder anhand der von einem Modell zurückgegebenen Neigungsscores oder Vorhersagen. In Anwendungen mit der Registerkarte „Definieren“, die so konfiguriert sind, dass mehrere Dimensionshierarchien möglich sind (welche als *Elemente* bezeichnet werden), können Sie auch aus anderen Dimensionsmitgliedern eine Auswahl treffen. [Für weitere Informationen siehe Thema Projektkonfiguration in Kapitel 3 auf S. 21.](#)

Jedes Segment wird durch einen oder mehrere Ausdrücke definiert. Nur wenn alle Bedingungen erfüllt sind, wird der Wert „wahr“ zurückgegeben. Ein über Alter und Geschlecht definiertes Segment würde beispielsweise nur Datensätze beinhalten, die beide Bedingungen erfüllen. Für alle übrigen Datensätze würde der Wert *falsch* zurückgegeben.

So erstellen und bearbeiten Sie Segmentregeln:

- ▶ Klicken Sie in der Anwendung, der Sie die Regel hinzufügen möchten, auf das Symbol Neue Regel erstellen.
- ▶ Legen Sie einen Namen für die Regel fest.
- ▶ Klicken Sie auf den Pfeil nach unten, um für die Eingabespalte ein Feld, Modell oder eine Regel anzugeben.

Tabelle 5-1
Eingabespaltenoptionen

Symbol	Verwendung
	Wenn aktiviert, kann ein Dimensionsmitglied aus einem vorab definierten Element ausgewählt werden.
	Ermöglicht Ihnen, ein Feld im aktuellen Daten-Set auszuwählen.
	Ermöglicht Ihnen, ein vorhandenes Modell oder eine vorhandene Regel im Repository auszuwählen.
	Ermöglicht Ihnen, ein neues Modell zu erstellen.
	Ermöglicht die Erstellung eines Ausdrucks.

- ▶ Wählen Sie die Funktion, die Sie benutzen möchten. In der folgenden Tabelle sind die verfügbaren Optionen aufgelistet.

Tabelle 5-2
Verfügbare Operatoren

Werttyp:	Verfügbare Operatoren
Zeitstempel, Datum, Uhrzeit	<ul style="list-style-type: none"> ■ = (gleich) ■ != (nicht gleich) ■ > (größer als) ■ < (kleiner als) ■ >= (größer oder gleich) ■ <= (kleiner oder gleich) ■ ZWISCHEN ■ ISNIL ■ IST NICHT NULL
A-Bereich	<ul style="list-style-type: none"> ■ = (gleich) ■ != (nicht gleich) ■ > (größer als) ■ < (kleiner als) ■ >= (größer oder gleich) ■ <= (kleiner oder gleich)

Werttyp:	Verfügbare Operatoren
	<ul style="list-style-type: none"> ■ ZWISCHEN ■ ISNIL ■ IST NICHT NULL ■ Element von (ONE OF) ■ Kein Element von (NOT ONE OF)
A-Satz	<ul style="list-style-type: none"> ■ = (gleich) ■ != (nicht gleich) ■ ISNIL ■ IST NICHT NULL ■ Element von (ONE OF) ■ Kein Element von (NOT ONE OF)
A-Flag	<ul style="list-style-type: none"> ■ = (gleich) ■ != (nicht gleich) ■ ISNIL ■ IST NICHT NULL

- ▶ Legen Sie einen Wert als für den Eingabetyp geeignet fest. Für nicht-numerische Werte wird eine Dropdown-Liste mit Auswahlmöglichkeiten bereitgestellt.

Alternativ können Sie auf das Symbol Vorhandene Regel suchen klicken, um nach Regeln zu suchen, die bereits definiert worden sind. [Für weitere Informationen siehe Thema Regeln exportieren und wiederverwenden auf S. 46.](#)

Segmentregelsätze

- Segmentregeln können auch zu Sätzen zusammengefasst werden. Der Segmentregelsatz gibt den Wert „wahr“ zurück, wenn *mindestens eines* der einzelnen Segmente „wahr“ ist. Im Gegensatz dazu müssen von den Ausdrücken *innerhalb* eines Segments alle „wahr“ sein.
- Sie können Segmentregelsätze aus einer beliebigen Anwendung, die Regeln verwendet, oder aus IBM® SPSS® Rules Management exportieren, falls verfügbar. [Für weitere Informationen siehe Thema Regeln exportieren und wiederverwenden auf S. 46.](#)

Selektionsregeln definieren

Eine Selektionsregel ist ein Satz aus einer oder mehreren Segmentregeln oder Regelsätzen, denen jeweils das Attribut *Einschließen* oder *Ausschließen* zugewiesen ist. Selektionsregeln werden in Anwendungen vielerorts eingesetzt, um Datensatzgruppen zur Modellierung, zum Scoring oder für andere Verarbeitungsvorgänge auszuwählen.

- ▶ Klicken Sie in der Anwendung, der Sie die Regel hinzufügen möchten, auf das Symbol Neue Regel erstellen. Das Symbol ist überall dort verfügbar, wo Selektionsregeln benutzt werden können.
- ▶ Legen Sie einen Namen für die Regel fest, wählen Sie aus, ob es sich um eine einschließende oder ausschließende Regel handeln soll (auch später möglich), und fügen Sie je nach Bedarf

einen oder mehrere Ausdrücke hinzu. [Für weitere Informationen siehe Thema Segmentregeln definieren auf S. 43.](#)

- ▶ Klicken Sie auf OK, um die Regel zu speichern.
- ▶ Klicken Sie auf die Symbole in der Spalte Einschließen/Ausschließen, wenn Sie diese Einstellung für die einzelnen Regeln umschalten möchten. Für referenzierte Selektionsregelsätze kann diese Einstellung nicht geändert werden. Einschluss- und Ausschlussregeln können nicht in einer Regel kombiniert werden.

Alternativ können Sie auf das Symbol Vorhandene Regel suchen klicken, um nach Regeln zu suchen, die bereits definiert worden sind. [Für weitere Informationen siehe Thema Regeln exportieren und wiederverwenden auf S. 46.](#)

Selektionsregeln evaluieren

Klicken Sie auf das Symbol Selektionsregellogik anzeigen, um den Auswahlmechanismus anzuzeigen (logisches UND/ODER).

- Wenn mehrere Einschluss-Segmente angegeben werden, werden bei der Auswertung der Selektionsregeln diejenigen Datensätze eingeschlossen, die *mindestens einem* der Segmente zugeordnet werden. Beispielsweise könnten Kunden auf Grund der Bonität ODER der Länge der Kundenbeziehung eingeschlossen werden.
- Wenn mehrere Ausschluss-Segmente angegeben werden, werden Datensätze ausgeschlossen, die mindestens einem der Segmente zugeordnet werden. Auf diese Weise können Kunden beispielsweise auf Grund ihres Alters ODER der Bonität ausgeschlossen werden.
- Im Falle eines Konflikts haben Ausschlussregeln Vorrang. Kunden, die einem ausgeschlossenen Segment zugeordnet sind, werden auch dann ausgeschlossen, wenn sie aufgrund einer Zuordnung zu einem anderen Segment eingeschlossen werden müssten.
- Bei der Erstellung von Selektionsregelsätzen in IBM® SPSS® Rules Management müssen alle Regeln vom selben Typ sein – entweder *Einschließen* oder *Ausschließen*. Einschluss- und Ausschlussregeln können nicht in einem Regelsatz vereinigt werden.

Regeln exportieren und wiederverwenden

Standardmäßig werden Regeln innerhalb einer Anwendung erstellt und mit dieser gespeichert. Alternativ können Regeln exportiert oder in IBM® SPSS® Collaboration and Deployment Services Repository gespeichert und dann von mehreren Anwendungen genutzt werden. In diesem Fall speichern die verschiedenen Anwendungen einen Verweis auf eine bestimmte beschriftete Version der gemeinsamen Regel, und Änderungen an dieser Regel werden automatisch für alle Anwendungen übernommen, die diese Version benutzen.

Regeln exportieren

- ▶ Klicken Sie in einer Anwendung mit einer lokalen Regel, die Sie mit anderen Anwendungen nutzen möchten, auf das Symbol Exportieren, um die Regel in IBM SPSS Collaboration and Deployment Services Repository zu speichern, wo sie bei Bedarf wiederverwendet werden kann.

- Beim Export mehrerer Regeln können Sie entscheiden, ob Sie sie zusammen in einem Regelsatz oder als einzelne Regeln exportieren wollen. Im Fall von Einzelregeln wird jedes Segment als eigene Regel exportiert.
- Wenn Sie beim Export mehrerer Regeln Mit Selektionen speichern wählen, wird ein Selektionsregelsatz exportiert. In diesem Fall können Sie entweder Einschluss- oder Ausschlussregeln exportieren, nicht aber beides. Wenn Sie sowohl Einschluss- als auch Ausschlussregeln benötigen, müssen diese als getrennte Regelsätze exportiert werden. Falls Mit Selektionen speichern nicht ausgewählt ist, wird ein Segmentregelsatz exportiert.
- Wenn Sie beim Export von einem Aggregationsregel-Editor Mit Risikopunkten exportieren wählen, wird ein Aggregationsregelsatz erstellt. Falls diese Option nicht ausgewählt ist, wird ein Segmentregelsatz exportiert.

Vorhandene Regeln verwenden

- ▶ Klicken Sie in einer Anwendung, in der Sie eine gemeinsame Regel nutzen möchten, auf das Symbol Vorhandene Regel suchen, um IBM SPSS Collaboration and Deployment Services Repository nach Regeln zu durchsuchen, die in der aktuellen Anwendung wiederverwendet werden können.
- ▶ Um eine bestimmte beschriftete Version zu referenzieren, klicken Sie auf den Pfeil neben dem Regelnamen und wählen die gewünschte Version. Wenn keine Version ausgewählt wurde, wird die jeweils neueste verwendet. (Stellen Sie sich *NEUESTE* als eine spezielle Beschriftung vor, die bei der Erstellung einer neuen Version automatisch auf diese übertragen wird.) Es empfiehlt sich, eine bestimmte beschriftete Version zu verwenden, um unerwartete Ergebnisse zu verhindern, die bei Erstellung einer neuen Version einer Regel auftreten können.

Referenzierte Regeln können in vielen Fällen mit lokalen Regeln kombiniert werden, unterscheiden sich aber von diesen visuell durch Fettdruck und ein Symbol, das den Regel(satz)typ kennzeichnet.

- ▶ Alternativ können referenzierte Regeln in lokale Regeln konvertiert und in der aktuellen Anwendung gespeichert werden. Klicken Sie dazu auf den Regelnamen einer referenzierten Regel und wählen Sie dann im Dialogfeld „Regelvorschau“ die Option „In Einzelregeln konvertieren“.

Hinweise zur gemeinsamen Nutzung von Regeln

Die Verwendung gemeinsam genutzter Regeln unterliegt den folgenden Bedingungen:

- Regeln können für die Wiederverwendung in beliebigen Anwendungen exportiert werden. Nach dem Export können sie jedoch nur noch in IBM® SPSS® Rules Management und in keiner anderen Anwendung bearbeitet werden. Beachten Sie, dass IBM SPSS Rules Management möglicherweise nicht an allen Standorten installiert ist oder nicht allen Benutzern zur Verfügung steht.
- Die Verwendung einer bestimmten beschrifteten Version (nicht der *NEUESTEN*) wird sehr empfohlen – insbesondere in einer Produktionsumgebung –, da spätere Änderungen an der neuesten Version sonst zu unerwarteten Konsequenzen führen können. In bestimmten Fällen, z. B. bei der Referenzierung einer Regel innerhalb einer Regel, muss eine beschriftete Version verwendet werden.

- Einzelregeln können per Referenz aus IBM SPSS Collaboration and Deployment Services Repository hinzugefügt werden oder über das Dialogfeld „Regelvorschau“ in lokale Regeln umgewandelt werden.
- Einzelregeln können in einem beliebigen Regeleditor referenziert und mit lokalen Regeln kombiniert werden. Beim Export als Regelsatz werden verschachtelte Referenzen zu Einzelregeln beibehalten.
- Segment- und Selektionsregelsätze können in den meisten Fällen wie Einzelregeln referenziert und kombiniert werden. Der kombinierte Regelsatz kann jedoch erst exportiert werden, nachdem eventuelle verschachtelte Regelsätze in Einzelregeln umgewandelt wurden. Ein exportierter Regelsatz kann daher Referenzen auf verschachtelte Regeln enthalten, *aber nicht auf verschachtelte Regelsätze*.
- Aggregationsregelsätze können ebenfalls referenziert, jedoch nicht mit anderen Regeln kombiniert werden. Sie dürfen keine verschachtelten Referenzen auf andere Regeln enthalten. Wenn ein Aggregationsregelsatz referenziert wird, ersetzt die referenzierte Regel den gesamten Inhalt des Aggregationsregel-Editors. (Falls andernfalls kein Aggregationsregelsatz referenziert wird, kann der Aggregationsregel-Editor lokale Regeln, referenzierte Segmentregeln und Segmentregelsätze gleichermaßen enthalten.)
- Für IBM SPSS Rules Management gelten die gleichen Einschränkungen wie für exportierte Regeln. Verschachtelte Regeln sind zulässig, verschachtelte Regelsätze dagegen nicht. Der Grund ist, dass IBM SPSS Rules Management für die Erstellung und Bearbeitung gemeinsamer Regeln entwickelt wurde (ähnlich dem Export aus einer der anderen Anwendungen).
- Um den Einschränkungen für referenzierte Regelsätze auszuweichen – ob in IBM SPSS Rules Management oder anderswo –, können Regelsätze immer in einzelne lokale Regeln umgewandelt werden.

Anmerkungen hinzufügen

Klicken Sie in einer Aggregation oder Zuordnung auf das Symbol Anmerkungen hinzufügen, um eine Spalte mit Freitextfeldern hinzuzufügen, in die Sie Anmerkungen für die Regelausgabe eingeben können. Dies können z. B. zusätzliche Angaben für Call Center-Mitarbeiter sein, welche Ansprechperson für Schadensmeldungen zuständig ist, die als verdächtig gekennzeichnet sind.

Anmerkungen können zusammen mit dem für eine Dimension festgelegten Wert Zurückgeben mit benutzt werden, um zusätzliche Informationen bereitzustellen. Während der Wert für Zurückgeben mit auf die zurückgegebene Entscheidung oder eine Dimension bezogen ist, gibt die Anmerkung an, welche Regeln als Eingabe für diese Entscheidung wirksam waren.

Verwenden von Regeln aus dem ILOG Business Rule Management System

Regeln, die in einem Business Rules Management System wie ILOG erstellt wurden, können in IBM® Analytical Decision Management-Anwendungen referenziert und verwendet werden. Diese externen Regeln können auf dieselbe Weise ausgewählt und verwendet werden wie andere Regeln, nämlich durch Klicken auf das Symbol Vorhandene Regel suchen an der gewünschten

Stelle in der Anwendung. Externe Regeln sind an einem gesonderten Symbol zu erkennen, können jedoch ansonsten auf die standardmäßige Art und Weise verwendet werden.

Informationen zum Erstellen externer Regeln zur Verwendung in IBM Analytical Decision Management finden Sie im *Handbuch für Anwendungsentwickler*.

Die Regel wird auf dieselbe Weise in die Anwendung eingefügt wie jede andere Regel. Beachten Sie, dass die Regelausgaben dem Kontext entsprechen müssen, in dem die Regel verwendet wird. Beispielsweise ist für Zuordnungen ein Wahr-/Falsch bzw. Ja-/Nein-Wert (boolescher Wert) als Eingabe für die Entscheidung erforderlich.

Klicken Sie wie bei anderen Regeln auf den Namen einer in der Anwendung verwendeten externen Regel, um Details zu der Regel (einschließlich Eingabe und Felder) sowie die Regelbeschreibung und andere Eigenschaften anzuzeigen.

Wenn eine externe Regel mehrere Ausgabefelder aufweist, wird eine Eingabeaufforderung angezeigt, über die Sie das gewünschte Feld auswählen können.

Wenn eine externe Regel anstelle eines Modells verwendet wird, wird kein Zielfeld angezeigt und alle konfigurierten Ausgaben werden in der Dropdown-Liste angezeigt.

Tipps zur Verwendung von externen Regeln

- Regeln, die den Wert *true* (wahr) oder *false* (falsch) bzw. *yes/no* (ja/nein) oder *0/1* zurückgeben, können in einer beliebigen Auswahl oder Zuordnung, als Eingabe zu einer beliebigen Regel im Regeleditor oder als Feld vom Typ Zurückgeben mit verwendet werden, das einem bestimmten Dimensionsobjekt zugeordnet ist.
- Regeln, die kategoriale Ausgaben (*rot, weiß* oder *blau; Norden, Süden, Osten, Westen* usw.) zurückgeben, können im Abschnitt Mit einem Modell bestimmen der Registerkarte „Definieren“, als Eingabe im Regeleditor oder als Feld vom Typ Zurückgeben mit, das einem Dimensionsobjekt zugeordnet ist, verwendet werden.
- Regeln mit numerischen Ausgaben können im Abschnitt Mit einem Modell bestimmen der Registerkarte „Definieren“, als Eingabe für eine Priorisierung, als Eingabe im Regeleditor oder als Feld vom Typ Zurückgeben mit, das einem Dimensionsobjekt zugeordnet ist, verwendet werden.

Vorhersagemodelle erstellen

Informationen zu Vorhersagemodellen

Mithilfe von Modellen können wahrscheinliche zukünftige Ereignisse durch Erkennung von Mustern in Vergangenheitsdaten prognostiziert werden. Sie können ein Modell z. B. verwenden, um auf der Grundlage von Merkmalen wie Einkommen, Alter und Mitgliedschaft in bestimmten Organisationen vorherzusagen, welche Kunden mit der geringsten Wahrscheinlichkeit kündigen oder mit der höchsten Wahrscheinlichkeit auf ein bestimmtes Angebot reagieren.

Modelle können genauso wie Regeln eingesetzt werden. Während Regeln jedoch auf Unternehmensrichtlinien, Geschäftslogik oder anderen Voraussetzungen basieren, werden Modelle auf Grundlage der tatsächlichen Beobachtung früherer Ergebnisse erstellt und können Muster aufdecken, die andernfalls schwer erkennbar wären. Während Regeln in Anwendungen allgemeine betriebswirtschaftliche Logik kodieren, ermöglichen Modelle neue Erkenntnisse und Prognosen.

Daten zur Erstellung und Auswertung von Modellen

Für den Modellierungsprozess kommen zwei verschiedene Datentypen zum Einsatz:

- Zur Erstellung des Modells werden Informationen über den Gegenstand der Vorhersage benötigt. Wenn Sie z. B. die Kündigungsrate vorhersagen wollen, benötigen Sie Informationen über Kunden, die in der Vergangenheit gekündigt haben. Solche Daten werden oft als Vergangenheits- oder analytische Daten bezeichnet. Sie müssen eines oder alle Felder des Projektdatenmodells sowie ein zusätzliches Feld enthalten, welches das vorherzusagende Ergebnis erfasst. Dieses zusätzliche Feld wird als **Zielfeld** der Modellierung verwendet.
- Um ein Modell zur Vorhersage zukünftiger Ergebnisse verwenden zu können, benötigen Sie Daten über die fragliche Gruppe oder Population, z. B. potentielle Kunden oder eingehende Schadensmeldungen. Diese werden oft als **operative** Daten oder **Scoring**-Daten bezeichnet. Das Projektdatenmodell basiert typischerweise auf diesen Daten.
- Wenn das Zielfeld Teil des Datenmodells ist, sollte für dieses Feld die Spalte Operativ nicht ausgewählt werden, da es beim Scoring des Modells nicht verfügbar ist. Nachdem das Modell mit Vergangenheitsdaten erstellt wurde, hat das Scoring das Ziel, das Modell auf neue Daten anzuwenden, wo das Ergebnis nicht im Voraus bekannt ist.

Erstellen von Modellen

Auf der Registerkarte „Modellierung“ können Sie Datenquelle, Zielfeld und andere Optionen einstellen. Sie können auch festlegen, welcher Modelltyp erstellt werden soll.

Beachten Sie: Wenn der Anwendung (Stream-Datei) in IBM® SPSS® Modeler **Parameter** hinzugefügt wurden, steht auf den Registerkarten „Modellierung“ und „Scoring“ ein Parameter-Link zur Verfügung. Wenn Sie auf den Parameter-Link klicken, werden die definierten

Parameter sowie deren aktuelle Werte angezeigt. Sie können die Parameterwerte ändern, um die Modellerstellung oder das Scoring zu beeinflussen.

Sie können auf eine der folgenden Weisen ein neues Modell erstellen:

Erstellen eines Vorhersagemodells

Zur Erstellung eines Vorhersagemodells müssen nur eine Datenquelle und ein Zielfeld festgelegt werden. Sie können außerdem Optionen zur Vorbereitung und Partitionierung der Daten sowie zur Auswahl der zu benutzenden Felder und Datensätze festlegen. Das Modell wird gebildet, indem verschiedene für den Zieltyp geeignete Methoden eingesetzt werden und automatisch diejenige Verfahrensweise oder Kombination ermittelt wird, die bei den gewählten Daten die besten Ergebnisse liefert.

Beim Scoring gibt das Modell ein oder mehrere Felder mit Vorhersagen sowie die zugehörigen Konfidenzwerte zurück. Beispiel: Wenn das Zielfeld ein Flag-Feld mit der Bezeichnung *Antwort* ist, gibt das Modell ein Feld mit der Bezeichnung *XF-Antwort* zurück, das die für dieses Feld vorhergesagten Werte enthält.

- ▶ Geben Sie eine Datenquelle an. Dies kann eine analytische Datenquelle sein, in der das Ergebnis, das Sie vorhersagen möchten, aufgezeichnet ist.
- ▶ Legen Sie ein Zielfeld fest, welches das Ergebnis erfasst, das Sie vorhersagen wollen. Das Zielfeld könnte z. B. angeben, welche Kunden gekündigt, frühere Angebote akzeptiert, Versicherungsbetrug begangen haben o. Ä.

Alle anderen Felder werden automatisch als Eingabe aufgenommen und können für die Vorhersage dieses Werts nützlich sein.

- ▶ Legen Sie die gewünschten optionalen Einstellungen fest. [Für weitere Informationen siehe Thema Optionale Modelleinstellungen auf S. 58.](#)
- ▶ Klicken Sie auf **Modell erstellen**.

Während das Modell erstellt wird, können Sie den Browser schließen oder andere Aufgaben ausführen. Nach Abschluss des Erstellungsvorgangs kann auf das Modell über die Galerie zugegriffen werden. [Für weitere Informationen siehe Thema Galerie in Kapitel 3 auf S. 22.](#) Falls die Modellerstellung länger als erwartet dauert, können Sie auf die Schaltfläche **Stop** klicken, um zur zuletzt gespeicherten Version zurückzukehren.

Nach Abschluss der Modellerstellung werden Verteilungs- und Prädiktoreinfluss-Diagramme angezeigt. [Für weitere Informationen siehe Thema Ergebnisse der Vorhersagemodellierung auf S. 52.](#)

- ▶ Optional können Sie mit der Auswertungs- und Testfunktion kontrollieren, wie das Modell bei Ihren Testdaten abschneidet. [Für weitere Informationen siehe Thema Auswertung von Modellen auf S. 62.](#)

Ergebnisse der Vorhersagemodellierung

Wenn ein Vorhersagemodell fertiggestellt wird, werden Verteilungs- und Variableneinflussdiagramme angezeigt, die eine schnelle visuelle Einschätzung der Leistungsfähigkeit des Modells sowie der Bedeutung der verschiedenen Prädiktoren für das Endergebnis ermöglichen.

Verteilungsdiagramme

Die Verteilungsdiagramme des Modells zeigen die beobachteten Werte auf der horizontalen Achse und die vorhergesagten Werte auf der vertikalen Achse oder Überlagerung an. Daran können Sie ablesen, wie oft jeder Wert oder jede Antwort vorhergesagt wurde und wie oft diese Vorhersage richtig war, was eine schnelle visuelle Bewertung der Modellgenauigkeit erlaubt.

- Für Flag- oder Set-Ziele wird die Verteilung in Form eines Balkendiagramms mit einem Balken für jede Kategorie des Zielfelds angezeigt. Die Höhe eines Balkens gibt an, wie viele Male das jeweilige Ergebnis vom Modell vorhergesagt wird. Die farbigen Bereiche der einzelnen Balken drücken aus, wie viele Male das entsprechende Ergebnis tatsächlich im Testdatensatz beobachtet wurde. Wenn das Modell hinreichend präzise ist, sollte der größte Bereich der Balken mit der entsprechenden Antwort übereinstimmen.
- Für Bereichsziele wird die Verteilung durch ein klassiertes Streudiagramm dargestellt. Auch in diesem Fall werden die vorhergesagten Werte auf der horizontalen Achse und die beobachteten Werte auf der vertikalen Achse aufgetragen. Idealerweise gruppieren sich die Punkte entlang einer diagonalen Linie. Je breiter die Punkte (Ergebnisse) verteilt sind und je weiter sie von einer linearen Verteilung abweichen, desto weniger genau ist das Modell.

Prädiktoreinfluss-Diagramme

Normalerweise ist es sinnvoll, sich bei der Modellierung auf die wichtigsten Prädiktorfelder zu konzentrieren und diejenigen zu verwerfen bzw. zu ignorieren, die die geringste Bedeutung haben. Das Prädiktor-Einflussdiagramm unterstützt Sie dabei, indem es die relative Wichtigkeit der einzelnen Prädiktoren bei der Modellschätzung darstellt. Da die Werte relativ sind, ist die Summe der Werte für alle angezeigten Prädiktoren 1,0. Der Einfluss der Prädiktoren steht in keiner Beziehung zur Modellgenauigkeit. Sie bezieht sich lediglich auf die Bedeutsamkeit der einzelnen Prädiktoren für die Erstellung von Vorhersagen, nicht darauf, ob die Vorhersage genau ist oder nicht.

Die Berechnung des Prädiktor-Einflusses kann erheblich länger dauern als die Modellerstellung, insbesondere bei großen Daten-Sets.

Falls bei der Modellerstellung die Option zur automatischen Datenpartitionierung aktiviert wurde, wird der Testdatensatz automatisch auf der Seite „Auswerten“ ausgewählt.

Falls die Datenpartitionierung nicht aktiviert wurde, müssen Sie für die Modellauswertung eine Datenquelle auswählen. Hierbei wird empfohlen, eine andere als die zur Modellerstellung verwendete Datenquelle auszuwählen. Beachten Sie, dass die Datenpartitionierung standardmäßig aktiviert ist. Dies ist daher nur notwendig, falls Sie diese Einstellung geändert haben.

Erstellen eines Vorhersageregelmodells (interaktiv)

Mit Vorhersageregelmodellen (interaktiv) können Sie Segmente ausfindig machen, die in Bezug auf ein bestimmtes Resultat eine höhere oder geringere Wahrscheinlichkeit aufweisen. Sie könnten beispielsweise nach Kunden suchen, deren Abwanderung besonders unwahrscheinlich ist oder die mit der größten Wahrscheinlichkeit auf ein Angebot ansprechen. Das resultierende Modell ist einer Liste von Selektionsregeln mit Einschluss- und Ausschluss-Attributen ähnlich, die angeben, wie die verschiedenen Segmente gescort wurden.

Vorhersageregelmodelle (interaktiv) sind Selektionsregeln ähnlich, werden aber auf andere Weise gescort. Im Fall von Selektionsregeln werden ausgeschlossene Datensätze entfernt, so dass die Ausgabe weniger Datensätze enthält. Im Fall eines Modells werden keine Datensätze entfernt, sondern zusätzliche Felder für die Vorhersage- und Neigungswerte der Datensätze hinzugefügt. Eingeschlossene Segmente werden als *Wahr* (oder *Falsch* oder was auch immer Sie vorhersagen wollen) gescort. Ausschlossene Segmente werden als *NULL* gescort, ebenso wie alle Datensätze, die dem Rest zugeordnet werden, da Sie keine Regel erfüllen.

Jedem Datensatz wird außerdem ein Neigungswert zugeordnet, der die Wahrscheinlichkeit der Zielantwort angibt. Dadurch ist es möglich, z. B. die 10% der Kunden mit der höchsten Antwortwahrscheinlichkeit zu ermitteln. Vorhersageregelmodelle (interaktiv) sind somit gut zur Erstellung von Mailinglisten geeignet und im Customer Relationship Management weit verbreitet, unter anderem in Call Center- und Marketinganwendungen.

- Die Vorhersageregelmodelle (interaktiv) sind nur für kategoriale Ziele verfügbar.
- Sie können Segmente anzeigen, bearbeiten, neu anordnen und auswählen, welche beim Scoring ein- bzw. ausgeschlossen werden sollen. Sie können beispielsweise eine Gruppe von Kunden von zukünftigen Angeboten aus- und andere einschließen und sofort ablesen, wie dies die Gesamttrefferquote beeinflusst.
- Wie auch bei Selektionsregeln können Segmente nach Bedarf manuell hinzugefügt oder bearbeitet werden. Oft ist es jedoch einfacher, sie vom Modell ermitteln zu lassen. Es stehen eine Reihe von Optionen zur Verfügung, die Ihnen die Feinabstimmung des Resultats erleichtern.
- Die Regeln werden in der angegebenen Reihenfolge angewendet. Dabei legt die erste zutreffende Regel das Ergebnis für einen bestimmten Datensatz fest. Für sich genommen können sich Regeln bzw. Bedingungen überschneiden, durch die Reihenfolge der Regeln wird die Mehrdeutigkeit jedoch aufgelöst. Wenn keine Regel zutrifft, wird der Datensatz der Restregel zugewiesen.

So rufen Sie ein Vorhersageregelmodell (interaktiv) ab:

- ▶ Geben Sie eine Datenquelle an. Dies kann eine analytische Quelle sein, in der das Ergebnis, das Sie vorhersagen möchten, erfasst wird.
- ▶ Legen Sie ein Zielfeld fest, welches das Ergebnis erfasst, das Sie vorhersagen wollen.
- ▶ Geben Sie den gesuchten Wert oder die gesuchte Antwort an. Wenn der Name des Zielfelds beispielsweise *Antwort* lautet, können Sie den Wert *Wahr* oder *Falsch* auswählen, um Kunden aufzufinden, die das Angebot angenommen bzw. abgelehnt haben.

Beispielsweise kann das Zielfeld angeben, welche Kunden gekündigt haben, auf frühere Angebote eingegangen sind, Schadensmeldungen fingiert haben o. Ä. Alle anderen Felder werden automatisch als Eingaben aufgenommen, die sich bei der Vorhersage dieses Wertes als hilfreich erweisen könnten.

- ▶ Legen Sie die gewünschten optionalen Einstellungen fest. [Für weitere Informationen siehe Thema Optionale Modelleinstellungen auf S. 58.](#)
- ▶ Klicken Sie auf Erstellung beginnen.
- ▶ Geben Sie an, ob Sie Segmente mit hoher oder niedriger Wahrscheinlichkeit in Bezug auf das angegebene Ziel finden möchten. Sie könnten z. B. nach Gruppen oder Kunden suchen, die mit der höchsten oder geringsten Wahrscheinlichkeit kündigen, und sie dementsprechend einschließen.
- ▶ Geben Sie die maximale Anzahl zu suchender Segmente an. Üblicherweise wird aus Gründen der Einfachheit und Leistung eine möglichst niedrige Zahl gewählt.
- ▶ Klicken Sie auf Modell erweitern, um eine Liste mit ein oder mehreren Segmentregeln zu erhalten, die Gruppen mit höherer oder geringerer Wahrscheinlichkeit aufzeigen.
- ▶ Klicken Sie auf das Einschließen- bzw. Ausschließen-Symbol, um festzulegen, wie die Regeln verwendet und die entsprechenden Segmente gescort werden sollen.

Eingeschlossene Segmente erhalten einen Score, der dem Zielwert entspricht (z. B. *Antwort=wahr*). Ausgeschlossene Segmente erhalten den Score *NULL*, werden aber nicht aus der Datei entfernt.

- ▶ Optional können Sie Segmente nach Bedarf manuell hinzufügen und bearbeiten. [Für weitere Informationen siehe Thema Segmentregeln definieren in Kapitel 5 auf S. 43.](#)
- ▶ Optional können Sie mit der Auswertungs- und Testfunktion kontrollieren, wie das Modell bei Ihren Testdaten abschneidet.

Tipps

- Mit den Pfeilschaltflächen können Sie Regeln nach oben oder unten verschieben. Jeder Datensatz wird gemäß dem ersten Treffer gescort. Das Ergebnis hängt daher von der Reihenfolge der Regeln ab. Wenn ein Datensatz z. B. mehrere Regeln erfüllt, wird er gemäß der ersten zutreffenden Regel gescort.
- Ändern Sie eine oder mehrere Optionen, und klicken Sie erneut auf Modell erweitern, um zusätzliche Regeln zu suchen. Sie könnten z. B. den Schwellenwert für die minimale Segmentgröße senken oder das Konfidenzintervall verkleinern, um zusätzliche Segmente zu erhalten.
- Wenn Sie auf Modell erweitern klicken, wird dadurch nicht die Tabelle gelöscht oder das Modell von Grund auf neu erstellt – es werden lediglich weitere Segmentregeln zur Aufnahme in die Tabelle gesucht.

Erstellen eines Clustermodells

Mit Clustermodellen können Datensätze in eine bestimmte Anzahl von Clustern unterteilt werden. So können Sie natürliche Gruppen in Ihren Daten bestimmen.

Clustermodelle konzentrieren sich auf die Ermittlung ähnlicher Datensätze und Beschriftung der Datensätze anhand der Gruppe, in die sie gehören. Dies geschieht ohne Vorkenntnisse zu Gruppen und ihren Eigenschaften. Vielleicht wissen Sie nicht einmal, nach wie vielen Gruppen Sie suchen sollen. Hierin liegt der Unterschied der Clustermodelle zu anderen Techniken des Maschinlernens: Es gibt keine zuvor definierte Ausgabe und kein Zielfeld für das vorherzusagende Modell. Diese Modelle werden häufig als **nicht überwachte Lernmodelle** bezeichnet, da es keinen externen Standard gibt, mit dem die Klassifizierungsleistung des Modells beurteilt werden könnte. Für diese Modelle gibt es keine *richtigen* oder *falschen* Antworten. Ihr Wert wird durch die Möglichkeit bestimmt, interessante Gruppierungen in den Daten zu erfassen und sinnvolle Beschreibungen dieser Gruppierungen zu liefern.

Methoden zur Clusterbildung basieren auf dem Messen der Entfernungen zwischen Datensätzen und Clustern. Die Datensätze werden den Clustern auf eine Weise zugewiesen, die die Entfernung zwischen den Datensätzen minimiert, die demselben Cluster angehören.

Cluster-Modelle werden häufig verwendet, um Cluster oder Segmente zu erstellen, die dann als Eingaben in nachfolgenden Analysen verwendet werden. Ein häufiges Beispiel dafür sind die von Marktforschern verwendeten Marktsegmente, mit denen der Gesamtmarkt in homogene Untergruppen aufgeteilt wird. Jedes Segment weist besondere Eigenschaften auf, die sich auf den Erfolg der Marktforschung auswirken. Wenn Sie Data-Mining zur Optimierung Ihrer Marketingstrategie verwenden, können Sie Ihr Modell in der Regel erheblich verbessern, indem Sie die entsprechenden Segmente ermitteln und diese Segmentinformationen für Ihre Vorhersagemodelle verwenden.

So erstellen Sie ein Clustermodell:

- ▶ Geben Sie eine Datenquelle an. Dies kann eine analytische Quelle sein, in der das Ergebnis, das Sie vorhersagen möchten, erfasst wird.
- ▶ Legen Sie die gewünschten optionalen Einstellungen fest. [Für weitere Informationen siehe Thema Optionale Modelleinstellungen auf S. 58.](#)
- ▶ Klicken Sie auf Cluster suchen.
- ▶ Optional können Sie manuelle Cluster hinzufügen. [Für weitere Informationen siehe Thema Verwenden von manuellen Clustern auf S. 55.](#)
- ▶ Optional können Sie mit der Auswertungs- und Testfunktion kontrollieren, wie das Modell bei Ihren Testdaten abschneidet.

Verwenden von manuellen Clustern

Bei Bedarf können Cluster manuell definiert werden. Im Anschluss an die Verwendung der manuellen Cluster wird dann die automatische Clusterbildung ausgeführt, um nach Clustern von Datensätzen zu suchen, die von keinem der manuell definierten Cluster erkannt wurden. Erstellen Sie manuelle Cluster über einen vom ersten Treffer ausgehenden segmentbasierten Zuordnungsregelsatz. Anhand der von Ihnen erstellten Regeln werden Cluster definiert. Die verbleibenden Datensätze, die keinem Cluster zugeordnet werden konnten, werden dann im Rahmen der automatischen Clustermodellierung gruppiert.

Beachten Sie, dass manuelle Cluster nicht in Ergebnisübersichtsdiagrammen berücksichtigt werden.

- ▶ Klicken Sie im Abschnitt Manueller Cluster auf das Symbol Neue Regel erstellen, und fügen Sie eine oder mehrere Regeln zum Definieren von manuellen Clustern hinzu.
- ▶ Ändern Sie bei Bedarf die Beschriftungsnamen der Cluster.
- ▶ Klicken Sie auf das Symbol Datensätze in manuellen Clustern, um die Gesamtanzahl der in den manuellen Clustern enthaltenen Datensätze zu bestimmen. Daraufhin wird die Spalte Anzahl für jeden definierten manuellen Cluster aktualisiert.

Erstellen eines Assoziationsmodells

Assoziationsmodelle ermöglichen eine Vorhersage darüber, welche Elemente höchstwahrscheinlich nah beieinander liegen und wie stark die Beziehung zwischen den Elementen ist.

Assoziationsregeln ordnen eine bestimmte Schlussfolgerung (z. B. den Kauf eines bestimmten Produkts) einer Menge von Bedingungen (z. B. dem Kauf mehrerer anderer Produkte) zu.
Beispiel: Die Regel

Bier \leq -Dosengemüse & TK-Fertiggericht (173, 17,0 %, 0,84)

besagt, dass *Bier* häufig vorkommt, wenn *Dosengemüse* und *TK-Fertiggericht* zusammen vorkommen. Die Regel ist zu 84 % zuverlässig und auf 17 % der Daten, also 173 Datensätze, anwendbar. Algorithmen für Assoziationsregeln finden automatisch die Zuweisungen, die Sie manuell finden könnten, wenn Sie Visualisierungstechniken, wie den Webknoten, anwenden.

Der Vorteil von Algorithmen für Assoziationsregeln im Vergleich zu Algorithmen für Standardentscheidungsbäume (C5.0 und C&R-Bäume) besteht darin, dass zwischen *beliebigen* Attributen Verbindungen bestehen können. Ein Entscheidungsbaumalgorithmus erstellt Regeln mit nur einer Schlussfolgerung, während Assoziationsalgorithmen viele Regeln zu finden versuchen, von denen jede zu einer anderen Schlussfolgerung kommen kann.

Der Nachteil von Assoziationsregeln besteht darin, dass sie versuchen, Muster innerhalb eines potenziell sehr großen Suchbereichs zu finden, also mehr Zeit für die Ausführung in Anspruch nehmen können als ein Entscheidungsbaumalgorithmus. Die Algorithmen verwenden eine Methode vom Typ **Generieren und Testen** zum Auffinden von Regeln, bei der einfache Regeln erstellt und mit dem Daten-Set verglichen werden. Die "guten" Regeln werden gespeichert und alle Regeln – nachdem sie verschiedenen Beschränkungen unterworfen wurden – werden dann weiter verfeinert. **Spezialisierung** ist der Prozess, bei dem einer Regel Bedingungen hinzugefügt werden. Diese neuen Regeln werden dann mit den Daten verglichen und validiert. Die "besten" oder interessantesten Regeln werden dann gespeichert. Der Benutzer legt normalerweise einen Grenzwert für die mögliche Anzahl Antezedenzen in einer Regel fest. Es werden außerdem verschiedene Techniken basierend auf der Informationstheorie oder effiziente Indizierungsschemata verwendet, um den potenziell großen Suchbereich zu reduzieren.

So erstellen Sie ein Assoziationsmodell:

- ▶ Wählen Sie eine Datenquelle aus. Dies kann eine analytische Quelle sein, in der das Ergebnis, das Sie vorhersagen möchten, erfasst wird.
- ▶ Geben Sie ein Datenformat an. [Für weitere Informationen siehe Thema Vergleich von Tabellen- und Transaktionsdaten auf S. 57.](#)
- ▶ Legen Sie die gewünschten optionalen Einstellungen fest. [Für weitere Informationen siehe Thema Optionale Modelleinstellungen auf S. 58.](#)
- ▶ Klicken Sie auf Modell erstellen.
- ▶ Optional können Sie mit der Testfunktion prüfen, wie das Modell bei Ihren Testdaten abschneidet. Beachten Sie, dass für Assoziationsmodelle keine Auswertungsfunktion verfügbar ist.

Vergleich von Tabellen- und Transaktionsdaten

Bei der Erstellung von Assoziationsmodellen kann zwischen den Datenformaten Tabellendaten und Transaktionsdaten gewählt werden. Standardmäßig sind Tabellendaten ausgewählt. Wenn Sie Transaktionsdaten auswählen, wählen Sie in den angezeigten Dropdownmenüs die Felder ID und Inhalt aus.

Transaktionsformat

Transaktionsdaten weisen einen eigenen Datensatz für jede Transaktion bzw. jedes Element auf. Wenn ein Kunde beispielsweise mehrere Einkäufe tätigt, handelt es sich bei jedem um einen eigenen Datensatz, wobei die zugehörigen Elemente durch eine Kunden-ID verknüpft sind. Dies wird auch manchmal als **Kassenrollen**-Format bezeichnet.

Kunde	Kauf
1	Marmelade
2	Milch
3	Marmelade
3	Brot
4	Marmelade
4	Brot
4	Milch

Tabellendaten

In Tabellendaten (auch als **Warenkorb**- oder **Wahrheitstabellen**-Daten bezeichnet) werden die Elemente durch gesonderte Flags dargestellt, wobei jedes Flag-Feld für das Vorliegen bzw. die Abwesenheit eines bestimmten Elements steht. Jeder Datensatz steht für ein komplettes

Set zugehöriger Elemente. Prinzipiell können Flag-Felder kategorial oder numerisch sein; bei manchen Modellen können jedoch genauere Anforderungen gelten.

Kunde	Marmelade	Brot	Milch
1	T	F	F
2	F	F	T
3	T	T	F
4	T	T	T

Optionale Modelleinstellungen

Beim Erstellen von Modellen können Sie aus den folgenden optionalen Einstellungen wählen. Änderungen dieser Einstellungen treten bei einem bestehenden Modell erst in Kraft, nachdem es neu erstellt wurde.

Daten zur Aktivierung der Modellvalidierung beim Erstellen der Datenquelle automatisch partitionieren. Teilt die Daten in separate Teilmengen oder Stichproben zum Trainieren und Testen des Modells ein. Wenn Sie das Modell auf der Grundlage einer Teilmenge aufbauen und anhand einer anderen Teilmenge testen, erhalten Sie eine ungefähre Vorstellung davon, inwieweit es sich auf andere Daten-Sets verallgemeinern lässt.

Daten automatisch bereinigen und zur zuverlässigen Modellerstellung aufbereiten. Findet und repariert Datenfehler, um die Modellierung schneller, voraussehbarer und zuverlässiger zu machen. Filtert problematische oder wahrscheinlich nicht nützliche Felder, indem beispielsweise durch intelligente Filter- und Probenentnahmeverfahren Verzerrungen durch fehlende Werte oder Ausreißer verhindert und nach Bedarf neue Attribute abgeleitet werden, um eine bessere Leistung zu erzielen. Wenn ein Modell zum ersten Mal mit einer neuen Datenquelle aufgebaut wird, folgt eine Analyse zur Identifikation von Problemen und Abhilfemaßnahmen, wodurch der Prozess lediglich beim ersten Durchgang eventuell verzögert wird. Bei späteren Durchläufen werden die Probleme beseitigt, die Analyse aber nur wiederholt, wenn die Datenquelle modifiziert wurde. Diese Einstellung kann für bestimmte Modelle deaktiviert werden, u. a. für solche, deren Datenvorbereitung von einem erfahrenen Anwender speziell angepasst wurde.

Zu benutzende Eingabedaten festlegen. Ermöglicht Ihnen, die Felder auszuwählen, die sie verwenden möchten. Meist werden dies diejenigen sein, die einen praktischen Bezug zum Gegenstand der Prognose haben, z. B. Alter oder Einkommen. Wenn Sie über große Datenmengen verfügen, können Sie das Modell durch Begrenzen der Feldanzahl vereinfachen. Felder wie Kunden- oder Kontakt-Nr. in den Daten sind für die Modellierung normalerweise nicht von Nutzen und sollten nicht ausgewählt werden. Felder, die mit anderen Daten identisch sind, können ebenfalls ausgeschlossen werden.

Wenn Sie auf ein verknüpftes Eingabefeld (einen Ausdruck) klicken, wird der Ausdrucksbetrachter für diesen Ausdruck angezeigt. Zum Bearbeiten eines Ausdrucks verwenden Sie die Registerkarte „Daten“. [Für weitere Informationen siehe Thema Ausdruckseditor in Kapitel 4 auf S. 41.](#)

Wählen Sie die zu verwendenden Selektionen. Legt fest, welche Datensätze bei der Modellierung einbezogen und welche ausgeschlossen werden sollen. Sie können je nach Bedarf vorhandene Regeln suchen oder neue Regeln erstellen. [Für weitere Informationen siehe Thema Selektionsregeln definieren in Kapitel 5 auf S. 45.](#)

Wenn globale Selektionen definiert wurden, werden diese hier ebenfalls angezeigt, und Sie können angeben, ob sie auch während der Modellierung angewendet werden sollen. Sie können festlegen, ob alle oder keine globalen Selektionsregeln verwendet werden sollen. Es ist nicht möglich, nur eine Teilmenge davon auszuwählen.

Segmentierungsoptionen

Beim Erstellen von Vorhersageregeln (interaktiv) steht eine Reihe von Segmentierungsoptionen zur Verfügung. Diese sind zusätzlich zu den Optionen für andere Modelltypen verfügbar.

Mindestsegmentgröße. Diese Einstellungen bestimmen die minimale Segmentgröße, wobei der größere der beiden Werte ausgewählt wird. Wenn z. B. der Prozentwert einer höheren Zahl entspricht als der absolute Wert, wird die prozentuale Einstellung ausgewählt.

- **Prozentsatz des vorherigen Segments.** Legt die minimale Gruppengröße als Prozentsatz der Datensätze fest. Die zulässige Minimaleinstellung liegt bei 0, die zulässige Maximaleinstellung bei 99,9.
- **Absoluter Wert.** Legt die minimale Gruppengröße als absolute Anzahl der Datensätze fest. Die zulässige Minimaleinstellung liegt bei 1, eine Maximaleinstellung ist nicht vorhanden.

Maximale Anzahl der Attribute. Legt die maximale Anzahl von Bedingungen pro Segmentregel fest. Die zulässige Minimaleinstellung liegt bei 1, eine Maximaleinstellung ist nicht vorhanden.

Wiederverwendung von Attributen zulassen. Wenn diese Option aktiviert ist, können in jedem Zyklus alle Attribute berücksichtigt werden, auch solche, die in vorherigen Zyklen verwendet wurden. Die Bedingungen für ein Segment werden zyklisch aufgebaut, wobei in jedem Zyklus eine neue Bedingung hinzugefügt wird. Die Anzahl der Zyklen wird mithilfe der Einstellung Maximale Anzahl der Attribute festgelegt.

Konfidenzintervall für neue Bedingungen. Legt das Konfidenzintervall für das Testen der Signifikanz des Segments fest. Diese Einstellung spielt eine wichtige Rolle für die Anzahl der gegebenenfalls zurückgegebenen Segmente sowie für die „Anzahl der Bedingungen pro Segment“-Regel. Je höher der Wert, desto kleiner ist der zurückgegebene Ergebnissatz. Die zulässige Minimaleinstellung liegt bei 50, die zulässige Maximaleinstellung bei 99,9.

Optionen für die automatische Clusterbildung

Beim Erstellen von Clustermodellen steht eine Reihe von Optionen für die automatische Clusterbildung zur Verfügung. Diese sind zusätzlich zu den Optionen für andere Modelltypen verfügbar.

Auswertungsfeld. Legt das Feld fest, das für die Auswertung verwendet werden soll. Wenn ein Auswertungsfeld ausgewählt wurde, wird die Sortiermethode im Knoten für die automatische Clusterbildung in IBM® SPSS® Modeler auf *Bedeutung der Auswertung* gesetzt. Wenn kein Auswertungsfeld ausgewählt wurde, bleibt die standardmäßige Sortiermethode bestehen. Beachten Sie, dass hier auch ein Analysefeld als Auswertungsfeld ausgewählt werden kann.

Gewünschten Bereich für die Anzahl der gefundenen Cluster einstellen. Legt die Mindest- und Höchstzahl an Clustern fest, die bei der automatischen Clusterbildung gefunden werden sollen.

Optionen für Assoziationsregeln und Scoring

Beim Erstellen von Assoziationsmodellen steht eine Reihe von Optionen für Assoziationsregeln und Scoring zur Verfügung. Diese sind zusätzlich zu einigen der Optionen für andere Modelltypen verfügbar.

Optionen für Assoziationsregeln.

- **Abdeckungsbereich (%).** Legt den Mindestabdeckungsbereich in Prozent fest, über den die allgemeine Gültigkeit der Regeln bestimmt wird. Ein Wert von 80 % beispielsweise bedeutet, dass alle Regeln für mindestens 80 % der Fälle gelten. Beachten Sie, dass höhere Werte weniger oder möglicherweise gar keine Regeln ergeben.
- **Konfidenz (%).** Legt die Mindestkonfidenz in Prozent fest.
- **Maximale Anzahl an Vorbedingungen.** Legt die maximale Anzahl der Vorbedingungen fest.
- **Bereits nach ID sortierte Daten (nur Transaktionsdaten).** Wählen Sie diese Option aus, um die standardmäßige Sortierung beim Start der Modellerstellung zu deaktivieren, damit zusammengehörige IDs aneinander angrenzen. Durch Auswahl dieser Option wird die Verarbeitungszeit verringert. Bei Verwendung von Tabellendaten hat diese Option keine Auswirkungen.

Scoring-Optionen.

- **Maximale Anzahl an Prognosen.** Legen Sie die maximale Anzahl der zu berücksichtigenden Prognosen fest. Diese Einstellung wird in Verbindung mit dem Dropdown-Menü „Regelkriterium“ verwendet, um „Top“-Prognosen zu erstellen. Dabei steht „Top“ für das höchste Niveau an Konfidenz, Unterstützung, Lift usw.
- **Regelkriterium.** Dient zur Auswahl des Maßes, das zur Ermittlung der Stärke der Regeln verwendet wird. Die Regeln werden nach der Stärke der hier ausgewählten Kriterien sortiert, um die Top-Prognosen für ein Element-Set auszugeben. Folgende Kriterien stehen zur Verfügung:

Konfidenz zeigt das Verhältnis von Regelunterstützung zu Regelbedingungs-Unterstützung an. Es wird also der Anteil der Datensätze mit der/den festgelegten Regelbedingungs-Unterstützung(en) angegeben, bei der/denen auch die Regelprognose(n) wahr ist/sind. Beispiel: Wenn 50 % der Trainingsdaten „Brot“ enthalten (Regelbedingungs-Unterstützung), jedoch nur 20 % sowohl „Brot“ als auch „Käse“ (Regelunterstützung), ist die Prognose für die Regel $\text{Brot} \rightarrow \text{Käse}$ $\frac{\text{KäseRegelunterstützung}}{\text{Regelbedingungs-Unterstützung}}$, in diesem Fall 40 %.

Abdeckungsbereich (%) zeigt die Regelbedingungs-Unterstützung an, also den Anteil der IDs, für den die Regelbedingungen auf der Grundlage der Trainingsdaten wahr sind. Beispiel: Wenn 50 % der Trainingsdaten den Kauf von Brot beinhalten, hat die Regel `Brot -> Käse` eine Regelbedingungs-Unterstützung von 50 %.

Regelunterstützung (%) zeigt den Teil der Datensätze an, für die die gesamte Regel, die Regelbedingung(en) und die Regelprognose(n) wahr sind. Beispiel: Wenn 20 % der Trainingsdaten sowohl „Brot“ als auch „Käse“ enthalten, beträgt die Regelunterstützung für die Regel `Brot -> Käse` 20 %.

Lift zeigt das Verhältnis zwischen der Konfidenz für die Regel und der A Priori-Wahrscheinlichkeit für die Regelprognose an. Beispiel: Wenn 10 % der Gesamtbevölkerung Brot kauft, weist eine Regel, die mit einer Konfidenz von 20 % vorhersagt, ob die Leute Brot kaufen, einen Lift von $20/10 = 2$ auf. Wenn eine andere Regel mit einer Konfidenz von 11 % besagt, dass die Leute Brot kaufen, weist diese Regel einen Lift von annähernd 1 auf, was bedeutet, dass die Regelbedingungen keinen großen Unterschied für die Wahrscheinlichkeit der Regelprognose ausmachen. Im Allgemeinen sind Regeln mit einem Lift, der sich von 1 unterscheidet, interessanter als Regeln mit einem Lift von annähernd 1.

Verwendbarkeit (%) ist ein Maß dafür, welcher Prozentsatz der Trainingsdaten die Regelbedingungen, nicht jedoch die Regelprognose erfüllt. Beim Einkauf von Produkten bedeutet dies im Grunde, welcher Prozentsatz des Kundenstamms das Produkt aus den Regelbedingungen besitzt (bzw. erworben hat), jedoch noch nicht das in den Regelprognosen festgelegte Produkt gekauft hat. Die Verwendbarkeitsstatistik ist definiert als $([\text{Regelbedingungs-Unterstützung in Anzahl Datensätze} - \text{Regelunterstützung in Anzahl Datensätze}] / \text{Anzahl Datensätze}) * 100$. Dabei ist Regelbedingungs-Unterstützung die Anzahl der Datensätze, bei denen die Regelbedingungen wahr sind, und Regelunterstützung die Anzahl der Datensätze, bei denen sowohl die Regelbedingungen als auch die Prognose wahr sind.

■ **Gesamte Transaktion prüfen.** Wählen Sie aus den folgenden Optionen aus:

Auf fehlende Vorhersagen prüfen - Stellt sicher, dass keine Regelprognosen vorhanden sind. Beispiel: Wenn das Ziel des Scorings darin besteht, eine Produktempfehlung für Möbel abzugeben, ist es unwahrscheinlich, dass bei einem Warenkorb, der bereits einen Esszimmertisch enthält, ein weiterer erworben wird. In solchen Fällen sollten Sie diese Option auswählen. Andererseits können bei verderblichen Produkten und Einwegartikeln (wie Käse, Säuglingsnahrung oder Papiertaschentüchern) Regeln, bei denen das Sukzedens bereits im Warenkorb vorhanden ist, sinnvoll sein. Im letzteren Fall ist die nützlichste Option möglicherweise die Option Nicht auf Vorhersagen prüfen.

Auf vorhandene Vorhersagen prüfen - Stellt sicher, dass im Warenkorb auch Regelprognosen vorhanden sind. Dieser Ansatz ist sinnvoll, wenn Sie versuchen, einen Einblick in bestehende Kunden oder Transaktionen zu gewinnen. Sie könnten beispielsweise die Regeln mit dem höchsten Lift ermitteln und dann untersuchen, auf welche Kunden diese Regeln zutreffen.

Bei der Option Nicht auf Vorhersagen prüfen werden beim Scoring alle Regeln einbezogen, egal ob Regelprognosen im Warenkorb vorhanden sind oder nicht.

- **Wiederholungsprognosen zulassen.** Wählen Sie diese Option aus, um mehrere Regeln mit derselben Regelprognose beim Scoring aufzunehmen. Bei Auswahl dieser Option würden beispielsweise folgende Regeln gescort werden:

Brot und Käse -> Wein

Käse und Obst -> Wein

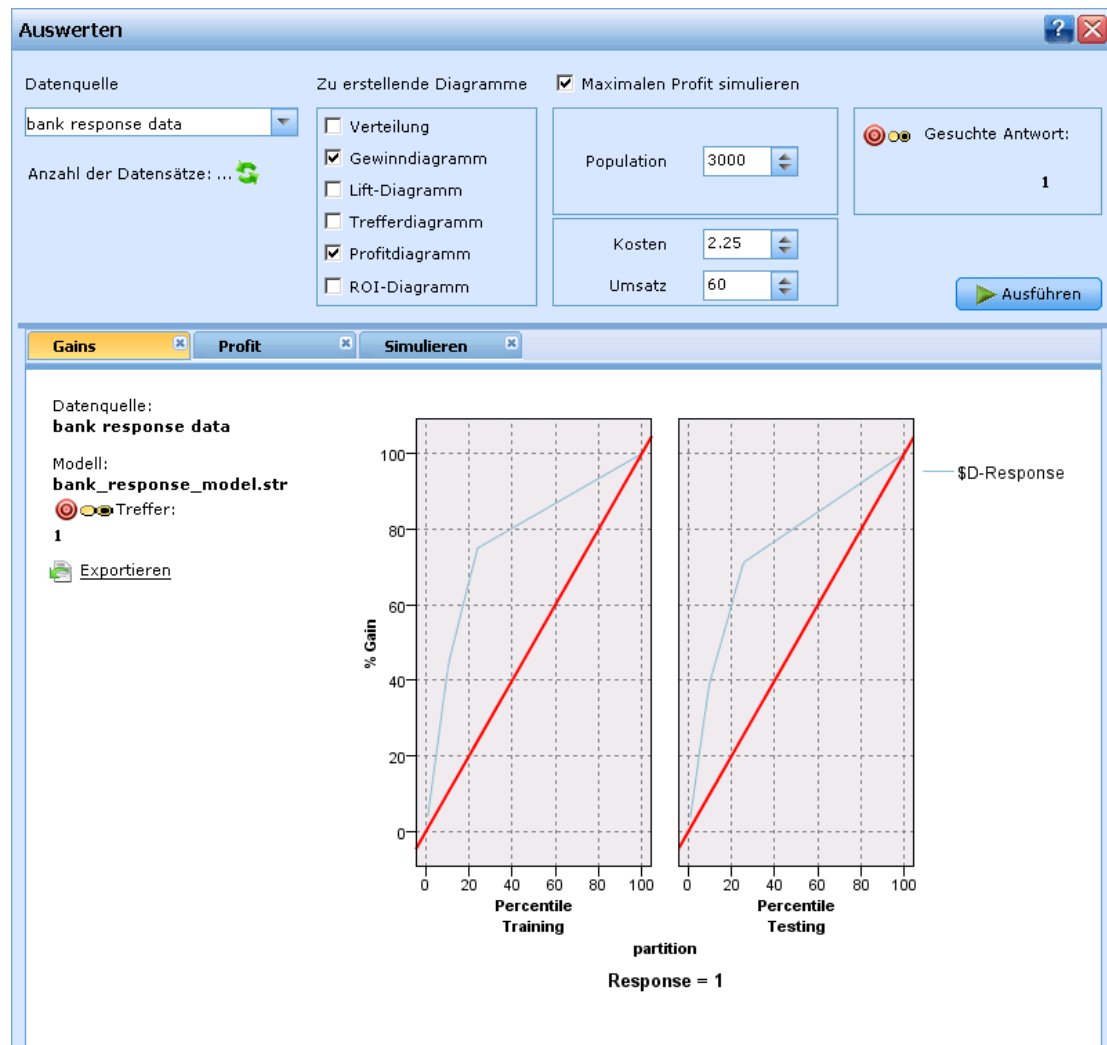
Deaktivieren Sie diese Option, um beim Scoring Wiederholungsprognosen auszuschließen.

- **Nicht übereinstimmende Elemente ignorieren.** Wählen Sie diese Option, um das Vorliegen zusätzlicher Elemente im Element-Set zu ignorieren. Wenn diese Option beispielsweise für einen Warenkorb ausgewählt wurde, der [Zelt und Schlafsack und Wasserkessel] enthält, gilt die Regel Zelt und Schlafsack -> Gaskocher trotz des zusätzlichen Elements (Wasserkessel) im Warenkorb.

Auswertung von Modellen

Klicken Sie nach dem Erstellen oder Öffnen eines fertigen Modells auf Auswerten, um Verteilungs-, Gains-, Lift-, Antwort-, Profit- und ROI-Diagramme zu erstellen und den maximal möglichen Profit zu simulieren. Dieses Dialogfeld steht nur zur Verfügung, wenn das gewählte Modell fertiggestellt ist. Beachten Sie, dass für Assoziationsmodelle keine Auswertungsfunktion verfügbar ist.

Abbildung 6-1
Gewinndiagramm



- ▶ Wählen Sie im oberen Bereich ein Daten-Set für die Auswertung aus. Wie bei der Modellerstellung muss dies ein analytisches Daten-Set sein, dessen tatsächliche Antworten bekannt sind. (Andernfalls wäre es nicht möglich, tatsächliche Antworten mit den vom Modell vorhergesagten Antworten zu vergleichen.)
- ▶ Wählen Sie die Diagramme aus, die Sie erstellen möchten.
- ▶ Wenn Sie den Profit simulieren, wählen Sie das Feld **Maximalen Profit simulieren** und geben die Population ein, auf die sich die Profitrechnung stützen soll. [Für weitere Informationen siehe Thema Maximalen Profit simulieren auf S. 69.](#)
- ▶ Wenn Sie ein Profit- oder ROI-Diagramm anfordern oder den Profit simulieren, geben Sie Kosten- und Umsatzwerte ein. Diese Werte geben die Kosten für jedes Angebot sowie das zu erwartende Umsatzvolumen pro positiver Antwort an.

- ▶ Wählen Sie den Antwortwert aus, der als „Treffer“ betrachtet werden soll. Hierbei muss es sich um einen der Werte für das bei der Modellerstellung verwendete Zielfeld handeln, z .B. *Antwort = wahr*, d.h. denjenigen, der für eine positive Antwort steht.
- ▶ Klicken Sie zum Anzeigen der Ergebnisse auf Ausführen.

Verteilungsdiagramme (alle Ziele)

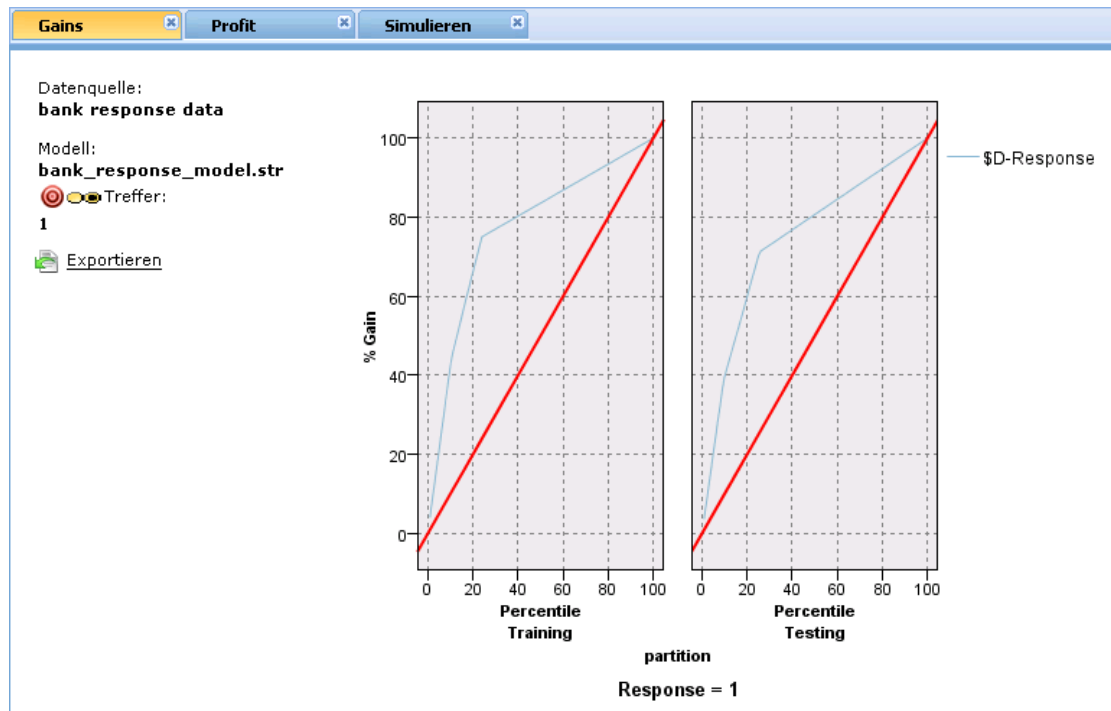
Verteilungsdiagramme stellen die beobachteten und erwarteten Werte dar und werden bei Fertigstellung des Modells standardmäßig angezeigt. Über das Dialogfeld „Auswerten“ erhalten Sie Verteilungsdiagramme für andere Daten-Sets als das Test-Daten-Set. Für Bereichsziele ist das Verteilungsdiagramm der einzige verfügbare Diagrammtyp.

Für Flag- und Set-Ziele können Sie außerdem die folgenden Diagramme auswählen:

Gewinndiagramme

Gains-Diagramme werden verwendet, um den durch den Einsatz des Modells erhaltenen „Gain“ oder „Lift“ darzustellen. Der „Gain“ ist als der Anteil der in jedem Inkrement enthaltenen Treffer im Verhältnis zur Gesamtzahl der vorhandenen Treffer definiert. Die gerade diagonale Linie bildet die ohne Modell für die gesamte Stichprobe erwarteten Antworten ab. In diesem Fall ist die Antwortquote konstant, da die Wahrscheinlichkeit einer Antwort für alle Personen gleich hoch ist. Um die Ausbeute zu verdoppeln, müssten Sie doppelt so viele Personen ansprechen. Die gekrümmte Linie zeigt an, in welchem Maß Sie dieses Ergebnis verbessern können, indem Sie diejenigen mit der höchsten Antwortwahrscheinlichkeit ansprechen. Je steiler die Kurve, desto höher ist der Gain.

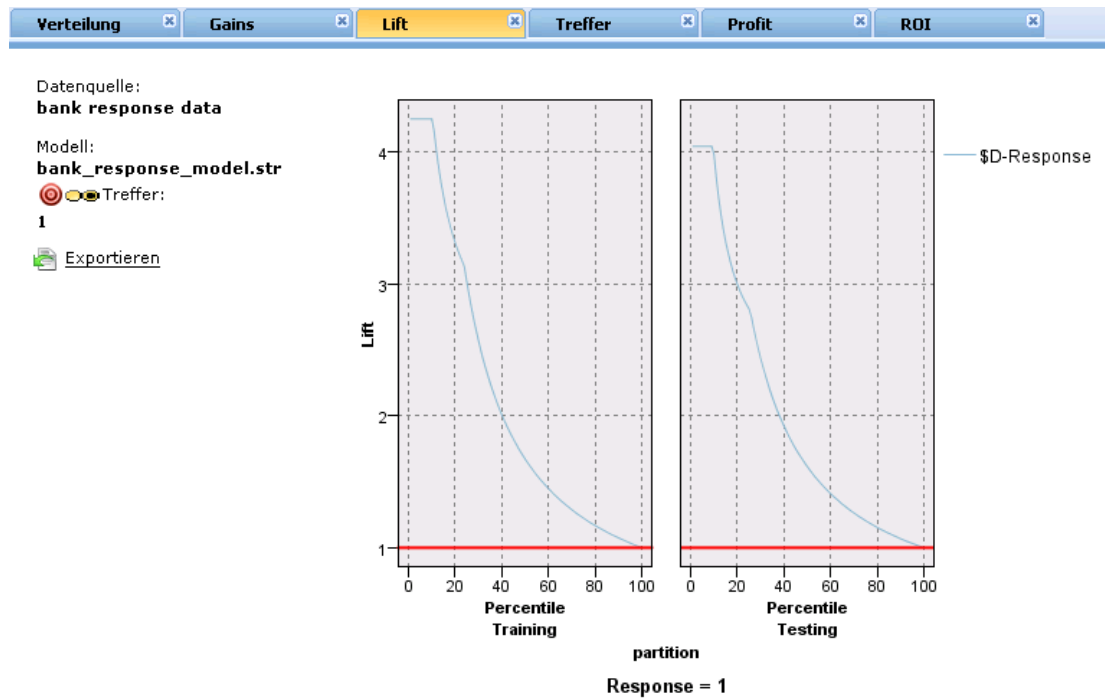
Abbildung 6-2
Gewinndiagramm



Lift-Diagramme

Das Lift-Diagramm vergleicht den Prozentsatz der in jedem Inkrement enthaltenen Datensätze, bei denen es sich um Treffer handelt, mit dem Prozentsatz aller Treffer im Trainings-Daten-Set.

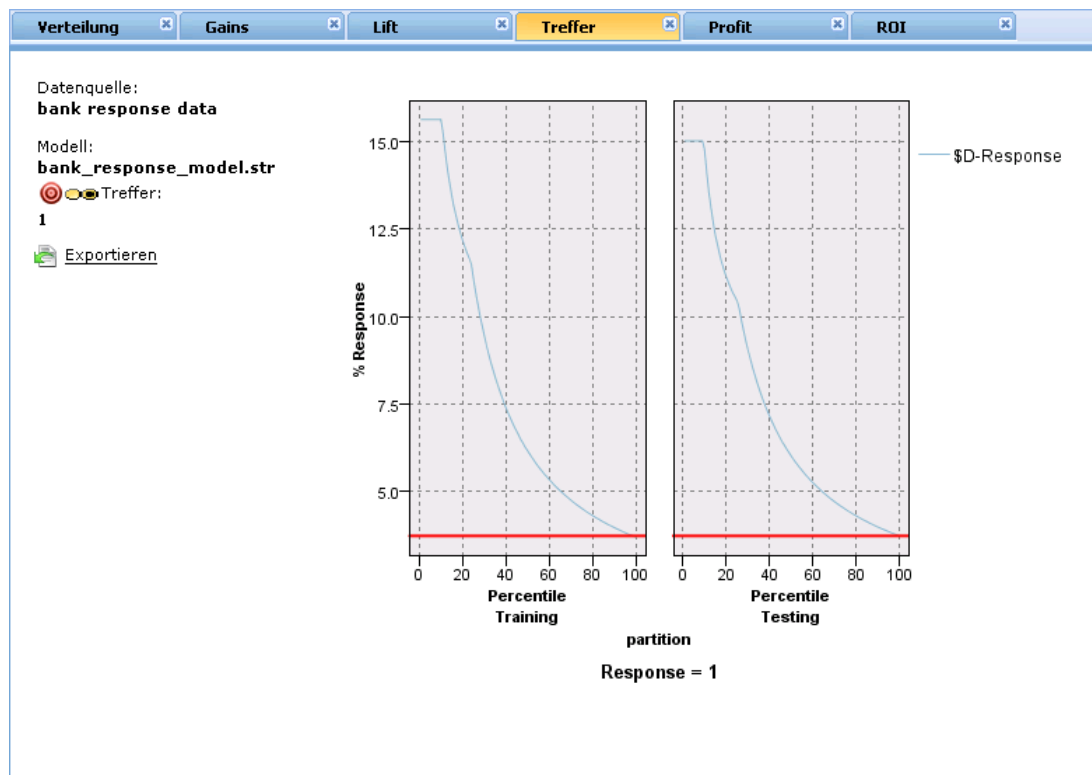
Abbildung 6-3
Lift Chart (Index)



Trefferdiagramme

Das Trefferdiagramm zeigt den Prozentsatz der Datensätze im Inkrement an, bei denen es sich um Treffer handelt.

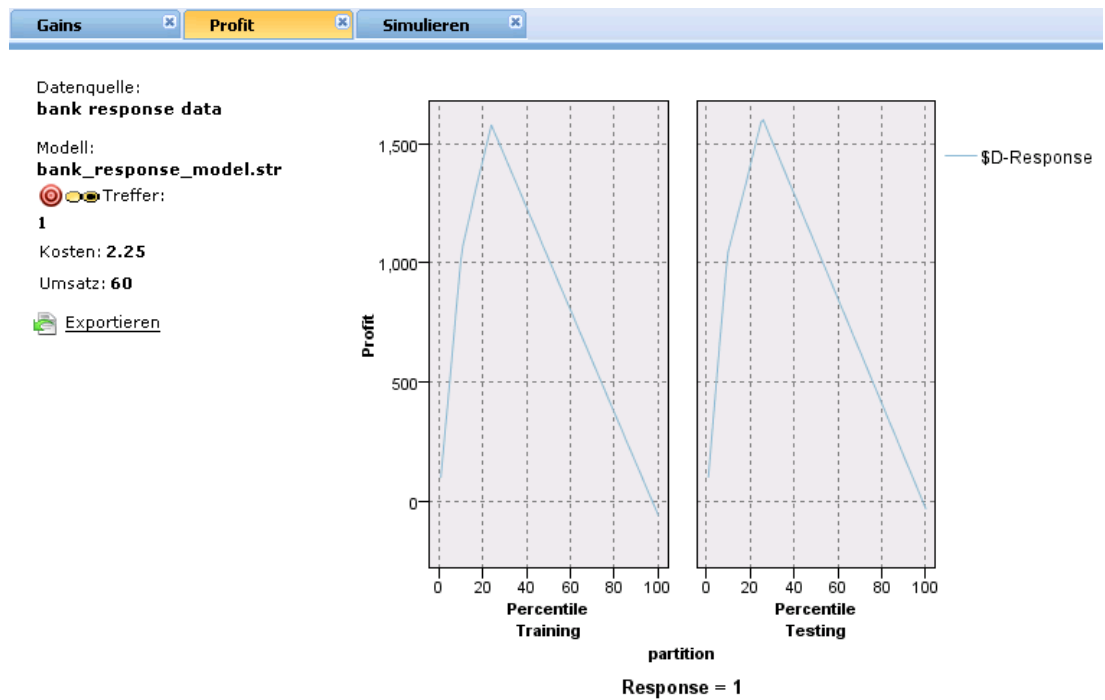
Abbildung 6-4
Trefferdiagramm



Profitdiagramme

Der Profit entspricht dem Umsatz für jeden Datensatz abzüglich der Kosten für den betreffenden Datensatz. Die Profite für ein Quantil entsprechen einfach der Summe der Profite für alle Datensätze im Quantil. Umsätze gelten definitionsgemäß nur für Treffer, Kosten dagegen für alle Datensätze. [Für weitere Informationen siehe Thema Maximalen Profit simulieren auf S. 69.](#)

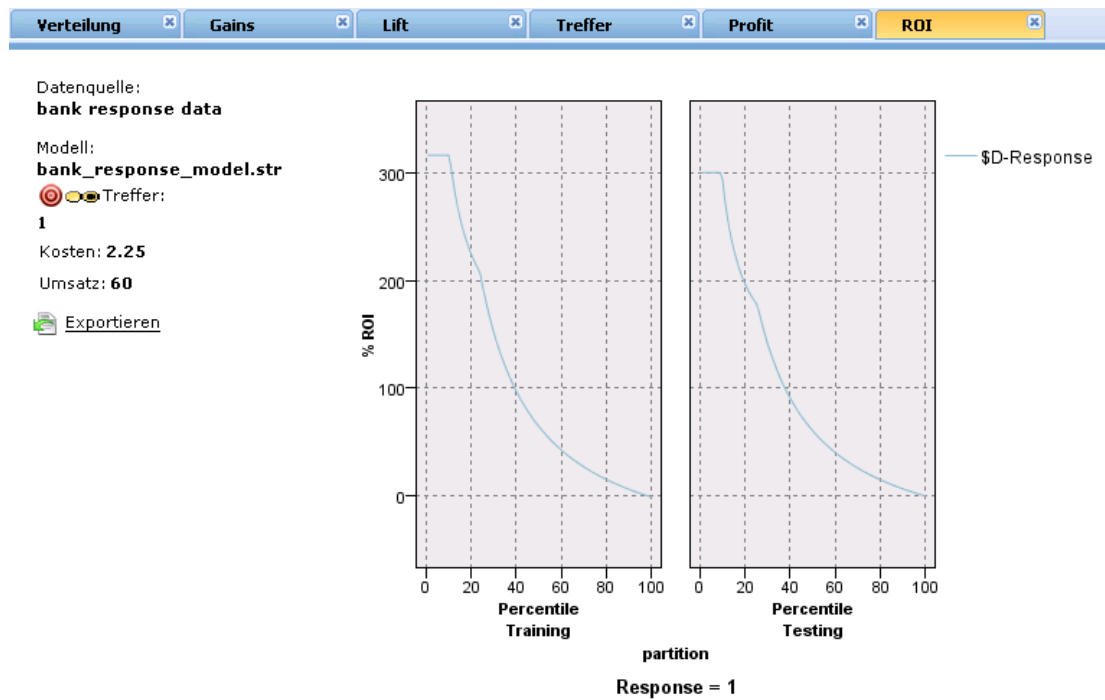
Abbildung 6-5
 Profitdiagramm



ROI-Diagramme

Der ROI-Wert („Return on Investment“ oder Kapitalrendite) weist gewisse Ähnlichkeiten mit dem Profit auf; auch hier basiert die Definition auf den Umsätzen und Kosten. Beim ROI werden die Profite mit den Kosten für jedes Intervall verglichen.

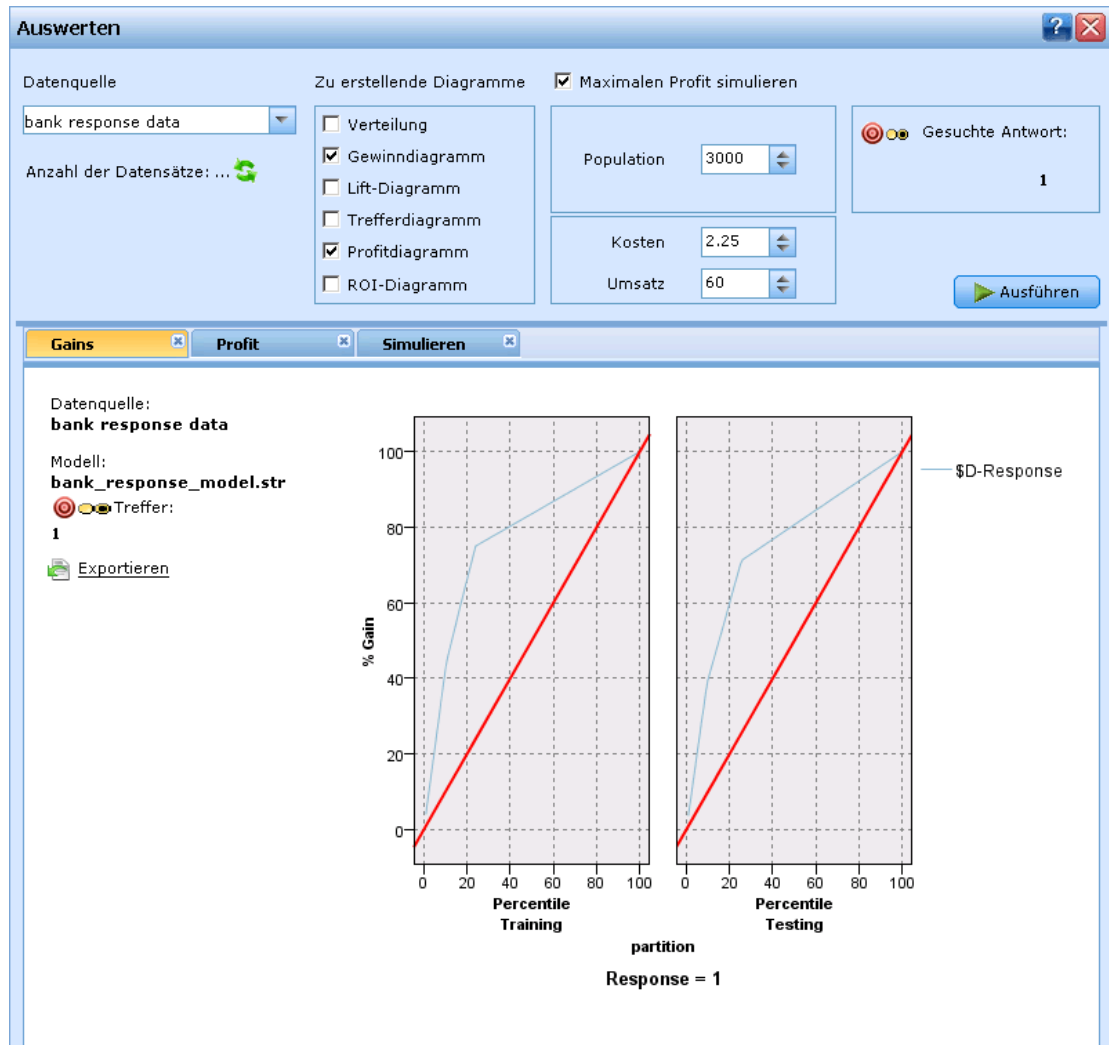
Abbildung 6-6
ROI-Diagramm



Maximalen Profit simulieren

Auf Grundlage der vom Modell vorhergesagten Antwortquote können Sie den maximalen Profit simulieren, der für eine gegebene Population möglich ist. Dazu müssen Sie die Kosten- und Umsatzwerte festlegen sowie die Größe der Population.

Abbildung 6-7
Simulationseinstellungen



So sagen Sie den maximalen Profit vorher:

- ▶ Geben Sie die Population oder Gesamtzahl der Datensätze an, auf die sich die Vorhersage stützen soll. *Anmerkung:* Wenn Sie ein Modell scoden, wird automatisch die Gesamtzahl der Datensätze in der Datenquelle oder Datensatzgruppe eingestellt.
- ▶ Geben Sie die Kosten für die einzelnen Datensätze an. Dies können beispielsweise die Kosten pro Person für Druck und Versand der Werbebroschüre einer Kampagne sein.
- ▶ Geben Sie den erwarteten Umsatz an, den eine positive Antwort generieren würde. Beispiel: Der Einzelhandelspreis eines Artikels einer Kampagne.
- ▶ Wählen Sie die Antwort aus, an der Sie interessiert sind.
- ▶ Klicken Sie auf Ausführen.

Das Diagramm der vorhergesagten Ergebnisse zeigt das beste Ergebnis im Hinblick auf maximalen Profit an sowie den Prozentsatz der Datensätze, die gemäß der vom Modell berechneten Antwortneigung in die Zielgruppe aufgenommen werden müssten, um diesen zu erreichen. [Für weitere Informationen siehe Thema Auswertung des Modells in Kapitel 2 auf S. 10.](#)

Unter dem Diagramm befindet sich ein Schieberegler, mit dem Sie den Prozentsatz der Datensätze, die zur Zielgruppe genommen werden sollen, auf einen Wert zwischen 0 und 100 % einstellen können, um die Auswirkungen auf Ihren Profit zu untersuchen.

Wenn die Ergebniswerte akzeptabel sind, können Sie die betreffenden Einstellungen für das Scoring Ihrer Datensätze übernehmen. Klicken Sie zu diesem Zweck auf [Zum Scoring verwenden](#). Der Prozentsatz der Datensätze, die in die Zielgruppe aufgenommen werden sollen, wird im Bereich „Neigungsoptionen“ der Registerkarte „Scores“ im Feld Top % angezeigt. [Für weitere Informationen siehe Thema Scoring von Modellen in eine Datenbanktabelle, eine Datei oder einen Cognos BI-Server in Kapitel 7 auf S. 75.](#)

Profitvergleich

Um zu überprüfen, ob die gewählten Datensätze ein zufriedenstellendes Ergebnis liefern, können Sie den zu erwartenden Profit aus diesen Datensätzen mit zufällig ausgewählten Datensätzen gleichen Prozentanteils und dem aus allen Datensätzen vergleichen.

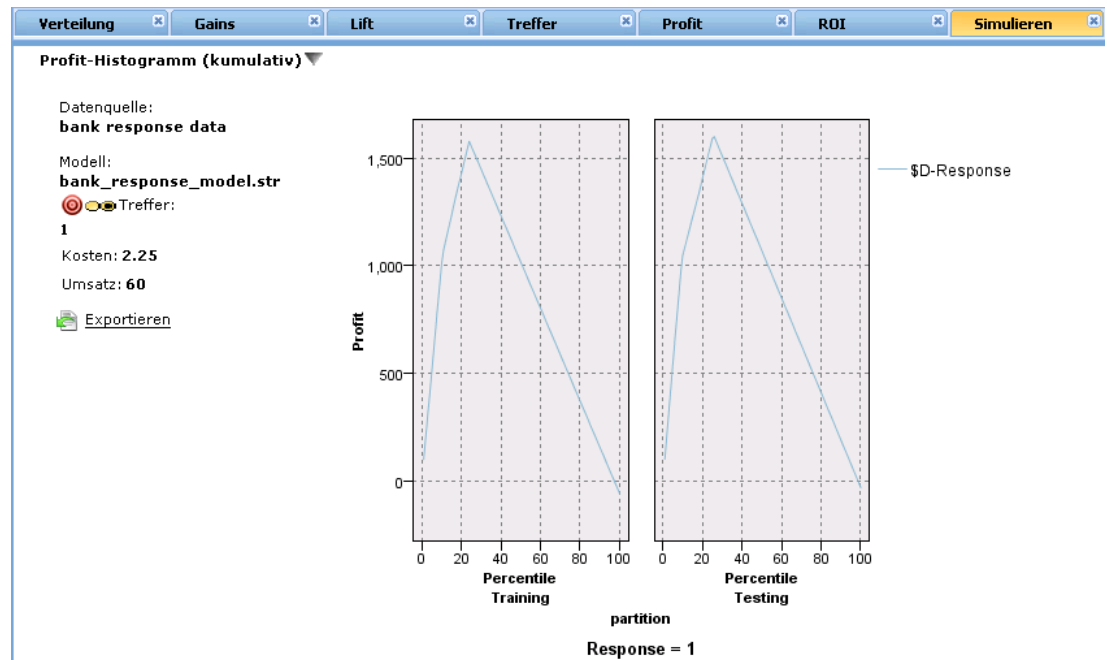
Die angezeigten Details beinhalten die Anzahl der Zieldatensätze, die wahrscheinliche Antwortquote, die Kosten, die Einnahmen und den Gesamtprofit. Sie können diese Daten für die Verwendung in anderen Anwendungen ins HTML-Format exportieren.

Profitdiagramm

Sie können den aus den ausgewählten Datensätzen vorhergesagten kumulierten oder nicht kumulierten Profit als Diagramm anzeigen.

Anmerkung: Dieses Diagramm wird beim ersten Auswertungsdurchlauf erstellt und bei späteren Änderungen mit dem Schieberegler im Profitsimulationsbereich nicht automatisch aktualisiert.

Abbildung 6-8
Profit-Histogramm (kumulativ)



Testen des Modells

Im Dialogfeld „Testen“ können Sie überprüfen, wie ein oder mehrere Testdatensätze gescort würden. Die Ergebnisse hängen von dem Kontext ab, in dem der Test ausgeführt wird, und können die von einem Modell oder einer Regel zurückgegebenen Scores widerspiegeln. Sie können z. B. einige ausgewählte Kundendatensätze betrachten und ablesen, welche Kunden mit hoher Wahrscheinlichkeit auf einen Werbebrief reagieren würden.

- ▶ Klicken Sie auf der Registerkarte „Modellierung“ auf Testen. Das Dialogfeld „Testen“ wird angezeigt.
- ▶ Wählen Sie die Datenquelle, die Sie benutzen möchten, oder stattdessen im Abschnitt „Datensätze“ die Option Benutzerdefinierte Daten auswählen, wenn Sie die Testdatensätze manuell eingeben wollen.
- ▶ Legen Sie ggf. fest, welche Selektionen verwendet werden sollen. Vielleicht wollen Sie bestimmte Datensätze in den Test einbeziehen oder ausschließen. Sie können je nach Bedarf vorhandene Regeln suchen oder neue Regeln erstellen. [Für weitere Informationen siehe Thema Selektionsregeln definieren in Kapitel 5 auf S. 45.](#)
- ▶ Wählen Sie bei Verwendung einer Datenquelle die Anzahl der Datensätze, die zurückgegeben werden sollen. Sie können die ersten n Datensätze bis maximal 100 zurückgeben lassen, die entweder vom Anfang der Daten oder aus einer Zufallsstichprobe der Datensätze eingelesen werden. Beim Testen eines Assoziationsmodells mithilfe von Transaktionsdaten werden Stichproben der Eingaben auf Transaktionsebene, nicht auf Datensatzebene gezogen.

- ▶ Wenn Sie benutzerdefinierte Daten verwenden, wählen Sie Benutzerdefinierte Daten auswählen und klicken auf Benutzerdefinierten Datensatz hinzufügen. Es werden alle im Projektdatenmodell definierten Felder aufgeführt, und Sie können für jedes einen Wert festlegen. Klicken Sie auf Speichern, wenn Sie fertig sind.
- ▶ Wählen Sie unter Anzuzeigende Felder auswählen die Felder, die in der Testausgabe erscheinen sollen.
- ▶ Klicken Sie auf Ausführen.
- ▶ In der Testzusammenfassung wird eine Liste mit erfolgreichen Zuordnungen für die einzelnen Datensätze angezeigt.
- ▶ Über das Informations-Symbol am Anfang einer Zeile erhalten Sie Details zur jeweiligen Zeile.

Verwenden von Modell-Scores in Anwendungen

Beim Scoring eines Modells wird das Modell auf bestimmte Daten oder eine bestimmte Grundgesamtheit angewendet, um ein Ergebnis oder eine Vorhersage zu erhalten, die als Eingabe für Entscheidungen verwendet werden kann. Sie können beispielsweise alle Kunden in Ihrer Datenbank scoren, um mögliche Interessenten für eine Direktwerbungskampagne zu ermitteln, oder die Kunden einzeln bei der Kontaktaufnahme mit dem Call Center scoren, um zu entscheiden, welche Produkte oder Dienstleistungen ihnen angeboten werden sollten. Je nach Anwendung können Scoring-Ergebnisse in eine Datenbanktabelle oder eine einfache Textdatei geschrieben oder als Eingaben für die Segment-, Auswahl und Zuordnungsregeln verwendet werden, die die Entscheidungen in einer Anwendung steuern.

Modell-Scores können auf folgende Weisen verwendet werden:

- Als Eingabe für Regeln. [Für weitere Informationen siehe Thema Regeln erstellen in Kapitel 5 auf S. 43.](#)
- Für Anwendungen, die Batch-Scoring unterstützen, können Scores in eine Datenbanktabelle oder eine einfache Textdatei geschrieben werden. [Für weitere Informationen siehe Thema Scoring von Modellen in eine Datenbanktabelle, eine Datei oder einen Cognos BI-Server in Kapitel 7 auf S. 75.](#)

Modellausgabefelder

Gemäß der Konvention beruhen die Namen der Scoring-Ausgabefelder auf dem Namen des Zielfelds, tragen jedoch zusätzlich ein benutzerdefiniertes Präfix. Wenn der Name des Zielfelds beispielsweise *Antwort* lautet, tragen die Ausgabefelder Namen wie *\$XF-Antwort* und *\$XFRP-Antwort*. Die jeweiligen Felder hängen, wie unten ausgeführt, vom Modelltyp sowie dem Messniveau für das ausgewählte Zielfeld ab. [Für weitere Informationen siehe Thema Messniveaus in Kapitel 4 auf S. 31.](#)

Flag-Ziele. Automatische Modelle mit Flag-Zielen geben die folgenden Felder aus:

- **\$XF-<Ziel>.** Der vorhergesagte Wert für die einzelnen Datensätze, angegeben als für das Feld definierte Wahr-/Falsch-Werte („true“ bzw. „false“).
- **\$XFRP-<Ziel>.** Neigungs-Score; gibt die Wahrscheinlichkeit des „wahren“ Werts für jeden Datensatz an. Neigungs-Scores unterscheiden sich von Konfidenz-Scores, die typischerweise als Vorhersage für den betreffenden Datensatz angegeben werden. So entspricht beispielsweise eine Vorhersage mit dem Wert „false“ (falsch), die eine hohe Konfidenz aufweist, einem niedrigen Neigung-Score, da dieser eine hohe Wahrscheinlichkeit dafür angibt, dass *keine* Antwort erfolgt.

Kategoriale Ziele. Automatische Modelle mit kategorialen Zielen (nominal oder ordinal) geben die folgenden Felder aus:

- **\$XS-<Ziel>.** Vorhergesagter Wert für jeden Datensatz.
- **\$XSC-<Ziel>.** Der Prognose zugeordneter Konfidenzwert.

Stetige Ziele. Automatische Modelle mit stetigen numerischen Zielen geben die folgenden Felder aus:

- **\$XR-<Ziel>.** Vorhergesagter Wert für jeden Datensatz.
- **\$XRE-<Ziel>.** Standardfehler der Prognose.

Vorhersageregelmole (interaktiv) Vorhersageregelmole (nur Flag-Ziele oder kategoriale Ziele) geben die folgenden Felder aus:

- **\$D-<Ziel>.** Vorhergesagtes Ergebnis für jeden Datensatz, wobei der Wert „true“ (wahr) für das Feld definiert ist, das für die gewünschte Antwort verwendet wird, und NULL jeden anderen Wert angibt.
- **\$DP-<Ziel>.** Wahrscheinlichkeit des gewünschten Ergebnisses für jeden Datensatz; entspricht der Wahrscheinlichkeit, die für das Segment beim Erstellen des Modells in der Ergebnistabelle angezeigt wird. [Für weitere Informationen siehe Thema Erstellen eines Vorhersageregelmole \(interaktiv\) auf S. 53.](#)
- **\$DI-<Ziel>.** Indexnummer, anhand der das Segment identifiziert wird, in das der Datensatz fällt (wird in den Modellierungsergebnissen in der Spalte ganz links angezeigt).
- **\$DRP-<Ziel>.** Neigungs-Score; gibt die Wahrscheinlichkeit der gewünschten Antwort an. Nur für Flag-Ziele verfügbar.

Scoren von Modellen

Scoring von Modellen in eine Datenbanktabelle, eine Datei oder einen Cognos BI-Server

Ein Modell zu scoren heißt, es auf die Stichprobe oder Population anzuwenden, für die Sie sich interessieren. Zum Beispiel könnten Sie Ihre Bestandskunden zur Bildung eines Modells verwenden, das vorhersagt, welche Kunden am wahrscheinlichsten auf eine Kampagne reagieren, und das Modell dann dazu einsetzen, Datensätze von Personen zu scoren, die derzeit nicht zu Ihren Kunden gehören, aber für die Zusendung eines Werbebriefs in Frage kommen.

Folgende grundlegende Arbeitsschritte müssen zum Scoren eines Modells durchgeführt werden:

- ▶ Wählen Sie das Ziel für die gescorten Daten aus.
- ▶ Wählen Sie, welche Informationen ausgegeben und wie sie gespeichert werden sollen.
- ▶ Wählen Sie die Datenquelle sowie alle Datensatzgruppen aus, die Sie scoren möchten.
- ▶ Legen Sie fest, welche Datensätze gescort werden sollen, und zeigen Sie eine Vorschau der erwarteten Ergebnisse an.
- ▶ Klicken Sie auf Jetzt scoren, um das Modell gemäß Ihren Einstellungen zu scoren.

Scores werden in ein oder mehrere Felder geschrieben, die den Daten hinzugefügt werden. Gemäß der Konvention beruhen die Namen dieser Felder auf dem Zielfeld, tragen jedoch zusätzlich ein Präfix. Wenn der Name des Zielfelds beispielsweise *Antwort* lautet, tragen die Ausgabefelder Namen wie *\$XF-Antwort* und *\$XFRP-Antwort*. [Für weitere Informationen siehe Thema Ausgabefelder auswählen auf S. 79.](#)

Scoring-Ziel auswählen

Sie können die Scoring-Daten wahlweise in eine Datenbank, in eine Datei oder an einen Cognos BI-Server ausgeben. Zu den unterstützten Dateitypen gehören das Textformat (*.txt, *.csv), das Excel-Format, das von IBM® SPSS® Statistics benutzte *.sav-Format, das SAS-Format sowie die von IBM® SPSS® Data Collection unterstützten Formate. [Für weitere Informationen siehe Thema Dateiziel auf S. 76.](#)

Datenbankziel

Wenn Sie angeben, dass die Scoring-Ausgabe an eine Datenbank gesendet werden soll, werden Sie aufgefordert, weitere Details einzugeben.

- ▶ Geben Sie in das Feld Datenbankname den Datenbanknamen ein, oder wählen Sie die gewünschte Datenbank aus.

- Geben Sie an, wie die Scores gespeichert werden sollen. Sie haben die Möglichkeit, eine neue Tabelle zu erstellen. In diesem Fall werden Sie aufgefordert, einen Namen für die Tabelle einzugeben. Alternativ können Sie die Daten entweder an eine vorhandene Tabelle anhängen oder eine vorhandene Datenquelle überschreiben. In diesem Fall werden Sie aufgefordert, die gewünschte Tabelle oder Datenquelle auszuwählen.

Wenn Sie die Daten an eine vorhandene Datenbank anhängen oder diese überschreiben, können Sie die Scoring-Felder auf Felder des vorhandenen Ziels abbilden. [Für weitere Informationen siehe Thema Scoring-Felder abbilden auf S. 79.](#)

Dateiziel

Scores können in einer Reihe verschiedener Dateitypen ausgegeben werden. Wählen Sie für den jeweiligen Dateityp eine vorhandene Datei aus bzw. erstellen Sie eine neue Datei, der die Scoring-Daten hinzugefügt werden sollen.

Wenn Sie die Daten an eine vorhandene Datei anhängen oder diese überschreiben, können Sie die Scoring-Felder auf Felder des vorhandenen Ziels abbilden. [Für weitere Informationen siehe Thema Scoring-Felder abbilden auf S. 79.](#)

Textdatei

Wenn Sie angeben, dass die Scoring-Ausgabe an eine Textdatei gesendet werden soll, können Sie weitere Details eingeben, um den Dateiinhalt festzulegen.

Schreibmodus. Wenn die Option Überschreiben aktiviert ist, werden alle vorhandenen Daten in der angegebenen Datei überschrieben. Ist die Option Anhängen aktiviert, wird die Ausgabe an die vorhandene Datei angehängt; die bereits vorhandenen Daten in dieser Datei werden also beibehalten.

- **Feldnamen einschließen.** Bei dieser Option werden die Dateinamen in die erste Zeile der Ausgabedatei geschrieben. Diese Option ist nur für den Schreibmodus Überschreiben verfügbar.

Neue Zeile nach jedem Datensatz. Bei dieser Option wird jeder Datensatz in eine eigene Zeile in der Ausgabedatei geschrieben.

Feldtrennzeichen. Dient zur Angabe des Zeichens, das als Trennzeichen zwischen den Feldwerten in der erzeugten Textdatei eingefügt werden soll. Die folgenden Optionen stehen zur Auswahl: Komma, Tabulator, Leerzeichen und Andere. Wenn Sie die Option Andere wählen, geben Sie das oder die gewünschten Trennzeichen in das Textfeld ein.

Symbolanführungszeichen. Hier können Sie die Art der Anführungszeichen angeben, die für Werte in symbolischen Feldern verwendet werden sollen. Die folgenden Optionen stehen zur Auswahl: Keine (die Werte werden nicht in Anführungszeichen eingeschlossen), Einfach ('), Doppelt (") und Andere. Wenn Sie die Option Andere wählen, geben Sie das oder die gewünschten Anführungszeichen in das Textfeld ein.

Kodierung. Gibt die verwendete Textkodierungsmethode an. Sie haben die Wahl zwischen der System-StandardEinstellung und UTF-8.

Dezimaltrennzeichen. Hier können Sie das Trennzeichen für die Dezimalstellen in den Daten festlegen.

- **Punkt (.).** Als Dezimaltrennzeichen wird ein Punkt verwendet.
- **Komma (,).** Als Dezimaltrennzeichen wird ein Komma verwendet.

Excel-Dateien

Typ. Wählen Sie den Excel-Dateityp aus, den Sie exportieren.

Feldnamen einschließen. Legt fest, ob Feldnamen in die erste Zeile des Arbeitsblattes eingefügt werden sollen.

SAS-Dateien

Typ. Wählen Sie das zu erstellende SAS-Dateiformat aus. Es stehen drei SAS-Dateiformate zur Auswahl: SAS für Windows/OS2 (*.sd2), SAS für UNIX (*.ssd) und SAS Version 7/8/9 (*.sas7bdat).

Feldnamen exportieren. Wählen Sie die Optionen zum Exportieren der Feldnamen und Beschriftungen aus, die in SAS genutzt werden sollen.

- **Namen und Variablenbeschriftungen.** Wählen Sie diese Option, um sowohl Feldnamen als auch Feldbeschriftungen zu exportieren. Die Namen werden als SAS-Variablenamen exportiert, die Beschriftungen entsprechend als SAS-Variablenbeschriftungen.
- **Namen als Variablenbeschriftungen.** Wählen Sie diese Option, um die Feldnamen in SAS als Variablenbeschriftungen zu verwenden. Um die mögliche Bildung ungültiger SAS-Namen zu vermeiden, wählen Sie stattdessen die Option Namen und Variablenbeschriftungen.

IBM SPSS StatisticsDateien

Feldnamen exportieren. Dient zur Angabe einer Methode für die Behandlung von Variablenamen und Beschriftungen nach dem Export als IBM® SPSS® Statistics S_{AV} -Datei. Beachten Sie, dass ungültige Namen automatisch durch Ersetzung der ungültigen Zeichen korrigiert werden.

- **Namen und Variablenbeschriftungen.** Wählen Sie diese Option, um sowohl Feldnamen als auch Feldbeschriftungen zu exportieren. Die Namen werden als SPSS Statistics-Variablenamen exportiert, die Beschriftungen entsprechend als SPSS Statistics-Variablenbeschriftungen.
- **Namen als Beschriftungen.** Wählen Sie diese Option, um die Feldnamen in SPSS Statistics als Variablenbeschriftungen zu verwenden. Um die mögliche Bildung ungültiger SPSS Statistics-Namen zu vermeiden, wählen Sie stattdessen die Option Namen und Variablenbeschriftungen.

Cognos BI Server als Ziel

Wenn Sie angeben, dass die Scoring-Ausgabe an einen Cognos BI-Server gesendet werden soll, werden Sie aufgefordert, weitere Details einzugeben.

Anmerkung: Es können nur relationale Daten exportiert werden, keine OLAP-Daten.

Um Daten nach IBM Cognos BI zu exportieren, müssen Sie folgende Angaben machen:

- **Cognos-Verbindung.** Die Verbindung zum IBM Cognos BI-Server.
- **ODBC-Verbindung.** Die vom IBM Cognos BI-Server verwendete Verbindung zum Cognos-Datenserver.

Die Verbindungen müssen auf dieselbe Datenbank verweisen und der Benutzername und das Passwort für den Cognos-Server, der für die Verbindung mit ODBC verwendet wird, müssen den ODBC-Details entsprechen.

Die eigentlichen Daten werden an den Datenserver exportiert und die Paketmetadaten an den Cognos BI-Server.

- ▶ Wenn Sie als Ziel den Cognos BI-Server auswählen und auf Verbinden klicken, werden Sie aufgefordert, die Details für die Serververbindung einzugeben.
- ▶ Geben Sie die Server-URL des IBM Cognos-Servers ein, den Sie für den Import bzw. Export von Daten verwenden möchten. Wenden Sie sich an Ihren Cognos-Systemadministrator, wenn Sie sich nicht sicher sind, welche URL Sie verwenden müssen.
- ▶ Wählen Sie den Modus für die Verbindung aus. Wenn Sie sich als ein bestimmter Benutzer (z. B. als Administrator) anmelden möchten, wählen Sie die Option Anmeldeinformationen festlegen aus und geben Sie den Namespace, den Benutzernamen und das Passwort für IBM Cognos ein:
 - Geben Sie den Namespace des IBM Cognos-Sicherheitsauthentifizierungsanbieters ein, über den die Anmeldung beim Server erfolgt. Der Authentifizierungsanbieter dient dazu, Benutzer, Gruppen und Rollen zu definieren und zu verwalten und den Authentifizierungsprozess zu steuern.
 - Geben Sie den Benutzernamen für IBM Cognos ein, der für die Anmeldung beim Server verwendet wird.
 - Geben Sie das Passwort (Passwort) ein, das zum angegebenen Benutzernamen gehört.
- ▶ Um sich stattdessen ohne Anmeldeinformationen anzumelden (bei dieser Variante können die Felder für Namespace, Benutzernamen und Passwort nicht ausgefüllt werden), verwenden Sie die Option Anonyme Verbindung verwenden.

Anmerkung: Bei einigen Serververbindungen sind anonyme Verbindungen nicht zulässig.

- ▶ Wählen Sie die Datenquelle aus.
- ▶ Geben Sie den Pfad und den Namen für den Ordner auf dem Cognos BI-Server ein, in dem das Exportpaket erstellt werden soll. Alternativ können Sie den erforderlichen Ordner auswählen.
- ▶ Geben Sie den Paketnamen des Pakets ein, das die exportierten Metadaten enthalten soll. Es muss sich dabei um ein neues Paket handeln. Ein Export in ein bestehendes Paket ist nicht möglich. [Für weitere Informationen siehe Thema Auswählen von Cognos-Objektdetails in Kapitel 4 auf S. 38.](#)
- ▶ Geben Sie zur Auswahl des ODBC-Ziels den Datenbanknamen in das Feld Datenbankname ein oder wählen Sie die gewünschte Datenbank aus.
- ▶ Geben Sie an, wie die Scores gespeichert werden sollen. Sie haben die Möglichkeit, eine neue Tabelle zu erstellen. In diesem Fall werden Sie aufgefordert, einen Namen für die Tabelle

einzugeben. Alternativ können Sie die Daten entweder an eine vorhandene Tabelle anhängen oder eine vorhandene Datenquelle überschreiben. In diesem Fall werden Sie aufgefordert, die gewünschte Tabelle oder Datenquelle auszuwählen.

Wenn Sie die Daten an eine vorhandene Datenbank anhängen oder diese überschreiben, können Sie die Scoring-Felder auf Felder des vorhandenen Ziels abbilden. [Für weitere Informationen siehe Thema Scoring-Felder abbilden auf S. 79.](#)

Scoring-Felder abbilden

Wenn Sie Scores an eine vorhandene Datenbanktabelle oder Datei anhängen oder diese überschreiben, müssen Sie die Scoring-Felder u. U. auf Felder in der vorhandenen Tabelle oder Datei abbilden. Beachten Sie, dass die Ausgabefelder und die entsprechenden abgebildeten Felder denselben Datentyp aufweisen müssen.

- Beim Schreiben von Scores in eine vorhandene Datenbanktabelle müssen alle Felder auf vorhandene Felder abgebildet werden, damit das Scoring fortgesetzt werden kann, unabhängig davon, ob die neuen Scores angehängt oder überschrieben werden. Klicken Sie bei Bedarf auf Felder abbilden, um Ausgaben für alle verbleibenden Felder anzugeben.
 - Beim Schreiben in eine vorhandene Datei ist ein Abbilden nur dann erforderlich, wenn in den Optionen für die Ausgabedatei die Option Anhängen ausgewählt ist. Beim Überschreiben einer vorhandenen Datei ist ein Abbilden nicht erforderlich, da die neue Datei die alte einfach ersetzt.
- ▶ Klicken Sie auf Felder abbilden. Ein Dialogfeld wird geöffnet, und das System versucht zu ermitteln, welche Abbildung für Sie am besten geeignet ist. Sie können jederzeit zu der vom System vorgeschlagenen Abbildung zurückkehren, indem Sie auf Zurücksetzen klicken.
 - ▶ Wählen Sie aus den verfügbaren Ausgabezielen und klicken Sie auf Abbilden, um sie auf die entsprechenden Scoring-Felder abzubilden.
 - ▶ Wenn Sie fertig sind, klicken Sie auf Speichern.

Ausgabefelder auswählen

Wählen Sie auf der Registerkarte „Scoren“ die Option Ausgabefelder angeben, um die Felder auszuwählen, die in der Scoring-Ausgabe enthalten sein sollen. Neben den tatsächlichen Scores sollten Sie normalerweise mindestens ein Feld einschließen, mit dem Sie jeden Datensatz identifizieren können, wie z. B. das Feld „Kunden-Nr.“. Optional können einige oder alle der bei der Modellierung verwendeten Eingabefelder eingeschlossen werden.

Modellausgabefelder

Gemäß der Konvention beruhen die Namen der Scoring-Ausgabefelder auf dem Namen des Zielfelds, tragen jedoch zusätzlich ein benutzerdefiniertes Präfix. Wenn der Name des Zielfelds beispielsweise *Antwort* lautet, tragen die Ausgabefelder Namen wie *\$XF-Antwort* und *\$XFRP-Antwort*. Die jeweiligen Felder hängen, wie unten ausgeführt, vom Modelltyp sowie

dem Messniveau für das ausgewählte Zielfeld ab. [Für weitere Informationen siehe Thema Messniveaus in Kapitel 4 auf S. 31.](#)

Flag-Ziele. Automatische Modelle mit Flag-Zielen geben die folgenden Felder aus:

- **\$XF-<Ziel>.** Der vorhergesagte Wert für die einzelnen Datensätze, angegeben als für das Feld definierte Wahr-/Falsch-Werte („true“ bzw. „false“).
- **\$XFRP-<Ziel>.** Neigungs-Score; gibt die Wahrscheinlichkeit des „wahren“ Werts für jeden Datensatz an. Neigungs-Scores unterscheiden sich von Konfidenz-Scores, die typischerweise als Vorhersage für den betreffenden Datensatz angegeben werden. So entspricht beispielsweise eine Vorhersage mit dem Wert „false“ (falsch), die eine hohe Konfidenz aufweist, einem niedrigen Neigung-Score, da dieser eine hohe Wahrscheinlichkeit dafür angibt, dass *keine* Antwort erfolgt.

Kategoriale Ziele. Automatische Modelle mit kategorialen Zielen (nominal oder ordinal) geben die folgenden Felder aus:

- **\$XS-<Ziel>.** Vorhergesagter Wert für jeden Datensatz.
- **\$XSC-<Ziel>.** Der Prognose zugeordneter Konfidenzwert.

Stetige Ziele. Automatische Modelle mit stetigen numerischen Zielen geben die folgenden Felder aus:

- **\$XR-<Ziel>.** Vorhergesagter Wert für jeden Datensatz.
- **\$XRE-<Ziel>.** Standardfehler der Prognose.

Vorhersageregelmodelle (interaktiv) Vorhersageregelmodelle (nur Flag-Ziele oder kategoriale Ziele) geben die folgenden Felder aus:

- **\$D-<Ziel>.** Vorhergesagtes Ergebnis für jeden Datensatz, wobei der Wert „true“ (wahr) für das Feld definiert ist, das für die gewünschte Antwort verwendet wird, und NULL jeden anderen Wert angibt.
- **\$DP-<Ziel>.** Wahrscheinlichkeit des gewünschten Ergebnisses für jeden Datensatz; entspricht der Wahrscheinlichkeit, die für das Segment beim Erstellen des Modells in der Ergebnistabelle angezeigt wird. [Für weitere Informationen siehe Thema Erstellen eines Vorhersageregelmodells \(interaktiv\) in Kapitel 6 auf S. 53.](#)
- **\$DI-<Ziel>.** Indexnummer, anhand der das Segment identifiziert wird, in das der Datensatz fällt (wird in den Modellierungsergebnissen in der Spalte ganz links angezeigt).
- **\$DRP-<Ziel>.** Neigungs-Score; gibt die Wahrscheinlichkeit der gewünschten Antwort an. Nur für Flag-Ziele verfügbar.

Daten und Datensatzgruppen auswählen

Wählen Sie in der Liste Zu scorende Datenquelle die gewünschte Datenquelle aus.

Datensatzgruppen

Wenn Sie nicht die gesamte Datenquelle scoren möchten, können Sie eine Datensatzgruppe auswählen, indem Sie eine vorhandene Regel anwenden oder speziell für dieses Modell eine neue Regel erstellen. [Für weitere Informationen siehe Thema Selektionsregeln definieren in Kapitel 5 auf S. 45.](#)

Datensätze für Scoring auswählen

Bei Modellen, die Neigungsscores zurückgeben, können Sie diese Werte zur Auswahl der zu scorenden Datensätze verwenden. Sie können beispielsweise die 30 % der Kunden mit der höchsten Neigung zu reagieren scoren und anhand dessen eine Mailingliste erstellen. Neigungsscores geben die Wahrscheinlichkeit des „wahren“ Werts für das Ziel an und sind für Vorhersagemodelle mit Flag-Zielen sowie für Vorhersageregelmole (interaktiv) verfügbar. (Bei Vorhersageregelmole (interaktiv) ist die ausgewählte Antwort der „wahre“ Wert, und alle anderen sind falsch.) Bei anderen Modelltypen sind diese Optionen nicht verfügbar.)

Es gibt drei Möglichkeiten, Datensätze nach Neigung auszuwählen:

- **Den Prozentsatz mit den höchsten Neigungswerten.** Standardmäßig werden die höchsten zehn Prozent ausgewählt. Mit dieser Option können Sie Ihre Scoring-Ergebnisse nach Profit simulieren. [Für weitere Informationen siehe Thema Maximalen Profit simulieren in Kapitel 6 auf S. 69.](#)
- **Eine vorgegebene Anzahl der Datensätze mit den höchsten Neigungswerten.** Standardmäßig werden die höchsten 1.000 Datensätze ausgewählt (soweit verfügbar).
- **Durch Angabe des minimalen und maximalen Neigungswerts.** Zum Beispiel alle Datensätze mit einer Antwortwahrscheinlichkeit zwischen 75 % und 100 %.

Scoring von Modellen

Wenn Sie alle Einstellungen zum Scoren festgelegt haben, können Sie eine Vorschau der Ergebnisse für das Modell anzeigen. Dies kann sich besonders dann als nützlich erweisen, wenn Ihr Daten-Set sehr groß ist, da Sie so schnell überprüfen können, ob die Ergebnisse wie gewünscht angezeigt werden, ohne auf die Auswertung des gesamten Daten-Sets warten zu müssen.

Anmerkung: Falls die Daten ungültig oder unvollständig sind, wird bei der Anforderung der Vorschau eine Warnmeldung ausgegeben.

Bei der Datenvorschau handelt es sich um eine Tabelle, in der Datensätze reihenweise und Variablen spaltenweise angezeigt werden.

Scoren

Um zu ermitteln, wie viele Datensätze gescort werden, klicken Sie auf das Aktualisieren-Symbol im unteren Teil der Ansicht. Es wird dann die Anzahl der zur Modellbildung verwendeten Datensätze ausgegeben, wobei die Gesamtzahl der Datensätze der Datenquelle sowie eventuelle von Ihnen festgelegte optionale Einstellungen berücksichtigt werden.

Wenn Sie auf Jetzt scoren klicken, wird eine Fortschrittsanzeige im unteren Teil der Registerkarte angezeigt. Das Scoring kann jederzeit angehalten werden, solange die Fortschrittsanzeige angezeigt wird.

Anmerkung: Falls die Daten ungültig oder unvollständig sind, wird bei der Initiierung des Auswertungsvorgangs eine Warnmeldung ausgegeben.

Wenn die Daten erfolgreich gescort wurden, erhalten Sie eine Mitteilung, wo die Ergebnisse gespeichert wurden – in der von Ihnen angegebenen Datenbank oder Datei.

Anwendungen verwalten

Administratoren können IBM® Analytical Decision Management-Anwendungen für Geschäftsnutzer vorkonfigurieren. Diese Konfigurierbarkeit erleichtert Geschäftsnutzern die Arbeit und erlaubt eine schnellere und effektivere Bedienung der Anwendung, was letztlich zur Lösung betriebswirtschaftlicher Problemstellungen beiträgt. Sie ermöglicht es Administratoren außerdem, den Zugriff auf bestimmte Funktionen einzuschränken.

Verwenden Sie IBM® SPSS® Collaboration and Deployment Services Deployment Manager, um bei Bedarf bestimmten Benutzern oder Gruppen Administratorrechte zu erteilen. Nur Benutzer, die zur *Decision Management-Verwaltung* berechtigt sind, haben Zugriff auf die in diesem Abschnitt beschriebenen Funktionen.

- ▶ Öffnen Sie Deployment Manager.
- ▶ Wählen Sie aus dem Menü „Extras“ die Option Server-Verwaltung.
- ▶ Wählen Sie die Registerkarte „Server-Verwaltung“ und verbinden sich mit Ihrem Server.
- ▶ Öffnen Sie unter Benutzer und Gruppen das Local User Repository.
- ▶ Klicken Sie auf Neue Gruppe, um eine neue Gruppe für Administratoren und eine neue für Geschäftsnutzer zu erstellen. Zum Beispiel:
 - Decision Management-Administratoren.
 - Decision Management-Benutzer.

Falls gewünscht, können Sie zu diesem Zeitpunkt auch neue Benutzer einrichten. Fügen Sie die gewünschten Benutzer bei der Erstellung der jeweiligen Gruppe hinzu. Benutzer der Administratorengruppe beispielsweise sind zur Verwaltung der IBM Analytical Decision Management-Anwendungen in Ihrem Unternehmen berechtigt.
- ▶ Wechseln Sie zu „Rollen“ und klicken auf Neue Rolle, um eine neue administrative Rolle und eine neue Geschäftsnutzerrolle einzurichten. Zum Beispiel:
 - IBM Analytical Decision Management-Administrator.
 - IBM Analytical Decision Management-Benutzer.

Weisen Sie den neuen Rollen zumindest die folgenden Vorgänge zu.

Tabelle 8-1
Rollenaktionen

Rolle	Erforderliche Vorgänge
Decision Management-Administrator	Auf Inhalte und Verzeichnisse zugreifen Decision Management-Verwaltung Bericht dynamisch ausführen Modell scoren Arbeit senden Alle Versionen anzeigen
Decision Management-Benutzer	Auf Inhalte und Verzeichnisse zugreifen Bericht dynamisch ausführen Modell scoren Arbeit senden Neueste anzeigen* <i>ODER</i> Alle Versionen anzeigen*

**Hinweis:* Die Aktion Show latest oder die Aktion Show All Versions ist erforderlich, wenn die Benutzer auf Projektversionen zugreifen müssen, die sie nicht selbst erstellt haben. Andernfalls werden die Projekte im schreibgeschützten Modus geöffnet, wenn der aktuelle Benutzer keinen Zugriff auf die neuste Version hat.

- ▶ Wählen Sie die Rolle „Decision Management-Administrator“ und klicken auf Benutzer und Gruppen bearbeiten. Fügen Sie die soeben erstellte Gruppe „Decision Management-Administratoren“ hinzu.
- ▶ Wählen Sie die Rolle „Decision Management-Benutzer“ und klicken auf Benutzer und Gruppen bearbeiten. Fügen Sie die soeben erstellte Gruppe „Decision Management-Benutzer“ hinzu.

Allgemeine Optionen

Bestimmte allgemeine Optionen stehen Administratoren auf allen Registerkarten der IBM® Analytical Decision Management-Benutzeroberfläche zur Verfügung. Diese Optionen ermöglichen es Administratoren, Registerkarten für Geschäftsnutzer zu sperren oder auszublenden.

- Oben auf jeder Registerkarte befindet sich die Option Registerkarte <RK-Name> für Nicht-Administratoren ausblenden, wobei <RK-Name> der Name der Registerkarte ist. Dadurch wird die gesamte Registerkarte für Geschäftsnutzer ausgeblendet. Wenn Sie z. B. nicht möchten, dass Benutzer die Registerkarte „Daten“ einer Anwendung aufrufen können, melden Sie sich in der Anwendung an, gehen Sie zur Registerkarte „Daten“ und wählen Sie die Option Registerkarte „Daten“ für Nicht-Administratoren ausblenden.
- Oben auf jeder Registerkarte befindet sich die Option Alle <RK-Name>-Optionen sperren, wobei <RK-Name> der Name der Registerkarte ist. Dadurch werden alle Steuerelemente der Registerkarte gesperrt, so dass Geschäftsnutzer nur zum Lesen berechtigt sind. Wenn Sie z. B. Geschäftsnutzern die Einsicht der aktuellen Dateneinstellungen, aber nicht deren Änderung ermöglichen wollen, können Sie diese Option auf der Registerkarte „Daten“ einer Anwendung einstellen. Diese Einstellung erweist sich auch dann als sehr nützlich, wenn Sie bestimmte Optionen als Hilfestellung für Geschäftsnutzer voreinstellen wollen.
- Administratoren können auch einzelne Einstellungen einer Registerkarte durch Klicken auf das Schloss-Symbol sperren, wo verfügbar. Bei Anwendungen mit der Registerkarte „Definieren“ können die Administratoren beispielsweise den Dimensionshierarchiebaum sperren, sodass dieser nicht von Geschäftsnutzern geändert werden kann. Durch Klicken auf

das Schloss-Symbol wird das zugehörige Steuerelement ge- bzw. entsperrt. Geschäftsnutzer sehen ein verriegeltes Schloss-Symbol, wenn der Administrator ein Steuerelement gesperrt hat. Das entriegelte Schloss-Symbol wird jedoch nie eingeblendet.

Einzelne gesperrte Optionen können auch mit grauem Hintergrund oder in anderer Farbe angezeigt werden, um anzuzeigen, dass sie nicht ausgewählt werden können. Beim Versuch, sie auszuwählen, erscheint eine Fehlermeldung. Wenn eine ganze Registerkarte gesperrt wurde, wird ein entsprechender Hinweis oben auf der Registerkarte angezeigt.



Vorkonfiguration von Anwendungs-Verknüpfungen

Der Anwendungs-Startbereich gibt Geschäftsnutzern eine einfache Methode zum Öffnen von Anwendungen und Hinzufügen eigener Verknüpfungen zu bestimmten mit Beschriftung versehenen Versionen einer Anwendung zur Hand. Administratoren können festlegen, welche Verknüpfungen zu Anwendungen im Startbereich aller Benutzer eingeblendet werden. Melden Sie sich mit Administratorrechten an und gehen Sie wie folgt vor:

- ▶ Wählen Sie Standardmäßig im Startbereich für Anwendungen, die im Startbereich aller Benutzer eingeblendet werden sollen.
- ▶ Wenn Sie auf das Schloss-Symbol einer Anwendungs-Verknüpfung klicken, können Geschäftsnutzer diese nicht aus ihrem Startbereich entfernen. Beachten Sie jedoch, dass Geschäftsnutzer immer noch andere Anwendungs-Verknüpfungen hinzufügen und entfernen können. Der Benutzer sieht bei gesperrten Verknüpfungen anstelle des Schließen-Symbols ein verriegeltes Schloss-Symbol. [Für weitere Informationen siehe Thema Anwendungen starten in Kapitel 3 auf S. 20.](#)

Ausblenden der Hauptseite

Jede Anwendung stellt auf der Hauptseite die verschiedenen Schritte des Arbeitsablaufs in Form von grafischen Symbolen dar. Durch Klicken auf eine Grafik wechseln Sie zum entsprechenden Abschnitt der Benutzeroberfläche. Geschäftsnutzer benötigen jedoch nicht unbedingt immer Zugriff auf alle Arbeitsschritte einer Anwendung.

Administratoren können die Hauptseite für Geschäftsnutzer ausblenden. In einer bestimmten Anwendung benutzen Geschäftsnutzer z. B. vielleicht nur die Registerkarte „Berichte“. In diesem Fall könnte der Administrator die Hauptseite und alle anderen Registerkarten für Geschäftsnutzer ausblenden. Wenn sich also ein Geschäftsnutzer anmeldet, erreicht er sofort die Registerkarte „Berichte“ und bekommt die anderen nie zu sehen.

- ▶ Wählen Sie auf der Hauptseite Anwendungs-Hauptseite ausblenden.
- ▶ Legen Sie im Dropdown-Menü fest, welche Registerkarte zuerst eingeblendet werden soll, wenn ein Benutzer die Anwendung öffnet.

Beachten Sie, dass für den Fall, dass Sie die ausgewählte Registerkarte später ausblenden, die nächste nicht ausgeblendete Registerkarte der Anwendung angezeigt würde.

Sperrung der Datenoptionen

Daten

Die Registerkarte „Daten“ definiert die von der Anwendung für Analysen, Simulationen und Tests sowie fürs Scoring und andere operative Einsätze genutzten Daten-Sets. Diese Datenquellen sind relativ zur Installation Ihres Servers definiert – daher sind auch alle Pfadangaben zu Datendateien, ODBC-Datenquellen u. Ä. relativ zum Server zu verstehen.

Als Administrator möchten Sie vielleicht einige oder alle Daten konfigurieren, bevor die Geschäftsnutzer die Arbeit mit ihren Anwendungen aufnehmen. Administratoren haben die folgenden Möglichkeiten:

- Bei der Erstellung einer neuen Datenquelle Datenquelle sperren wählen, um nach der Erstellung eine Sperrung zu bewirken.
- Einzelne Datenquellen durch Klicken auf das Schloss-Symbol neben der Datenquelle in der Tabelle oder im Dialogfenster „Datenquellen-Editor“ sperren und entsperren.
- Durch Auswahl eines Datenmodells aus dem Dropdown-Menü Projektdatenmodell und Klicken auf das Schloss-Symbol das in der Anwendung benutzte Datenmodell einstellen und sperren.

Beachten Sie, dass Geschäftsnutzer eine gesperrte Datenquelle aus den Projekt-Daten-Sets in den Bereich „Eigene Daten-Sets“ kopieren können, sie dann aber in beiden Sets gesperrt bleibt. [Für weitere Informationen siehe Thema Verwaltung von Datenquellen in Kapitel 4 auf S. 28.](#)

Sperrung der Optionen zum Scoring-Ziel

Bei Anwendungen, die zur Verwendung des Abschnitts „Jetzt scoren“ (Registerkarte „Bereitstellen“ oder „Scores“) konfiguriert sind, ist ein Abschnitt verfügbar, der Geschäftsnutzern die Festlegung des Speicherziels der Scoring-Daten ermöglicht (Datenbank oder Datei). Wenn die Scores z. B. in einer Datenbank gespeichert werden sollen, geben Sie den Datenbanknamen und die Methode zum Speichern der Scores an (z. B. Erstellung einer neuen Tabelle in der Datenbank).

Als Administrator können Sie die Optionen für das Scoring-Ziel sperren. Sie können z. B. nur die Hinzufügung der Scores in eine bestehende Tabelle einer bestimmten Datenbank zulassen – anstatt Geschäftsnutzern die Wahl zwischen Überschreiben bestehender Daten, Hinzufügung zu einer bestehenden Tabelle, Erstellung einer neuen Tabelle oder Speicherung in einer Datei zu geben.

- ▶ Zum Sperren aller Optionen für das Scoring-Ziel klicken Sie auf das Schloss-Symbol neben dem Dropdown-Menü „Ziel“ auf der Registerkarte „Bereitstellen“. Geschäftsnutzer können die Zielloptionen dann einsehen, aber nicht modifizieren.

Herunterladen von Projekt-Metadaten zur Verwendung beim Erstellen externer Regeln

In einem Business Rules Management System wie ILOG erstellte Regeln können in IBM® Analytical Decision Management-Anwendungen referenziert und verwendet werden, vorausgesetzt, sie wurden so entwickelt, dass sie dasselbe Datenmodell unterstützen, das auch

im aktuellen IBM Analytical Decision Management-Projekt verwendet wird, und können als Webservice für die Anwendung bereitgestellt werden. Um dies zu erreichen, kann der IBM Analytical Decision Management-Administrator eine .ZIP-Datei mit Metadaten für das aktuelle Projekt herunterladen. Die heruntergeladene Datei wird lokal auf dem System gespeichert, das vom Administrator zum Auslösen des Downloads verwendet wird.

Abbildung 8-1
Symbol zum Herunterladen der Metadaten



Die .ZIP-Datei enthält folgende Dateien:

- **XML-Schemadefinition (*.XSD)**. Enthält Definitionen der Feldtypen im Datenmodell für das aktuelle IBM Analytical Decision Management-Projekt. Diese Datei kann in ein externes Entwicklungstool wie ILOG Rules Studio importiert werden, um Regeln zur Verwendung mit dem aktuellen Projekt zu entwickeln.
- **Web Service Description Language-Vorlage (*.WSDL)**. Enthalten zur Unterstützung bei der Entwicklung von Webservices, die das Datenmodell verwenden. Die .WSDL-Datei wird als Vorlage bereitgestellt, mit der Webservices erstellt werden können, die für die Vermittlung von Nachrichten zwischen der IBM Analytical Decision Management-Anwendung und einem externen Regeldienst zuständig sind. Details zum vermittelnden Webservice können manuell vom Integrator eingegeben werden.

Beachten Sie: Es ist möglich, ILOG-Regeln zu erstellen und zu nutzen, die verschiedene Typen an IBM Analytical Decision Management ausgeben (*boolean* (boolesch) zur Verwendung bei Auswahlen, *numeric* (numerisch) beispielsweise zur Verwendung in Aggregationsregeln sowie *string* (Zeichenkette). Um jedoch einen anderen Typ als *string* (Zeichenkette) bei der „dataresponse“-Regel zu erstellen, muss die XML-Schemadefinition (.XSD) manuell bearbeitet werden.

Auch Regelprojekte können mehrere Regeln enthalten, die jeweils verschiedene Typen ausgeben. Es ist erforderlich, mehrere Versionen der .XSD-Datei beizubehalten, um dies zu unterstützen (beispielsweise eine Version, bei der die Variable „dataresponse“ auf „numeric“ (Numerisch) gesetzt wurde, eine weitere für „string“ (Zeichenkette), eine für „boolean“ (Boolesch) usw.).

Die von IBM Analytical Decision Management heruntergeladene .XSD-Datei enthält zwei komplexe Datentypen: einen, der für das IBM Analytical Decision Management-Projekt Datenmodell steht, das für die Eingabeparameter von ILOG-Regelprojekten verwendet werden soll, und einen, der für einen Regelrückgabetyt steht, der für die Ausgabeparameter von ILOG-Regelprojekten verwendet werden soll. Der Rückgabetyt, (dessen Namen immer ein „response“-Suffix enthält) definiert ein `value`-Element, das standardmäßig den Typ *string* (Zeichenkette) aufweist. Wenn Sie ein Regelprojekt mit einer anderen Art von Rückgabewert erstellen möchten, müssen Sie die heruntergeladene .XSD-Datei manuell bearbeiten, bevor Sie das ILOG-Regelprojekt importieren, damit das Element `value` wunschgemäß definiert wird. Im Folgenden sehen Sie einige Beispiele:

```
<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>

```

```

<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:long"/>
  </xsd:sequence>
</xsd:complexType>

```

Informationen zum Erstellen externer Regeln zur Verwendung in IBM Analytical Decision Management finden Sie im *Handbuch für Anwendungsentwickler*.

Hilfe aufrufen

Es gibt drei Hilfsangebote:

- Für jede mit IBM® Analytical Decision Management ausgelieferte vorgefertigte Anwendung steht eine eigene Hilfe zur Verfügung. Um die Hilfe aufzurufen, klicken Sie auf das Hilfe-Symbol eines Bildschirms oder Dialogfelds:



- Infotext, welcher in kleinen Popup-Fenstern angezeigt wird und Erklärungen zur aktuellen Anwendung oder zum Geschäftsproblem enthält. Bei Infotexten handelt es sich um eine zusätzliche anpassbare, kontextsensitive Hilfefunktion, die für eine bestimmte Site oder Anwendung auf einfache Weise konfiguriert werden kann. Infotexte können durch Klicken auf die kleinen Infotext-Symbole der Benutzeroberfläche aufgerufen werden.



- Standard-Tool Tips, die für einzelne Steuerelemente einer Seite oder eines Dialogfelds bereitgestellt werden. Bewegen Sie den Mauszeiger auf die Schaltfläche, den Link oder ein anderes Steuerelement, um den zugehörigen Tool Tip anzuzeigen.

Verfügbare Dokumentation

Dokumentation für jedes Produkt bzw. jede Anwendung ist auf dem Installationsdatenträger oder in Form elektronischer Download-Dateien verfügbar. Die Dokumentation wird bei der Installation im Installationsverzeichnis von IBM® SPSS® Collaboration and Deployment Services (z. B. *C:\Program Files\IBM\SPSS\Deployment\5.0\help\de\DecisionManagement*) abgelegt.

Die Dokumentation ist zudem im PDF-Format online auf <http://www-01.ibm.com/support/docview.wss?uid=swg27024118> verfügbar, und Versionshinweise sind auf <http://www-01.ibm.com/support/docview.wss?uid=swg27024117> zu finden.

Online finden Sie eine zusammengefasste Sammlung der gesamten Dokumentation im [Informationscenter zu IBM Analytical Decision Management 7](#). In der folgenden Tabelle finden Sie Links zu allen verfügbaren Dokumenten im Informationscenter.

Tabelle 9-1
IBM Analytical Decision Management Dokumentation

Dokumentation	Beschreibung
Installationsanleitung für Windows Installationsanleitung für UNIX	Anleitungen zum Installieren und Starten des Produkts.
Handbuch für Anwendungsentwickler	Anleitungen zum Entwickeln und Konfigurieren von Anwendungen, darunter ein vollständiges Verzeichnis der XML-Schemata, sowie Einzelheiten zum Anpassen der Benutzeroberfläche.

Dokumentation	Beschreibung
Handbuch für Anwendungsbenuer	Enthält allgemeine Informationen über die Benutzeroberfläche für Geschäftsnutzer.
IBM SPSS Modeler Advantage	Mit IBM® SPSS® Modeler Advantage können Geschäftsnutzer vom Potenzial der Vorhersagemodellierung profitieren. Mithilfe von Vorhersagemodellen können Sie Muster in Vergangenheitsdaten aufdecken und die Wahrscheinlichkeit künftiger Ereignisse beurteilen.
IBM Analytical Decision Management for Customer Interactions	IBM® Analytical Decision Management for Customer Interactions entscheidet, welche Aktionsangebote Kunden unterbreitet werden, die sich an Ihr Unternehmen wenden, und übermittelt Empfehlungen für ein Call Center, eine Website oder ein Geschäft in Echtzeit. Durch Kombination der Logik der Geschäftsregeln mit den durch prädiktive Modellierung gewonnenen Erkenntnissen ermittelt die Anwendung die für jeden Kunden jeweils einträglichste Entscheidung.
IBM Analytical Decision Management for Campaign Optimization	IBM® Analytical Decision Management for Campaign Optimization ähnelt IBM Analytical Decision Management for Customer Interactions . Jedoch arbeitet IBM Analytical Decision Management for Campaign Optimization mit Optimierung und maximiert so den Return-on-Investment für Ihre Kampagne, indem unter Berücksichtigung Ihres Budgets, der Kapazitäten des Marketingkanals und der geltenden Kontaktrichtlinien die besten Angebote für einzelne Kunden bestimmt werden.
IBM Analytical Decision Management for Claims	IBM® Analytical Decision Management for Claims kombiniert die Logik der Geschäftsregeln mit den aus prädiktiven Modellen gewonnenen Erkenntnissen und ermöglicht es so Unternehmen, eingehende Schadensmeldungen in Echtzeit zu verarbeiten und die für jede Meldung beste Vorgehensweise zu ermitteln. Z. B. können Schadensmeldungen entweder im Schnellverfahren bearbeitet und zügig ausgezahlt, normal bearbeitet oder an eine spezielle Prüfungsabteilung übergeben werden.
IBM SPSS Rules Management	IBM® SPSS® Rules Management ist eine zentrale Schnittstelle zur Erstellung und Bearbeitung gemeinsamer Regeln, die anwendungsübergreifend genutzt werden können, um Datensätze auszuwählen und zu verarbeiten sowie den Entscheidungsprozess entsprechend zu automatisieren.

Freigabe von Streams

In IBM® Analytical Decision Management erstellte Modelle und Projekte werden als Stream-Dateien gespeichert und können, sofern verfügbar, bearbeitet und in IBM® SPSS® Modeler verwendet werden. Beispielsweise könnten Sie einen Modell- oder Projekt-Stream mit IBM® SPSS® Modeler Advantage oder einer anderen IBM Analytical Decision Management-Anwendung erstellen und anschließend den Stream in SPSS Modeler anpassen, bevor Sie ihn über IBM Analytical Decision Management bereitstellen. Alternativ können in SPSS Modeler erstellte Streams in IBM Analytical Decision Management geöffnet werden, vorausgesetzt, es wurden Breitstellungseinstellungen angegeben, wie weiter unten beschrieben.

Öffnen von Streams in IBM Analytical Decision Management

Öffnen von Streams in IBM Analytical Decision Management bzw. IBM SPSS Modeler Advantage:

- Kann kein Bereitstellungstyp angegeben wurde („none“), kann der Stream zwar geöffnet werden, es steht jedoch nur die Registerkarte „Daten“ zur Verfügung.
- Beim Bereitstellungstyp Nur Scoring steht eventuell auch die Registerkarte „Scoren“ zur Verfügung, vorausgesetzt, die Scoring-Verzweigung beginnt mit einem einzelnen Datenquellenknoten, dessen Typ von IBM SPSS Modeler Advantage bzw. IBM Analytical Decision Management unterstützt wird (keine zusammengeführten Daten-Sets).
- Beim Bereitstellungstyp Modellaktualisierung steht eventuell auch die Registerkarte „Modellierung“ zur Verfügung, vorausgesetzt, die Modellierungsverzweigung beginnt mit einem einzelnen Datenquellenknoten, dessen Typ von IBM SPSS Modeler Advantage bzw. IBM Analytical Decision Management unterstützt wird, und der ausgewählte Modellierungsknoten ist entweder ein Knoten für die automatisierte Modellierung (für Vorhersagemodelle unterstützt) oder ein Entscheidungslistenknoten (für Vorhersageregelmole (interaktiv) unterstützt). Außerdem muss ein Typknoten vorhanden sein, der das Zielfeld sowie andere bei der Modellierung verwendete Informationen definiert.

Angabe von Bereitstellungseinstellungen in IBM SPSS Modeler

In SPSS Modeler werden die Deployment-Einstellungen wie folgt im Dialogfeld „Stream-Eigenschaften“ angegeben:

- ▶ Wählen Sie die folgenden Befehle aus den SPSS Modeler-Menüs aus:
Extras > Stream-Eigenschaften > Bereitstellung
- ▶ Wählen Sie in der Liste Deployment-Typ die Option Nur Scoring bzw. Modellaktualisierung aus.
- ▶ Wählen Sie nach Bedarf einen Scoring-Knoten, Modellierungsknoten, und Modellknoten aus.

Wählen Sie Überprüfen, um zu bestätigen, dass alle erforderlichen Optionen angegeben wurden, oder Speichern, um den Stream in IBM® SPSS® Collaboration and Deployment Services Repository zu speichern.

Bei in IBM SPSS Modeler Advantage bzw. IBM Analytical Decision Management erstellten Streams werden die Bereitstellungsinformationen automatisch festgelegt.

Tipps zur Verwendung von IBM SPSS Modeler Streams in IBM Analytical Decision Management

- Es kann einfacher sein, Streams mit IBM SPSS Modeler Advantage bzw. IBM Analytical Decision Management zu erstellen und dann in SPSS Modeler zu bearbeiten, als sie ohne Vorlage mit SPSS Modeler zu erstellen. Auf diese Weise erstellte Streams sind so strukturiert, dass sie in IBM Analytical Decision Management voll funktionsfähig sind.
- Bei allen in IBM SPSS Modeler Advantage erstellten Streams sind Superknoten an bestimmten Positionen in der Verzweigung enthalten, die von erfahrenen Benutzern angepasst werden können. Es wird dringend empfohlen, etwaige Anpassungen oder Ergänzungen am Stream innerhalb dieser Superknoten vorzunehmen, um sicherzustellen, dass sie im Stream beibehalten und von IBM SPSS Modeler Advantage erkannt werden.
- Bei allen Knoten für lokale Regeln, die eine Ausdrucksregel unmittelbar unterhalb jedes Quellenknotens enthalten, werden die Felder in den Ausdrucksmanager aufgenommen. Beachten Sie, dass lokale Regeln nur in IBM Analytical Decision Management oder IBM SPSS Modeler Advantage erstellt werden können und nicht von SPSS Modeler.
- Wenn es oberhalb des Modellierungsknotens (und unterhalb des Typknotens) einen Partitionsknoten gibt, ist Daten automatisch partitionieren, um die Modellauswertung von Entwicklungs-Datenquellen zu ermöglichen (für weitere Evaluierung und Tests) in IBM SPSS Modeler Advantage aktiviert und der Zustand des Kontrollkästchens wird danach festgelegt, ob der Partitionsknoten aktiviert ist oder nicht.
- Wenn sich unterhalb des Typknotens ein ADP-Knoten (Automatische Datenvorbereitung) befindet, ist Daten automatisch bereinigen und zur zuverlässigen Modellerstellung aufbereiten aktiviert und der Zustand des Kontrollkästchens wird danach festgelegt, ob der ADP-Knoten aktiviert ist oder nicht.
- Wenn sich oberhalb des Modellierungsknotens Regel- oder Auswahlknoten befinden, werden diese dazu verwendet, den Abschnitt Zu benutzende Selektionen festlegen mit Daten zu füllen.
- Wenn Stream-Parameter definiert wurden (Registerkarte „Parameter“ im Dialogfeld „Stream-Eigenschaften“) wird der Link „Parameter“ auf den Registerkarten „Modellierung“ und „Scoren“ in IBM SPSS Modeler Advantage angezeigt. Die Benutzer können auf diesen Link klicken, um diese Werte anzugeben.

Starten von IBM SPSS Modeler Advantage aus IBM SPSS Modeler

Auf dem Startbildschirm von SPSS Modeler können Sie festlegen, dass ein Modell mit IBM SPSS Modeler Advantage erstellt werden soll.

Eingabehilfen

Eingabehilfen unterstützen Menschen mit Behinderung, z. B. mit eingeschränkter Mobilität oder mit Sehbehinderung, sodass diese technologische Produkte erfolgreich verwenden können. IBM ist bemüht, Produkte anzubieten, die unabhängig von Alter oder Geschicklichkeit von allen Menschen genutzt werden können. Bei diesem Produkt werden die standardmäßigen Navigationstasten von Webbrowsern verwendet.

Da IBM® Analytical Decision Management eine browserbasierte Anwendung ist, werden Eingabehilfen für Menschen mit Behinderung, z. B. Sehbehinderte, über die Webbrowser-Einstellungen konfiguriert. Für weitere Informationen zu Bedienmöglichkeiten wie z. B. Tastaturkürzeln, Bildschirmleseprogrammen o. Ä. ziehen Sie die Dokumentation Ihres Webbrowsers heran.

Drittanbietersoftware

IBM Analytical Decision Management kann die Verwendung von Drittanbietersoftware erforderlich machen, die nicht von der IBM-Lizenzvereinbarung abgedeckt wird. IBM gibt keine Gewährleistung ab hinsichtlich der Eingabehilfen dieser Produkte. Wenden Sie sich an den Hersteller, um Informationen zu den Eingabehilfen seiner Produkte zu erhalten.

IBM und Eingabehilfen

Im [IBM-Center zu menschlichen Fähigkeiten und Eingabehilfen](#) finden Sie weitere Informationen zum Engagement von IBM im Bereich Eingabehilfen.

Zugriffsmöglichkeiten auf die Hilfe

Es steht eine alternative Version der mit dem Produkt installierten Benutzerhilfe zur Verfügung, die Bildschirm-Lesesysteme unterstützt. Um zu dieser alternativen Version zu wechseln, müssen Sie eine Stapelverarbeitungsdatei ausführen, die die vorhandene Hilfe in ein Format umwandelt, das leichter von Bildschirm-Lesesystemen verarbeitet werden kann. In der alternativen Version befindet sich das Inhaltsverzeichnis rechts, und die Hilfethemen werden links angezeigt.

So aktivieren Sie die Unterstützung für Bildschirm-Lesesysteme

- ▶ Öffnen Sie das Hauptverzeichnis der Hilfe auf dem Server (z. B. *C:\Program Files\IBM\SPSS\Deployment\5.0\help\en\DecisionManagement\ClaimsManagement\userhelp*).
- ▶ Um zu dem für Bildschirm-Lesesysteme geeigneten Format zu wechseln, doppelklicken Sie auf die Stapelverarbeitungsdatei *format-for-screen-readers.bat*.

Um wieder zum Standardformat zurückzukehren, führen Sie die Stapelverarbeitungsdatei *restore-default-format.bat* aus.

Anmerkung: Das [Informationscenter zu IBM Analytical Decision Management 7](#) weist Eingabehilfen auf.

Hinweise

Diese Informationen wurden für Produkte und Dienstleistungen entwickelt, die weltweit angeboten werden.

IBM bietet die in diesem Dokument behandelten Produkte, Dienstleistungen oder Funktionen möglicherweise nicht in anderen Ländern an. Wenden Sie sich an Ihren lokalen IBM-Händler, um Informationen zu den derzeit in Ihrem Raum verfügbaren Produkten und Dienstleistungen zu erhalten. Bezugnahmen auf IBM-Produkte, -Programme oder -Dienstleistungen deuten nicht darauf hin, dass nur das entsprechende Produkt, Programm oder die Dienstleistung verwendet werden kann. Sämtliche in Bezug auf die Funktionalität gleichwertigen Produkte, Programme oder Dienstleistungen, die nicht gegen die Rechte von IBM am geistigen Eigentum verstoßen, können stattdessen verwendet werden. Es obliegt jedoch der Verantwortung des Benutzers, den Betrieb eines nicht von IBM stammenden Produkts, Programms bzw. einer Dienstleistung zu evaluieren und zu überprüfen.

IBM verfügt möglicherweise über Patente oder hat Patentanträge gestellt, die sich auf in diesem Dokument beschriebene Inhalte beziehen. Durch die Bereitstellung dieses Dokuments werden Ihnen keinerlei Lizenzen an diesen Patenten gewährt. Lizenzanfragen können schriftlich an folgende Adresse gesendet werden:

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

Bei Lizenzanfragen in Bezug auf DBCS-Daten (Double-Byte Character Set) wenden Sie sich an die für geistiges Eigentum zuständige Abteilung von IBM in Ihrem Land. Schriftliche Anfragen können Sie auch an folgende Adresse senden:

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

Der folgende Abschnitt findet in Großbritannien und anderen Ländern keine Anwendung, in denen solche Bestimmungen nicht mit der örtlichen Gesetzgebung vereinbar sind: INTERNATIONAL BUSINESS MACHINES STELLT DIESE VERÖFFENTLICHUNG IN DER VERFÜGBAREN FORM OHNE GARANTIEN BEREIT, SEIEN ES AUSDRÜCKLICHE ODER STILLSCHWEIGENDE, EINSCHLISSLICH JEDOCH NICHT NUR DER GARANTIEN BEZÜGLICH DER NICHT-RECHTSVERLETZUNG, DER GÜTE UND DER EIGNUNG FÜR EINEN BESTIMMTEN ZWECK. Manche Rechtsprechungen lassen den Ausschluss ausdrücklicher oder implizierter Garantien bei bestimmten Transaktionen nicht zu, sodass die oben genannte Ausschlussklausel möglicherweise nicht für Sie relevant ist.

Diese Informationen können technische Ungenauigkeiten oder typografische Fehler aufweisen. An den hierin enthaltenen Informationen werden regelmäßige Änderungen vorgenommen. Diese Änderungen werden in neuen Ausgaben der Veröffentlichung aufgenommen. IBM kann jederzeit und ohne vorherige Ankündigung Optimierungen und/oder Änderungen an den Produkten und/oder Programmen vornehmen, die in dieser Veröffentlichung beschrieben werden.

Jegliche Verweise auf Drittanbieter-Websites in dieser Information werden nur der Vollständigkeit halber bereitgestellt und dienen nicht als Befürwortung dieser. Das Material auf diesen Websites ist kein Bestandteil des Materials zu diesem IBM-Produkt und die Verwendung erfolgt auf eigene Gefahr.

IBM kann die von Ihnen angegebenen Informationen verwenden oder weitergeben, wie dies angemessen erscheint, ohne Ihnen gegenüber eine Verpflichtung einzugehen.

Lizenznehmer dieses Programms, die Informationen dazu benötigen, wie (i) der Austausch von Informationen zwischen unabhängig erstellten Programmen und anderen Programmen und (ii) die gegenseitige Verwendung dieser ausgetauschten Informationen ermöglicht wird, wenden sich an:

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

Diese Informationen sind je nach den entsprechenden Bedingungen und in manchen Fällen gegen Zahlung einer Gebühr erhältlich.

Das in diesem Dokument beschriebene lizenzierte Programm und sämtliche dafür verfügbaren lizenzierten Materialien werden von IBM gemäß dem IBM-Kundenvertrag, den Internationalen Nutzungsbedingungen für Programmpakete der IBM oder einer anderen zwischen uns getroffenen Vereinbarung bereitgestellt.

Alle in diesem Dokument enthaltenen Leistungsdaten wurden in einer kontrollierten Umgebung ermittelt. Daher können die unter anderen Betriebsumgebungen erzielten Ergebnisse erheblich abweichen. Einige Messungen wurden möglicherweise an Systemen im Entwicklungsstadium vorgenommen und es besteht keine Garantie, dass spätere allgemein verfügbare Systeme dieselben Messwerte aufweisen. Außerdem wurden einige Messwerte möglicherweise mittels Extrapolation geschätzt. Die tatsächlichen Ergebnisse können abweichen. Die Benutzer dieses Dokuments sollten die entsprechenden Daten für ihre jeweilige Umgebung überprüfen.

Informationen zu Nicht-IBM-Produkten stammen von den Herstellern dieser Produkte, ihren veröffentlichten Verlautbarungen oder aus anderen öffentlich verfügbaren Quellen. IBM hat diese Produkte nicht getestet und kann daher die Richtigkeit der Angaben zu Leistung und Kompatibilität oder anderer Behauptungen in Bezug auf Nicht-IBM-Produkte nicht bestätigen. Fragen zu den Funktionen von Nicht-IBM-Produkten sind an die Hersteller dieser Produkte zu richten.

Alle Aussagen in Bezug auf die zukünftige Ausrichtung oder die zukünftigen Vorhaben von IBM können ohne Vorankündigung geändert oder widerrufen werden und stellen lediglich Zielsetzungen dar.

Diese Informationen enthalten Beispiele zu Daten und Berichten, die im täglichen Geschäftsbetrieb Verwendung finden. Um diese so vollständig wie möglich zu illustrieren, umfassen die Beispiele Namen von Personen, Unternehmen, Marken und Produkten. Alle diese Namen sind fiktiv und jegliche Ähnlichkeit mit Namen und Adressen realer Unternehmen ist rein zufällig.

Bei der Anzeige dieser digitalen Informationsversion sind die Fotografien und Farbillustrationen möglicherweise nicht sichtbar.

Marken

IBM, das IBM-Logo, ibm.com und SPSS sind Marken der IBM Corporation und in vielen Ländern weltweit registriert. Eine aktuelle Liste der IBM-Marken finden Sie im Internet unter <http://www.ibm.com/legal/copytrade.shtml>.

Adobe, das Adobe-Logo, PostScript und das PostScript-Logo sind eingetragene Marken oder Marken von Adobe Systems Incorporated in den USA und/oder anderen Ländern.

Linux ist eine eingetragene Marke von Linus Torvalds in den USA, anderen Ländern oder beidem.

Microsoft, Windows, Windows NT und das Windows-Logo sind Marken der Microsoft Corporation in den USA, anderen Ländern oder beidem.

UNIX ist eine eingetragene Marke der The Open Group in den USA und anderen Ländern.

Java und alle Java-basierten Marken sowie Logos sind Marken von Sun Microsystems, Inc. in den USA, anderen Ländern oder beidem.

SAS ist eine eingetragene Marke von SAS Institute Inc. in den USA und/oder anderen Ländern.

Andere Produkt- und Servicenamen können Marken von IBM oder anderen Unternehmen sein.

Index

- Allgemeine Optionen, 84
- Anmerkungen
 - Zu Regeln hinzufügen, 48
- Antwortdaten, 8
- Anwendungen
 - Starten, 7
 - Testen, 72
- Anwendungs-Hauptseite, 22
- Anwendungs-Startbereich, 20, 85
 - Anpassen, 20
 - Anwendungen hinzufügen, 20
 - Dialogfeld für die Projektkonfiguration, 21
- Anwendungs-URL, 20
- Anwendungs-Verknüpfungen, 20, 85
- Assoziationsmodelle, 56, 60
 - Tabellen- im Vergleich zu Transaktionsdaten, 57
 - Tabellendaten, 57
 - Transaktionsdaten, 57
- Assoziationsregelmodelle, 56
- Ausblenden der Hauptseite, 85
- Ausdruckseditor, 41
- Ausdrucksmanager, 41
- Ausgabefelder in Modellen, 79
- Auswahlregeln, 45
- Auswerten von Modellen, 10
 - Clustermodelle, 14, 16
 - Interaktive Listenmodelle, 12
 - Vorhersageregelmodelle, 12
- Auswertung von Modellen, 62

- bank_customer_data.txt, 8
- bank_response_data.txt, 8
- Bearbeiten von Zusatzfeldern, 39
- Beispiel für Assoziationsmodellierung, 16
- Beispielanwendungen, 6
- Beispieldateien, 6
- Beispieldaten, 6
- Benutzervoreinstellungen, 23
- Bildschirm-Lesesysteme, 93

- Cluster-Modellierung, Beispiel, 14
- clustering, 55
- Clustermodelle, 54, 59
- Cognos BI Server als Ziel in Modellen, 77
- Cognos BI-Quellen, 37

- Dateien für Beispielanwendungen, 6
- Dateien speichern, 24–26
- Dateiziel in Modellen, 76
- Daten für Beispielanwendungen, 6
- Datenauswahl in Modellen, 80
- Datenbankquellen, 36
- Datenbankziel in Modellen, 75
- Datendateien, 6
- Datenoptionen, 86

- Datenquellen
 - Ausdruckseditor, 41
 - Dateiquellen, 33
 - Datenbank, 36
 - Definieren, 8, 28, 32
 - Eingabefelder , 40
 - Excel, 34
 - Hinzufügen von Zusatzfeldern, 39
 - IBM Cognos BI, 37
 - IBM SPSS Collaboration and Deployment Services
 - Enterprise View, 36
 - IBM SPSS Statistics, 35
 - Messniveaus, 31
 - textbasiert, 33
 - Vorschau, 31
 - Zuordnen von Feldern, 41
- Datensatzauswahl in Modellen, 81
- Datentypen, 31
- Datumsformate, 23
- Demo-Objekte, 6
- Dezimaltrennzeichen, 33
- Diagramm für die Bedeutsamkeit der Variablen, 52
- Diagramme in Modellen, 52
- Dialogfeld für die Projektkonfiguration, 21
- Direktwerbung – Beispiel, 5

- Eigene Datenquellen, 28
- Eingabefelder, 8
 - Definieren, 40
 - zuordnen, 41
- Eingabehilfen, 93
- Erstellen von Modellen
 - Zielfeld, 10
- Excel-Datendateien, 34
- Externe Regeln
 - Herunterladen von Projekt-Metadaten, 86
 - Verwendung in Anwendungen, 48

- Felder
 - ableiten, 41
 - Definieren, 8
 - target, 8
 - Trennzeichen, 33
 - zuordnen, 41
- Felder abbilden, 79
- Flag-Daten, 31
- Freigabe von Streams zwischen IBM Analytical Decision Management und IBM SPSS Modeler, 91

- Galerie
 - lang laufende Aufgaben, 22
 - nicht fertig gestellte Aufgaben, 22
- Gemeinsame Regeln, 46
- Geschäftsregeln, 43
- Gewinndiagramme, 10, 64

- Hauptseite, 22, 85
- Hilfe, 89
 - Alternative Version, 93
 - Eingabehilfen, 93
 - für Bildschirm-Lesesysteme, 93
- Hilfe aufrufen, 89
 - Verfügbare Dokumentation, 89
- Hinzufügen von Zusatzfeldern, 39
- Hochladen von Stream-Dateien, 24
- Hochladen von Streams, 24

- IBM Cognos BI-Quellen, 37
- IBM Cognos BI-Quellknoten
 - Symbole, 39
- IBM SPSS Collaboration and Deployment Services
 - Enterprise View-Datenquellen, 36
- IBM SPSS Collaboration and Deployment Services Repository
 - Speichern in, 24–26
- IBM SPSS Modeler Advantage
 - Beispiel, 5
- IBM SPSS Statistics Datendateien, 35
- ILOG-Regeln
 - Herunterladen von Projekt-Metadaten, 86
 - Verwendung in Anwendungen, 48
- Infotext, 89
- Interaktive Listenmodelle, 53, 59

- k*-Means-Modelle, 55
- Kassenrollendaten, 57
- Kategoriale Daten, 31
- Kohonen, Modelle, 55
- Komma, 33
- Kundendaten, 8

- Lift-Diagramme, 65

- Manueller Cluster
 - clustering, 55
- Marken, 97
- Messniveaus, 31
- Modell testen, 72
- Modelle
 - association, 56
 - Auswahl eines Ausgabefelds, 79
 - Auswahl eines Dateiziels, 76
 - Auswahl eines Datenbankziels, 75
 - Auswahl eines Scoring-Ziels, 75
 - Auswahl von Cognos BI Server als Ziel, 77
 - clustering, 54
 - Datenauswahl, 80
 - Datensatzauswahl, 81
 - Diagramme, 52
 - Ergebnisse, 52
 - Erstellen, 10, 50
 - Evaluiieren, 10, 62
 - Interaktive Liste, 53
 - Optionale Einstellungen, 58–60
 - Profit simulieren, 69
 - Scoring, 13, 75, 81
 - Scoring-Felder abbilden, 79
 - Testen, 13, 72
 - Verwenden in Regeln, 43
 - Voreinstellungen anpassen, 23
 - Vorhersage, 50–51
 - Vorhersageregeln, 53
 - Zeit für die Erstellung beschränken, 23
 - Zielfeld, 10, 50
- Modelle testen, 13

- nicht fertig gestellte Objekte
 - Galerie, 22
- Nicht überwachtetes Lernen, 55
- Nicht verfeinerte Modelle, 56
- Nominale Daten, 32

- Objekte für Beispieldaten, 6
- Öffnen von Dateien, 24
- Online-Hilfe, 89
 - Alternative Version, 93
 - Eingabehilfen, 93
 - für Bildschirm-Lesesysteme, 93
- Optionen für das Scoring-Ziel, 86
- Ordinaldaten, 32

- Prädiktoreinfluss-Diagramm, 52
- Profit
 - maximieren, 5
 - Simulieren, 10
- Profitsimulation in Modellen, 69
- Profitsimulationen, 10, 67
- Profitsimulation in Modellen, 69
- Projektdateienmodell, 8, 28
- Projektdateiquellen, 28
- Projekte sperren, 26

- Rechtliche Hinweise, 95
- Regeln, 43
 - Anmerkungen, 48
 - Auswahl, 45
 - Erstellen, 43
 - Exportieren, 46
 - externe, 48, 86
 - gemeinsame, 46, 48, 86
 - ILOG, 48, 86
 - Modelle verwenden mit, 43
 - Segment, 43
 - Selektionslogik, 46
 - Wiederverwenden, 46, 48, 86
- Regeln ausschließen, 45
- Regeln einschließen, 45
- Regeln exportieren, 46
- ROI-Diagramme, 68

- Schlüsselwörter, 24
- Scoring von Modellen, 13, 75, 81
 - Ausgabefelder auswählen, 79
 - Cognos BI Server als Ziel, 77
 - Dateiziel, 76
 - Daten auswählen, 80
 - Datenbankziel, 75
 - Felder abbilden, 79
 - Scoring-Ziel auswählen, 75
- Scoring-Felder abbilden, 79
- Scoring-Ziel in Modellen, 75
- Segmentregeln
 - Modelle verwenden mit, 43
- Sequenzerkennung, 56
- Simulation
 - Standarddatum, 24
- Simulation des Profits in Modellen, 69
- Simulationsdatum, 24
- Sperrung der Datenoptionen, 86
- Sperrung der Optionen zum Scoring-Ziel, 86
- Startbereich, 7, 20, 85
 - Anpassen, 20
 - Anwendungen hinzufügen, 20
 - Dialogfeld für die Projektkonfiguration, 21
- Stetige Daten, 31
- Steuerung
 - Allgemeine Optionen, 84
- Streams
 - Hochladen, 24
- Symbole, IBM Cognos BI, 39

- Tabellendaten, 57
- Testen einer Anwendung, 72
- textbasierte Datendateien, 33
- Tool Tips, 89
- Transaktionsdaten, 57
- Trefferdiagramme, 66
- Trennzeichen, 33
- TwoStep-Cluster-Modelle, 55
- Typlose Daten, 32

- Überblick
 - Verwaltung, 83
- URL
 - Anwendung, 20

- Verfügbare Dokumentation, 89
- Verknüpfungen
 - Anwendung, 20
- Versionen beschriften, 26
- Versionsbeschriftungen, 26
- Verteilungsdiagramm, 52
- Verteilungsdiagramme, 64
- Verwaltung
 - Allgemeine Optionen, 84
 - Anwendungs-Verknüpfungen, 85
 - Datenoptionen, 86
 - Hauptseite, 85
 - Optionen für das Scoring-Ziel, 86
 - Überblick, 83
 - Verwaltung – Überblick, 83
 - Voreinstellungen, 23
 - Vorhersagemodelle, 50–51
 - Vorhersageregelmamodelle, 53, 59
 - Vorkonfiguration von Anwendungs-Verknüpfungen, 85
 - Vorschau von Daten, 31

 - Wahrheitstabellendaten, 57
 - Währungsformate, 23
 - Warenkorbdaten, 57

 - Zeitformate, 23
 - Zielfeld, 8
 - Erstellen von Modellen, 10, 50
 - ZIP-Dateien für Beispielanwendungen, 6
 - Zuordnen von Feldern, 41
 - Zusatzfelder hinzufügen/bearbeiten, 39