

IBM® Analytical Decision Management
Version 18 Release 0

IBM® SPSS® Modeler Advantage

IBM

Hinweis

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen“ auf Seite 101 gelesen werden.

Produktinformation

Diese Ausgabe bezieht sich auf Version 18, Release 0, Modifikation 0 von IBM Analytical Decision Management und alle nachfolgenden Releases und Modifikationen, bis dieser Hinweis in einer Neuausgabe geändert wird.

Diese Veröffentlichung ist eine Übersetzung des Handbuchs
IBM Analytical Decision Management, Version 18 Release 0, IBM SPSS Modeler Advantage,
herausgegeben von International Business Machines Corporation, USA

© Copyright International Business Machines Corporation 2010, 2015

Informationen, die nur für bestimmte Länder Gültigkeit haben und für Deutschland, Österreich und die Schweiz nicht zutreffen, wurden in dieser Veröffentlichung im Originaltext übernommen.

Möglicherweise sind nicht alle in dieser Übersetzung aufgeführten Produkte in Deutschland angekündigt und verfügbar; vor Entscheidungen empfiehlt sich der Kontakt mit der zuständigen IBM Geschäftsstelle.

Änderung des Textes bleibt vorbehalten.

Herausgegeben von:
TSC Germany
Kst. 2877
Dezember 2015

© Copyright IBM Corporation 2010, 2015.

Inhaltsverzeichnis

Vorwort	v	Überprüfen von Ausdrücken	51
Kapitel 1. IBM Analytical Decision Management	1	Verwenden von Ausdrücken, die einen booleschen Ausdruck zurückgeben	52
Produktübersicht	1	Kapitel 5. Verwenden von Regeln	53
Neuerungen in Version 18	1	Erstellen von Regeln	53
Vorgängerversionen	2	Definieren von Segmentregeln	53
Kapitel 2. Beispiel für die Modellierung der Kundenreaktion	11	Definieren von Auswahlregeln	55
IBM SPSS Modeler Advantage	11	Exportieren und Wiederverwenden von Regeln	56
Maximieren des Profits für Direktwerbungskampagnen	11	Hinzufügen von Anmerkungen.	57
Beispieldateien	11	Verwenden von Regeln aus IBM Operational Decision Management	58
Starten der Anwendung	13	Kapitel 6. Erstellen von Vorhersagemodellen	59
Einrichten von Datenquellen.	13	Informationen zu Vorhersagemodellen	59
Erstellen des Modells	15	Erstellen von Modellen	59
Auswerten des Modells	17	Erstellen eines Vorhersagemodells	59
Vorhersageregelmole	20	Ergebnisse der Vorhersagemodellierung	60
Testen des Modells	21	Erstellen eines Vorhersageregelmole (interaktiv)	61
Scoren des Modells	21	Erstellen eines Clustermodells	63
Beispiel für Clustermodellierung	22	Erstellen eines Assoziationsmodells	64
Beispiel für Assoziationsmodellierung	23	Optionale Modelleinstellungen	65
Kapitel 3. Navigation im Arbeitsbereich 27		Auswerten von Modellen.	69
Ändern der Anzeigesprache	27	Simulieren des maximalen Profits	71
Starten von Anwendungen	27	Testen des Modells	72
Homepage.	28	Verwenden von Modellscores in Anwendungen	72
Galerie	28	Kapitel 7. Scoren von Modellen	75
Anpassen von Voreinstellungen.	29	Scoring von Modellen in eine Datenbanktabelle, eine Datei, eine Analytic Server-Instanz oder einen IBM Cognos BI-Server	75
Öffnen und Speichern von Projekten, Modellen und Regeln	31	Auswählen von Daten und Datensatzgruppen.. . . .	77
Öffnen und Speichern über das Repository	31	Auswählen von Datensätzen für Scoring.	77
Arbeiten mit Beschriftungen.	32	Auswählen des Scoring-Ziels	77
Sperrern von Projekten	33	Auswählen von Ausgabefeldern	81
Eigenschaften.	33	Scoring von Modellen	82
Kapitel 4. Verwalten von Datenquellen 35		Scoring-Zusammenfassungsbericht	82
Erstellen einer neuen Datenquelle	37	Kapitel 8. Verwalten von Anwendungen 85	
Dateiquelle	38	Allgemeine Optionen	86
Datenbankquelle.	41	Vorkonfiguration von Anwendungsverknüpfungen	86
Datenansichtsquelle	41	Ausblenden der Homepage	87
IBM Cognos BI-Quelle.	42	Sperrern der Datenoptionen	87
Analytic Server-Quelle.	44	Sperrern der Optionen zum Scoring-Ziel	88
Auswählen von Eingabefeldern.	45	Herunterladen von Projektmetadaten zur Verwendung beim Erstellen externer Regeln	88
Vorschau von Daten	45	Bearbeiten der XML-Schemadefinition	89
Datenübersicht	45	Kapitel 9. Aufrufen von Hilfe	95
Messniveaus	46	Verfügbare Dokumentation	95
Bearbeiten von Datenfeldwerten	47		
Abbildern von Feldern	47		
Ableiten von Feldern	48		
Expression Builder	49		
Auswählen von Funktionen	50		
Auswählen von Feldern und Werten	51		

Anhang A. Gemeinsame Nutzung von Streams 97

Anhang B. Behindertengerechte Bedienung 99

Bemerkungen 101

Marken 102

Bedingungen für Produktdokumentation 102

Index 105

Vorwort

Informationen zu IBM Business Analytics

Die Software IBM Business Analytics bietet umfassende, einheitliche und genaue 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. Kombiniert mit umfassenden Branchenlösungen, bewährten Verfahren und professionellen Services können Unternehmen jeder Größe die Produktivität maximieren, sicher Entscheidungen automatisieren und bessere Ergebnisse erzielen.

Als Teil dieses Portfolios unterstützt die Software IBM SPSS Predictive Analytics Unternehmen dabei, zukünftige Ereignisse vorherzusagen und proaktiv Maßnahmen zu ergreifen, um bessere Geschäftsergebnisse zu erzielen. Kunden aus der Wirtschaft, dem öffentlichen Dienst und dem Bildungs- und Forschungsbereich weltweit vertrauen auf IBM® SPSS-Technologie als Wettbewerbsvorteil für Kundengewinnung und Kundenbindung und für die Erhöhung der Kundenumsätze bei gleichzeitiger Eindämmung von Betrugsfällen und Risiken. Durch Einbindung von IBM SPSS-Software in ihre täglichen Unternehmensaktivitäten werden Unternehmen zu "Predictive Enterprises", die Entscheidungen so ausrichten und automatisieren können, dass diese den Geschäftszielen entsprechen, und die dadurch einen messbaren Wettbewerbsvorteil erzielen können. Wenn Sie weitere Informationen erhalten oder Kontakt zu einem Ansprechpartner aufnehmen möchten, rufen Sie die Seite <http://www.ibm.com/spss> auf.

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 oder bei der Installation in einer der unterstützten Hardware-Umgebungen benötigen. Zur Kontaktaufnahme mit dem technischen Support besuchen Sie die Website von IBM unter <http://www.ibm.com/support>. Sie müssen bei der Kontaktaufnahme Ihren Namen, Ihre Organisation und Ihre Supportvereinbarung angeben.

Kapitel 1. IBM Analytical Decision Management

Produktübersicht

IBM Analytical Decision Management macht Vorhersageanalysen für tatsächliche betriebswirtschaftliche Probleme nutzbar und ermöglicht es Ihnen, speziell auf Ihre Kunden oder Ihre Branche zugeschnittene Anwendungen zu erstellen. Anwendungen sind in der Regel für die Behebung spezieller Probleme konfiguriert, basieren jedoch auf einem gemeinsamen Funktionsatz:

- 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.

Neuerungen in Version 18

In IBM Analytical Decision Management 18 (18.0.0.0) wurden die folgenden neuen Funktionen hinzugefügt:

Im Folgenden finden Sie eine Beschreibung der wichtigsten neuen Funktionen:

Szenariolaufzeit. Im Dialogfeld **Szenarios verwalten** gibt es ein neues Eigenschaftensymbol, über das Sie Details zu einem Szenario anzeigen können, beispielsweise die Startzeit, Endzeit und Laufzeit.

CPLEX-Protokoll. Wenn Sie das CPLEX-Protokoll zur Fehlerbehebung in einer Datei auf dem Dateisystem von IBM SPSS Modeler Server speichern wollen, steht jetzt für Anwendungen, die über eine Registerkarte **Optimieren** verfügen, die neue Option **CPLEX-Engineprotokoll in Datei umleiten** im Dialogfeld **Erweiterte Optionen** zur Verfügung. In dem seltenen Fall, dass eine CPLEX-Anwendung dem Scoring-Service bereitgestellt wird, sollten Sie die Protokollierung inaktivieren, um die Leistung zu verbessern.

Laden von Massendaten und Stapelcommitgröße. Bei der Ausgabe an eine Datenbank können Sie auf der Registerkarte **Bereitstellen** jetzt die Stapelgröße angeben und das Laden von Massendaten aktivieren. Dadurch wird die Leistung beim Scoring einer großen Tabelle und beim Schreiben der Ergebnisse in eine Datenbank verbessert.

Datenübersicht im Kontext. Die Datenübersicht wird jetzt in dem Kontext angezeigt, in dem das Feld innerhalb des Projekts verwendet wird. Beispielsweise werden zuvor im Projekt erstellte Regeln berücksichtigt.

Neue Einschränkung für Optimierung. Der Anwendung IBM Analytical Decision Management for Campaign Optimization wurde eine neue Einschränkung für *Minimale Anzahl an Angeboten, die für diesen Angebotstyp angenommen werden müssen* hinzugefügt.

Speicherverbesserungen. Für die Anwendung IBM Analytical Decision Management for Campaign Optimization wurden Speicherverbesserungen vorgenommen. Weitere Informationen finden Sie in **Bewährte Verfahren bei Leistungsfragen > CPLEX-Optimierung > Mindestspeicherbedarf** im Handbuch *Leistung und Fehlerbehebung*.

Vorgängerversionen

Version 17

In IBM Analytical Decision Management 17 (17.0.0.0) wurden die folgenden neuen Funktionen hinzugefügt:

Im Folgenden finden Sie eine Beschreibung der wichtigsten neuen Funktionen, die in Version 17 hinzugefügt wurden:

Expression Builder. Expression Builder wurde erweitert, um komplexere und leistungsfähigere Ausdrücke zu ermöglichen. Expression Builder in IBM Analytical Decision Management stimmt nun besser mit Expression Builder im Client von IBM SPSS Modeler überein. Siehe http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_user_guide_ddita/configurableapps/dm_expressions.dita.

Registerkarte "Abgeleitet". In den meisten Anwendungen von IBM Analytical Decision Management können Sie Ihrem Projektdatenmodell Zusatzfelder hinzufügen, z. B. Ausdrücke, Modellausgabe, über Segmentregeln erstellte Felder oder Felder aus Sekundärtabellen. Auf der neuen Unterregisterkarte **Abgeleitet** der Registerkarte **Daten** werden alle Felder angezeigt, die als Teil der Daten eingeschlossen werden sollen. Siehe http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_user_guide_ddita/configurableapps/dm_app_derived.dita.

Abschnitt "Sekundärtabellen". Den meisten Anwendungen von IBM Analytical Decision Management wurde ein neuer Abschnitt **Sekundärtabellen** hinzugefügt. Er ersetzt die verschiedenen Links **Zusatztabellen hinzufügen/bearbeiten**, die in Vorgängerreleases vorhanden waren. Siehe http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_user_guide_ddita/configurableapps/dm_app_additional_fields.dita.

Mehrfaktorauthentifizierung. IBM SPSS Collaboration and Deployment Services 7, IBM SPSS Modeler 17 und IBM Analytical Decision Management 17 wurde Funktionalität hinzugefügt, die Mehrfaktorauthentifizierung (MFA) unterstützt. Diese Erweiterungen können die Sicherheit für umfangreiche Unternehmensumgebungen verbessern. Beispielsweise können SPSS-Produkte in die vorhandene Kerberos-Single Sign-on-Infrastruktur eines Unternehmens und in Smartcard-Infrastrukturen integriert werden (z. B. wenn die Mitarbeiter eines Unternehmens Ausweise und Kartenleser für Computer- und Anwendungsauthentifizierung verwenden). Siehe http://www-01.ibm.com/support/knowledgecenter/SS6A3P_17.0.0/adm_install_guide_ddita/configurableapps/dms_install_security.dita.

Datenquellenzuordnung. Sofern möglich, ordnet IBM Analytical Decision Management jetzt automatisch kompatible Speichertypen zu, wenn ein Benutzer eine Zusatzdatenquelle hinzufügt. Wenn der Typ nicht kompatibel ist, darf der Benutzer Typen umwandeln.

Dateispeichertypen. Benutzer können jetzt den Speichertyp und das Format festlegen, wenn in Datendateien mit festem und variablem Format gelesen wird, um sicherzustellen, dass Werte ordnungsgemäß gelesen werden. Wenn z. B. beim Hinzufügen einer neuen Datenquelle in IBM Analytical Decision Management der Datenquellentyp **Datei** und wenn das Dateiformat **Felder unterschiedlicher Länge in Nur-Text** ist, enthält das Dialogfeld **Datenquelle** jetzt den neuen Abschnitt **Datenformat angeben**, in dem der Benutzer den Eingabetyp und das Format der einzelnen Felder angeben kann. Dies ist der im Client von IBM SPSS Modeler verfügbaren Funktion vergleichbar.

Projektübersicht. Für Anwendungen mit der Registerkarte **Bereitstellen** wurde der neue Abschnitt **Projektübersicht** hinzugefügt. In diesem Abschnitt können Sie das Projekt validieren und in ihm werden wichtige Projektübersichtsinformationen angezeigt, einschließlich gefundener Fehler, die sich auf die Bereitstellung auswirken können. In der Projektübersicht werden Informationen wie z. B. alle Attributreferenzen und alle Repository-Referenzen angezeigt, die als *.csv-Datei exportiert werden können. Sie können auch andere Registerkarten der Anwendung ansteuern, um Fehler zu korrigieren, und dann zur Projektübersicht auf der Registerkarte **Bereitstellen** zurückkehren.

Modellreferenzen. Wenn Sie Modelle in einer Anwendung von IBM Analytical Decision Management referenzieren, hat sich das Verhalten des Modelllinks geändert. Wenn Sie auf einen Modellreferenzlink klicken, wird das Dialogfeld des Modellviewers geöffnet, in dem Details zum Modell angezeigt werden können. Sie können jetzt auch mit der rechten Maustaste auf die Modellreferenz klicken, um die Option **Referenzieren** oder **Erstellen** auszuwählen. Wenn Sie auf eine Regelreferenz klicken, wird der Regelviewer geöffnet (so wie in Vorgängerreleases). Sie können auch mit der rechten Maustaste auf eine Regelreferenz klicken und **Referenzieren** auswählen.

Analytic Server-Berechtigungsachweisuffix. Administratoren können mit einer browserbasierten Instanz von IBM SPSS Collaboration and Deployment Services Deployment Manager ein Suffix konfigurieren, um es an Analytic Server-Berechtigungsachweis anzuhängen. Hierdurch wird der Name des Berechtigungsachweises abgeleitet, der für Analytic Server-Verbindungen in IBM Analytical Decision Management verwendet wird.

IBM Cognos-Berechtigungsachweis. Beim Importieren von IBM Cognos BI-Datenbankdaten oder einzelnen Listenberichten zur Verwendung in IBM Analytical Decision Management können Sie nun einen in IBM SPSS Collaboration and Deployment Services Repository gespeicherten IBM Cognos-Berechtigungsachweis auswählen, anstatt den Benutzernamen und das Kennwort einzugeben oder eine anonyme Verbindung zu verwenden.

Version 8

IBM SPSS Analytical Decision Management 8 bietet Benutzern erweiterte Möglichkeiten, Analysen in alltäglichen Geschäftsprozessen bereitzustellen, und unterstützt dabei ein breiteres Spektrum von Entscheidungen mit besseren Erkenntnissen und besserer Steuerung als je zuvor.

- Feinabstimmung und Visualisierung der kombinierten Auswirkung von Regeln, Modellen und der Optimierung auf Geschäftsergebnisse unter Verwendung einer erweiterten Simulationsfunktionen
- Lösung eines breiteren Spektrums von Problemen mit neuen vordefinierten Anwendungen und neuen Optimierungsverfahren wie Preisoptimierung und Sortimentoptimierung
- Verbesserte Integration in IBM Operational Decision Management und Big-Data-Unterstützung über Integration von IBM SPSS Analytic Server

Im Folgenden finden Sie eine Beschreibung der wichtigsten neuen Funktionen, die in Version 8 hinzugefügt wurden:

Neue Installationstechnologie. Die Produktinstallation wurde vereinfacht und verbessert. Es wird nun Installation Manager-Technologie verwendet. Weitere Informationen finden Sie im Installationshandbuch zu IBM SPSS Analytical Decision Management.

Erweiterungen der Benutzerschnittstelle. Zur Verbesserung der Benutzerfreundlichkeit wurden zahlreiche Änderungen an der Benutzerschnittstelle vorgenommen.

IBM Analytical Decision Management for Predictive Maintenance. IBM SPSS Analytical Decision Management for Predictive Maintenance ist eine neue vordefinierte Anwendung, die die Optimierung verwendet, um Gerätefehler zu antizipieren und zu verhindern, bevor sie auftreten. Sie können kostenintensive Ausfallzeiten minimieren und eine spärliche Infrastruktur und knappe Ressourcen optimal nutzen, indem Sie Fehler vorhersagen und die richtigen Korrekturmaßnahmen zur richtigen Zeit planen. Möglicherweise können Sie abschätzen, welche Teile zuerst ausfallen und diese rechtzeitig ersetzen. Durch Vorhersagen des Zeitpunkts eines Maschinenausfalls können Profite maximiert und Verluste durch den Ausfall von Maschinen vermieden werden.

IBM Analytical Decision Management for Demand Optimization. IBM Analytical Decision Management for Demand Optimization ist eine neue vordefinierte Anwendung, die die Optimierung verwendet, um kostenintensive Überbestände und Engpässe durch Zuordnen der richtigen Menge der einzelnen Produk-

te für die einzelnen Filialen zu verhindern. Sie können Profite maximieren und Verluste aufgrund von Bestandsengpässen und Gebühren für die Wiederauffüllung des Lagers vermeiden, indem Sie jeder Filiale die richtige Menge Lagerbestände zuordnen.

Simulationserweiterungen. Neue Simulationsfunktionen kombinieren die Funktionalität der vorherigen Schaltfläche **Simulieren** (früher auf der Registerkarte "Definieren") mit der vorherigen Was-wäre-wenn-Funktionalität (früher auf der Registerkarte "Optimieren", "Priorisieren" oder "Kombinieren") in einer einzigen Funktion. Die neue und verbesserte Simulationsfunktion hat eine eigene Symbolleiste in der rechten oberen Ecke der Benutzerschnittstelle und ermöglicht das Speichern von Szenarios mit dem Projekt und das Beibehalten von Szenarios über Sitzungen hinweg.

Bereitstellung und Scoring. Die Registerkarte "Bereitstellen" bzw. "Scoren" wurde um zahlreiche Funktionen erweitert.

Anwendungs-URL. IBM SPSS Collaboration and Deployment Services bietet die Möglichkeit, die URL für webbasierte Anwendungen anzupassen, einschließlich der URL zum Starten von IBM SPSS Analytical Decision Management. Diese Angaben können Sie von Ihrem Administrator erhalten. Weitere Informationen finden Sie im Kapitel *Kontextstammverzeichnisse der Anwendung* des Installations- und Konfigurationshandbuchs zu IBM SPSS Collaboration and Deployment Services Repository.

Erweiterungen für Statistics-Dateien. Komprimierte, kennwortgeschützte IBM SPSS Statistics-Datendateien (.sav und .zsav) werden nun in IBM SPSS Analytical Decision Management unterstützt.

Erweiterungen der Strukturbaumanzeige-funktion. Wenn Sie einzelne Modellergebnisse in IBM SPSS Modeler Advantage anzeigen, sind die Namen der einzelnen Modelle als Link definiert, über den ein neues Dialogfeld geöffnet wird, das Informationen zum Modell enthält, wie z. B. das Zielfeld und die Eingabefelder. Wenn es sich bei dem ausgewählten Modell um ein *Baummodell* handelt, wird zudem eine Strukturbaumanzeige-funktion angezeigt. Die Strukturbaumanzeige-funktion wurde erweitert. Sie können nun Aktionen wie Vergrößern und Verkleinern des Baums, Ändern der Baumausrichtung, Navigieren durch große Bäume und Einblenden oder Ausblenden einzelner Knoten im Baum ausführen.

Erweiterte Optionen auf der Registerkarte 'Optimieren'. Für Anwendungen, die mit komplexer mathematischer Optimierung (CPLEX) arbeiten, wie z. B. IBM SPSS Analytical Decision Management for Campaign Optimization, IBM SPSS Analytical Decision Management for Predictive Maintenance und IBM SPSS Analytical Decision Management for Demand Optimization, sind jetzt erweiterte Optionen verfügbar, die sich auf den CPLEX-Optimierungsalgorithmus beziehen. Sie können nun eine OPL-Modelldatei (OPL - Optimization Programming Language) exportieren und benutzerdefinierte Einstellungen hinzufügen.

Datenansichtsquelle. IBM SPSS Modeler 16 wurde ein neuer *Datenansichtsknoten* hinzugefügt, um Daten, die in einer **Analysedatenansicht** von IBM SPSS Collaboration and Deployment Services definiert wurden, in Streams einzufügen. Die neue Datenquelle **Datenansicht** kann auch auf der Registerkarte **Daten** in IBM SPSS Analytical Decision Management verwendet werden. Sie können eine Verbindung zur Verwendung mit IBM Operational Decision Management-Regeln erstellen. Beachten Sie, dass Administratoren bei komplexen Regeln, die ein hierarchisches Datenmodell verwenden, die Verwendung solcher Regeln einschränken können, da sie zu einer Leistungsver schlechterung führen.

Vor Version 8 erforderten Regeln unstrukturierte Eingabeparameter, die dem Projektdatenmodell von IBM SPSS Analytical Decision Management entsprachen. Mit dieser neuen Unterstützung der Analysedatenansicht können externe Regeln nun jedoch hierarchische Eingabeparameter haben. Weitere Informationen finden Sie in den Themen „Datenansichtsquelle“ auf Seite 41 und „Sperrungen der Datenoptionen“ auf Seite 87 und in Kapitel 5 des Handbuchs zu Application Designer.

Integration von IBM SPSS Analytic Server. In IBM SPSS Modeler 16 wurde IBM SPSS Analytic Server eingeführt und ermöglichte die Verwendung von HDFS (Hadoop Distributed File System) als Datenquelle. IBM SPSS Modeler-Benutzer können mithilfe des HDFS-Quellenknotens von Analytic Server Streams

erstellen und Modelle auf der Grundlage des Hadoop-Big-Data-Systems erstellen. Benutzer von IBM SPSS Analytical Decision Management können diese Modellstreams in IBM SPSS Modeler Advantage zum Anzeigen öffnen (sie können keine neuen Projekte mithilfe von Analytic Server-Datenquellen erstellen).

Neues Kapitel im Handbuch zu Application Designer. Dem Handbuch zu IBM SPSS Analytical Decision Management Application Designer wurde ein neues Kapitel hinzugefügt. Es stellt eine Beispielkonfiguration einer benutzerdefinierten Optimierungsanwendung bereit. Die IBM SPSS Analytical Decision Management for Campaign Optimization-Standardanwendung verwendet eine vorkonfigurierte OPL-Modelldatei (OPL - Optimization Programming Language) mit dem Namen *CampaignOptimization.mod*. In einigen Fällen sind die Standardanwendung Campaign Optimization und ihr OPL-Modell möglicherweise nicht zur Lösung des aktuellen Geschäftsproblems geeignet. Das neue Kapitel führt durch ein Beispiel für ein benutzerdefiniertes OPL-Modell, das mithilfe von IBM ILOG CPLEX Optimization Studio erstellt wurde, und eine benutzerdefinierte IBM SPSS Analytical Decision Management-Optimierungsanwendung für ein fiktives Unternehmen.

Tabellarische Auswertungsergebnisse. Bei der Ausführung von Auswertungen ist die neue Option **Tabellarische Ergebnisse aufnehmen** verfügbar. Durch die Inaktivierung der Option wird die Leistung verbessert, da die Ausführung von Auswertungen, die tabellarische Ergebnisse enthalten, länger dauert.

Neue Einstellung für die maximale Anzahl an nominalen Feldern. Dem Dialogfeld "Eigenschaften" wurde eine neue Einstellung zum Festlegen der maximalen Anzahl an Mitgliedern hinzugefügt, die für *nominale* Datenfelder zulässig sind. Dies wird auf der Registerkarte "Daten" verwendet. Weitere Informationen finden Sie in „Eigenschaften“ auf Seite 33.

Single Sign-on. ODBC-Verbindungen können nun Single Sign-on für IBM SPSS Modeler Server- und ODBC-Datenbankverbindungen mit IBM SPSS Analytical Decision Management verwenden. Es sind einige Konfigurationsschritte erforderlich. Weitere Informationen finden Sie im Thema Überlegungen zur Sicherheit in der Installationsdokumentation.

SQL-Pushback. Wenn die Quelldaten aus einer Datenbank stammen, wird die Ausführung an die Datenbank zurückgegeben, wann immer dies möglich ist (in einem Prozess, der als "SQL-Pushback" bezeichnet wird). Im Allgemeinen unterstützt IBM SPSS Analytical Decision Management dieselbe SQL-Pushback-Stufe wie IBM SPSS Modeler, jedoch mit den folgenden Einschränkungen:

- Wenn das Projekt zusätzliche Felder oder Auswahlregeln enthält, die Referenzen auf Modelle oder externe Regeln aufweisen, werden diese Referenzen nicht an die Datenbank zurückgegeben und dadurch kann verhindert werden, dass der für den Projektstream als Ganzes erreichbare Pushback-Grad erreicht wird.
- In anderen Anwendungen als IBM SPSS Modeler Advantage wird der Pushback nur bis zu Auswahlregeln (globalen Auswahlen und Mitgliedsauswahlen in der ersten Dimension) aktiviert. Der Pushback-Grad kann erhöht werden, wenn Sie ein Feld für ganze Zahlen als Entitäts-ID im Projektdatenmodell auswählen.

Version 7.0.0.1

In IBM SPSS Decision Management 7 Fixpack 1 (7.0.0.1) wurden die folgenden Funktionen hinzugefügt:

Datenübersicht. Auf der Registerkarte "Daten" aller Anwendungen, auf den Registerkarten "Modellierung" und "Scoren" in IBM SPSS Modeler Advantage sowie auf der Registerkarte "Bereitstellen" von Anwendungen, die die Funktion *Jetzt scoren* enthalten, ist eine neue Funktion zur Datenübersicht verfügbar. Weitere Informationen finden Sie im Thema „Datenübersicht“ auf Seite 45. Diese Funktion ist nur mit IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1) verfügbar. Sie ist nicht mit Version 15 verfügbar.

Scoring-Zusammenfassungsbericht. Nach dem Scoren eines Projekts werden in dem neuen Abschnitt *Scoring-Zusammenfassungsbericht* Details angegeben: die Position, an die die Daten gescort wurden, die Dateigröße (falls verfügbar), der Datenbank- und Tabellename (falls verfügbar), der Modell- oder Projektname, die Anzahl der gescorten Datensätze und in einigen Fällen ein Diagramm und/oder eine Tabelle mit einer Zusammenfassung des Modells. Der Abschnitt "Scoring-Zusammenfassungsbericht" ist in

IBM SPSS Modeler Advantage sowie auf der Registerkarte "Bereitstellen" von Anwendungen verfügbar, die die Funktion "Jetzt scoren" umfassen. Hinweis: Die Dateigröße wird nur bei Ausführung unter IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1) angezeigt. Sie ist nicht mit Version 15 verfügbar.

Lokale Kopie einer gescorten Datei herunterladen. Wenn Sie die Scoring-Ausgabe an eine Flatfile senden möchten, können Sie bei Bedarf auch eine lokale Kopie der Datei herunterladen. Administratoren können die browserbasierte Komponente IBM SPSS Collaboration and Deployment Services Deployment Manager verwenden, um die Funktion zu aktivieren bzw. zu inaktivieren. Ferner können Sie damit die maximal zulässige Dateigröße erhöhen oder verringern. Wenn die Schaltfläche **Datei herunterladen** inaktiv ist, überschreitet die Dateigröße die vom Administrator festgelegte maximale Größe oder der Administrator hat die Funktion inaktiviert. Die Funktion ist standardmäßig inaktiviert. Diese Funktion ist nur mit IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1) verfügbar. Sie ist nicht mit Version 15 verfügbar.

Neue optionale Einstellungen zum Erstellen von Modellen. Wenn Sie Modelle in IBM SPSS Modeler Advantage erstellen, stehen Ihnen neue optionale Einstellungen zur Verfügung. Hierzu zählen Optionen zur Datenpartitionierung, das Auswählen von Modellierungsverfahren (Algorithmen), die beim Erstellen des Modells einbezogen werden sollen, die Angabe der maximalen Anzahl von Modellen, die gespeichert und kombiniert werden sollen, sowie das Festlegen von Profitkriterien vor der Erstellung des Modells. Die verfügbaren Optionen hängen vom Modelltyp und dem ausgewählten Ziel ab. Weitere Informationen finden Sie im Thema „Optionale Modelleinstellungen“ auf Seite 65. Hinweis: Die folgenden neuen Funktionen sind nur mit IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1) verfügbar. Sie sind nicht mit Version 15 verfügbar.

- Auswählen der Modellierungsverfahren, die beim Erstellen des Modells einbezogen werden sollen
- Festlegen von Profitkriterien
- Anzeigen allgemeiner Details zu einzelnen Modellen
- Anzeigen zusätzlicher Details zu einem ausgewählten Modell

Strukturbaumanzeigefunktion. Nach dem Erstellen eines Modells in IBM SPSS Modeler Advantage und dem Anzeigen einzelner Modellergebnisse sind die einzelnen Modelle als Link definiert, der in einem neuen Dialogfeld geöffnet wird. Dieses enthält Informationen zum Modell, wie z. B. das Zielfeld und die Eingabefelder. Wenn es sich bei dem ausgewählten Modell um ein *Baummodell* handelt, wird jetzt zudem eine Strukturbaumanzeigefunktion angezeigt. Hinweis: Die Strukturbaumanzeigefunktion ist nur mit IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1) verfügbar. Sie ist nicht mit Version 15 verfügbar.

Auswertungserweiterungen. Das Dialogfeld "Auswerten" in IBM SPSS Modeler Advantage wurde um zahlreiche Funktionen erweitert. Beispielsweise können Sie mit der neuen Schaltfläche **Modellauswahl ändern** die Modelle so ändern, dass sie in die Auswertung einbezogen werden. Wenn Sie Ihre Auswahl getroffen haben, können Sie die aktuelle Modellauswahl bei Bedarf zudem wieder in die Tabelle mit den einzelnen Modellergebnissen kopieren. Einige der neuen Optionen sind nur für bestimmte Modelltypen verfügbar (die meisten neuen Optionen sind nur für den Modelltyp *Vorhersage* verfügbar, einige aber auch für den Modelltyp *Vorhersageregulierung*). Hinweis: Die folgenden neuen Funktionen im Dialogfeld "Auswerten" sind nur mit IBM SPSS Modeler Server 15 Fixpack 1 (15.0.0.1) verfügbar. Sie sind nicht mit Version 15 verfügbar.

- Leistungsdiagramme für ein numerisches Ziel
- Profitsimulation für ein numerisches Ziel
- Fehlklassifizierungstabellendiagramme (für alle Ziele)
- Tabellenansicht aller Diagramme
- Schaltfläche **Modellauswahl ändern**
- Option **Einzelne Modellevaluierungen für Leistungsdiagramme einschließen**

Microsoft Windows 8 und Internet Explorer 10. IBM SPSS Collaboration and Deployment Services 5 Fixpack 1 (5.0.0.1) bietet Unterstützung für Windows 8 und Internet Explorer 10. Dies wird mit IBM SPSS Collaboration and Deployment Services 4.2.1 oder 5 nicht unterstützt.

Erweiterungen der Benutzerschnittstelle. Zur Verbesserung der Benutzerfreundlichkeit wurden verschiedene geringfügige Änderungen an der Benutzerschnittstelle vorgenommen.

Version 7

In IBM SPSS Decision Management 7 (7.0.0.0) wurden die folgenden Funktionen hinzugefügt:

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 Marketingkampagne bzw. die Minimierung des Betrugs- oder 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.

Clustering-Modelle. Neu in IBM SPSS Modeler Advantage. Mit Methoden zur Clustermodellierung 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 verwenden und keine bestimmte Vorhersage liefern, die als "wahr" oder "falsch" ausgewertet werden kann. Mit IBM SPSS Decision Management können Sie Clustermodelle automatisch erstellen. Sie können auch Cluster kombinieren, umbenennen, bearbeiten oder löschen oder benutzerdefinierte Cluster auf der Grundlage bereits vorhandenen Fachwissens 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 Anwendungsansicht, 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. Ebenso wie IBM SPSS Modeler 15 unterstützt auch IBM SPSS Decision Management jetzt den Import von in Cognos vordefinierten 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 ein verfügbares Element (Hierarchie) auszuwählen, das in der Anwendung verwendet werden soll, 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 früheren Releases musste ein Objekt im Dimensionsbaum vorhanden sein, damit es zugeordnet werden konnte. 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.

Mehrere Zuordnungen. 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.

Verbesserungen der Benutzerschnittstelle. Viele der Änderungen und Neuerungen für die IBM SPSS Decision Management-Benutzerschnittstelle wurden eingeführt, um die Benutzerfreundlichkeit und Skalierbarkeit aller Anwendungen zu verbessern.

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

Regeleditorerweiterungen. Je nachdem, von wo der Regeleditor gestartet wird, können Benutzer jetzt den Regeltyp für Auswahlregeln festlegen (einschließend oder ausschließend), sie können Aggregationsregeln Risikopunkte zuweisen, eine Dimension für Zuordnungsregeln auswählen sowie die Operatoren *ODER* und *ODER aufteilen* 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 von IBM Operational Decision Management. In IBM SPSS Collaboration and Deployment Services Deployment Manager können Sie mithilfe des Dialogfelds "Neue externe Regel hinzufügen" Verweise auf IBM Operational Decision Management-Regelsets definieren, die anschließend in IBM SPSS Decision Management verwendet werden können. Bisher waren diese Regeln nur per Web-Service zugänglich. Mit diesem Release können Sie nun auf lokale Regeln in der Datenbank von IBM Operational Decision Management verweisen. Für die neue Option "Lokal" wird die Regel über den IBM SPSS Modeler Server ausgeführt (derselbe Server, der in Deployment Manager für die Verwendung durch IBM SPSS Decision Management definiert wurde).

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

Streamdateien hochladen. Sie können ab sofort IBM SPSS Modeler-Streamdateien aus dem lokalen Dateisystem öffnen und bei Bedarf im Repository speichern.

Mehrere Berechtigungsnachweise für IBM SPSS Modeler Server. Wenn in IBM SPSS Collaboration and Deployment Services Deployment Manager mehrere Berechtigungsnachweise für IBM SPSS Modeler Server definiert sind, können Sie nun auch mehrere Berechtigungsnachweise zur Verwendung mit IBM SPSS Decision Management angeben. Wenn sich anschließend der IBM SPSS Decision Management-Benutzer anmeldet, prüft der Server anhand des ersten Berechtigungsnachweises in der Liste, ob der Benutzer berechtigt ist, und arbeitet die Listeneinträge nacheinander ab, bis ein Berechtigungsnachweis mit der entsprechenden Berechtigung gefunden wird. Wenn für diesen Benutzer kein entsprechender Berechtigungsnachweis in der geordneten Liste vorhanden ist, kann der Benutzer IBM SPSS Decision Management nicht verwenden. Somit kann genau geregelt werden, welche Benutzer Zugang zu welchen Teilen des Dateisystems von IBM SPSS Modeler Server haben.

Version 6.2

In IBM SPSS Decision Management 6.2 sind folgende neue Funktionen enthalten:

Unterstützung für Cognos-Daten. Bei Verwendung mit IBM SPSS Modeler Server 14.2 oder neuer bietet IBM SPSS Decision Management 6.2 die Möglichkeit, beim Definieren von Datenquellen Daten von einem IBM Cognos Business Intelligence-Server zu lesen und Scoring-Ergebnisse wieder auf einen 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. Weitere Informationen finden Sie im Thema „IBM Cognos BI-Quelle“ auf Seite 42.

Integration mit dem ILOG Business Rule Management System. In einem Business Rules Management System wie ILOG erstellte Regeln können in IBM SPSS Decision Management-Anwendungen referenziert und verwendet werden, vorausgesetzt, sie wurden so entwickelt, dass sie dasselbe Datenmodell unter-

stützen, das auch im aktuellen IBM SPSS Decision Management-Projekt verwendet wird. Diese externen Regeln können ausgewählt und auf dieselbe Weise verwendet werden wie andere Modelle und Regeln in der IBM SPSS Decision Management-Anwendung. Weitere Informationen finden Sie im Thema „Verwenden von Regeln aus IBM Operational Decision Management“ auf Seite 58. Hinweis: ILOG wird jetzt als IBM Operational Decision Management bezeichnet.

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

Kapitel 2. Beispiel für die Modellierung der Kundenreaktion

IBM SPSS Modeler Advantage

IBM SPSS Modeler Advantage ist eine benutzerfreundliche Anwendung, die Fachanwendern das ganze Potenzial der Vorhersagemodellierung 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 verwendet werden, um Listen von Zielkunden oder relevanten Fällen zu erstellen, die als Ausgangsmaterial für die strategische Planung oder im Rahmen einer Vorhersageanwendung zusammen mit Regeln eingesetzt werden können.

Maximieren des Profits für Direktwerbungskampagnen

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 Profitschwelle.

Ihr Ziel ist es, diesen Verlust mit IBM SPSS Modeler Advantage in einen Profit 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 1. Ergebnisse der Testkampagne

Ergebnisse der Testkampagne	
Anzahl der Anschreiben	3000
Anzahl der Antworten	111
Rücklaufquote	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 fertiggestellte Version des Beispiels steht mit der Datei *bank_response_model.str* zur Verfügung. Genaueres zu Installation und Beispieldateien können Sie gegebenenfalls von Ihrem Administrator erfahren. Weitere Informationen finden Sie im Thema „Beispieldateien“.

Beispieldateien

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

Daten

Während der Installation werden zahlreiche Beispieldaten im Verzeichnis IBM SPSS Modeler Server\
Demos installiert (z. B. *C:\Programme\IBM\SPSS\ModelerServer\18\Demos*). Diese Daten können mit den vordefinierten Beispielanwendungen verwendet werden. Es handelt sich dabei um die folgenden Dateien:

- *bank_customer_data.txt*
- *bank_response_data.txt*
- *customer_contact_history.txt*
- *customer_summary_data.txt*
- *insurance_claim_data.txt*
- *insurance_fraud_data.txt*
- *inventory_orders.txt*
- *inventory_stock.txt*
- *machine_log_data.txt*
- *machine_summary_data.txt*
- *retail_purchase_data.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 Datenquellenpfade in den Beispielanwendungen entsprechend aktualisiert werden.

Beispiele

Es wird eine Reihe von Beispielanwendungsdateien, Modellen und Regeln für die Verwendung mit den vordefinierten Anwendungen von IBM Analytical Decision Management bereitgestellt. Die Beispiele werden in einer für den elektronischen Download verfügbaren Datei bereitgestellt. Gehen Sie wie folgt vor, um die Beispieldateien für eine oder mehrere vordefinierte Anwendungen in die Instanz von IBM SPSS Collaboration and Deployment Services Repository zu importieren.

1. Laden Sie die Datei *ADM_18_Demo.zip* an einen temporären Speicherort herunter und extrahieren Sie sie.
2. Starten Sie IBM SPSS Collaboration and Deployment Services Deployment Manager und melden sich bei Ihrem Content-Server an.
3. Klicken Sie mit der rechten Maustaste auf den Ordner für das Inhalts-Repository und wählen Sie die Option **Importieren** aus.
4. Klicken Sie im Dialogfeld **Importieren** auf **Durchsuchen** und wählen Sie die Datei *DecisionManagementExamples.pes* aus, die Sie in Schritt 1 an einen temporären Speicherort kopiert oder extrahiert haben. Klicken Sie auf **OK**.
5. 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.

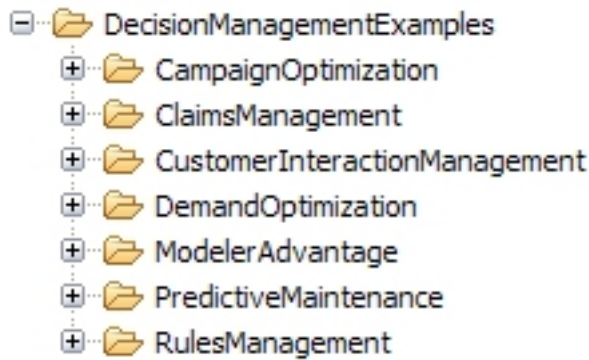


Abbildung 1. Ordner für Beispielartefakte

Beachten Sie, dass alle Beispieldateien für Anwendungen, Modelle und Regeln voraussetzen, dass im Verzeichnis *Demos* von IBM SPSS Modeler Server Datendateien installiert sind (z. B. *C:\Programme\IBM\SPSS\ModelerServer\18\Demos*). Wenn die Datendateien an einem anderen Speicherort installiert wurden, müssen die Datenquellenpfade in den Beispielanwendungen für die lokale Konfiguration entsprechend aktualisiert werden.

Starten der Anwendung

Die Startseite von IBM Analytical Decision Management 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.

1. Öffnen Sie die Startseite von IBM Analytical Decision Management, 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.
2. Zur Erstellung eines neuen Modells wählen Sie aus dem Dropdown-Menü der Anwendung in der Gruppe IBM SPSS Modeler Advantage die Option **Neu** aus und klicken auf **Wechseln**. Die Homepage von IBM SPSS Modeler Advantage wird angezeigt.



Abbildung 2. IBM SPSS Modeler Advantage - Homepage

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.

Einrichten von Datenquellen

Auf der Registerkarte "Daten" können Sie Datenquellen für die Beispieldatendateien aus dem Beispiel konfigurieren. Datasets können von Ihrem Administrator konfiguriert oder von Ihnen selbst hinzugefügt werden.

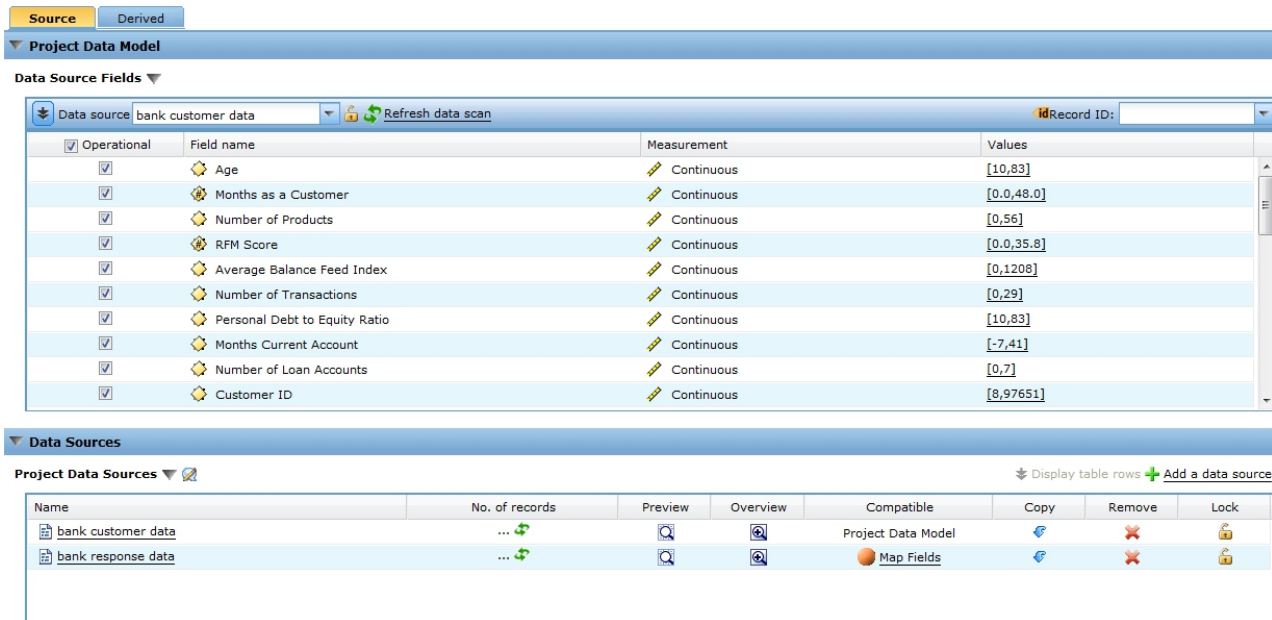


Abbildung 3. Registerkarte "Daten"

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 Beispieldatendateien im IBM SPSS Modeler Server-Ordner *Demos* abgelegt, auf den Sie über Ihre Anwendung zugreifen können (z. B. *C:\Programme\IBM\SPSS\ModelerServer\18\Demos*). Wenden Sie sich an den Administrator, falls Sie Hilfe beim Installieren dieser Dateien benötigen.

So definieren Sie das Projektdatenmodell:

1. Wählen Sie im Bereich "Projektdatenquellen" in der Dropdown-Liste **Datenquelle** die Option **Datenquelle hinzufügen** aus.
2. Geben Sie *bank customer data* als Datenquellennamen an.
3. Wählen Sie **Datei** und dann *bank_customer_data.txt*.
4. Achten Sie darauf, dass das Messniveau des Felds *Has Bad Payment Record* unter **Eingabefelder festlegen** auf **Flag** eingestellt ist. Die Einstellung "Flag" bedeutet, dass das Feld zwei mögliche Werte haben kann, die für *wahr* (True) und *falsch* (False) -stehen (oft als *T* und *F* oder in numerischen Feldern als 0 und 1 kodiert) und das Vorliegen bzw. Nichtvorliegen des jeweiligen Attributs ausdrücken.
5. Achten Sie außerdem darauf, dass das Messniveau des Felds *Gender* 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.)
6. Klicken Sie auf **Links**. Das Dialogfeld **Links** wird angezeigt.
7. 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 Datensets werden relativ zu dieser Quelle abgebildet.

So definieren Sie die Antwortdatenquelle:

Das Antwortdataset enthält weitgehend dieselben Felder, die auch im Projektdatenmodell erscheinen, aber zwei zusätzliche Felder mit der Bezeichnung *Campaign ID* und *Response*. 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.

1. Wählen Sie im Bereich "Projektdatenquellen" in der Dropdown-Liste **Datenquelle** die Option **Datenquelle hinzufügen** aus.
2. Wählen Sie *bank response data* als Datenquellennamen.
3. Wählen Sie **Datei** und dann *bank_response_data.txt*.
4. Achten Sie darauf, dass das Messniveau des Felds *Response* unter **Eingabefelder festlegen** auf **Flag** eingestellt ist.
5. Heben Sie unter **Eingabefelder festlegen** die Auswahl des Kontrollkästchens neben dem Feld *Campaign ID* auf, 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.)
6. Klicken Sie auf **Speichern**.

Da einige der Felder des Datenmodells im Antwortdataset fehlen, wird für diese Datenquelle in der Spalte **Kompatibel** ein orangefarbenes Symbol angezeigt, das darauf hinweist, dass manche, aber nicht alle Felder übereinstimmen. Das Dataset 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 "Modell" können Sie Datenquelle, Zielfeld und andere Optionen für die Erstellung und Auswertung des Modells festlegen.

Model type
 Predictive model Change Model...

Data source
 bank response data + +

Records used to build: ... ↻

Target 🎯
 Response

Evaluate
Test

Optional Settings ▶

▶ Build Model

▼ Predictive Model Results

Display: Combined model results Individual model results

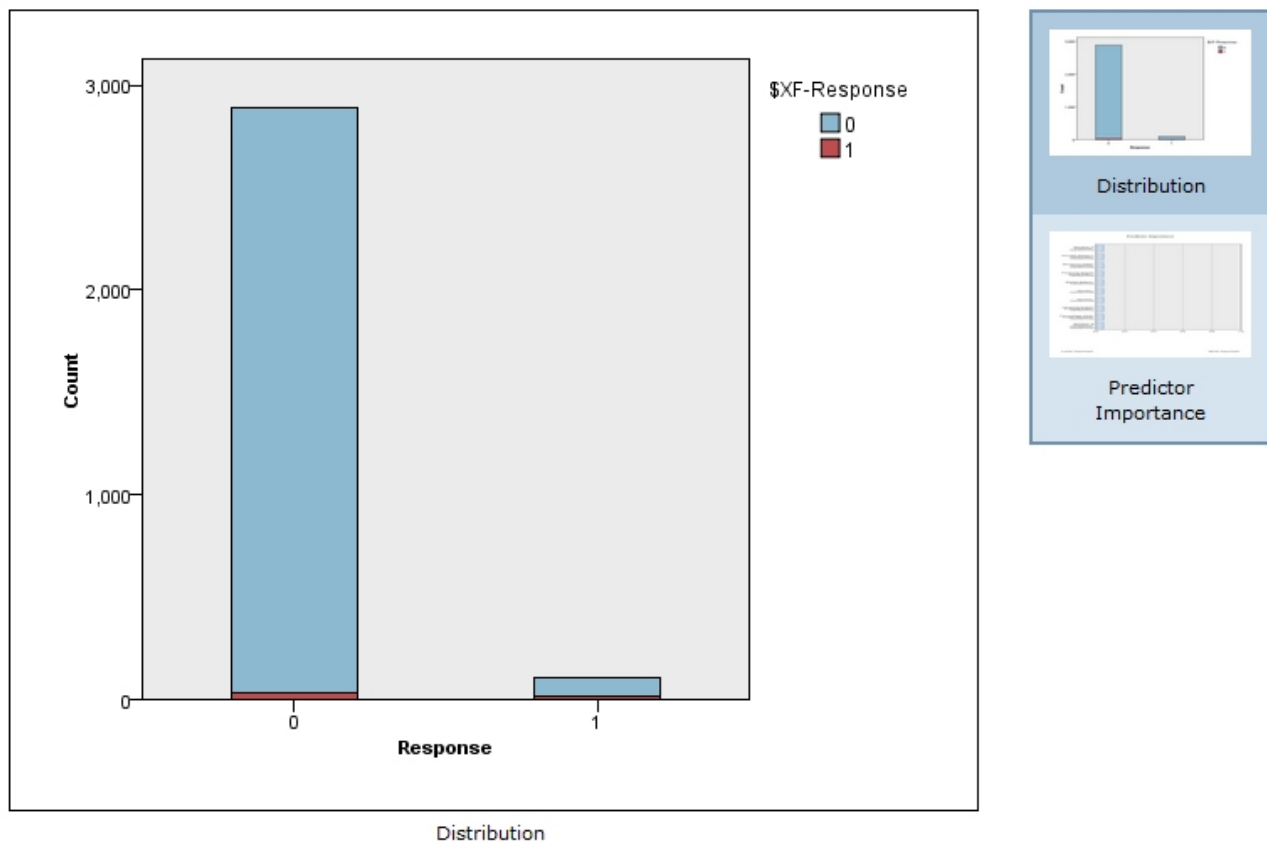


Abbildung 4. Modell (Registerkarte)

1. Wählen Sie *bank response data* als Datenquelle und *Response* als Zielfeld.
2. Wenn die Nachricht **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 Gesamttrücklaufquote aus. Von den 3000 Kunden des Datasets 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 Gesamttrücklaufquote gibt es relativ wenige vorhergesagte

Antworten (rote Bereiche). Der Balken für den Wert "falsch" (*Response = 0*) ist jedoch überwiegend blau, was bedeutet, dass der größte Teil derjenigen, für die keine Antwort vorhergesagt wurde, auch tatsächlich nicht geantwortet haben.

Um einen besseren Eindruck davon zu bekommen, wie das Modell funktioniert, fahren Sie mit dem nächsten Abschnitt („Auswerten des Modells“) fort.

Auswerten 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.

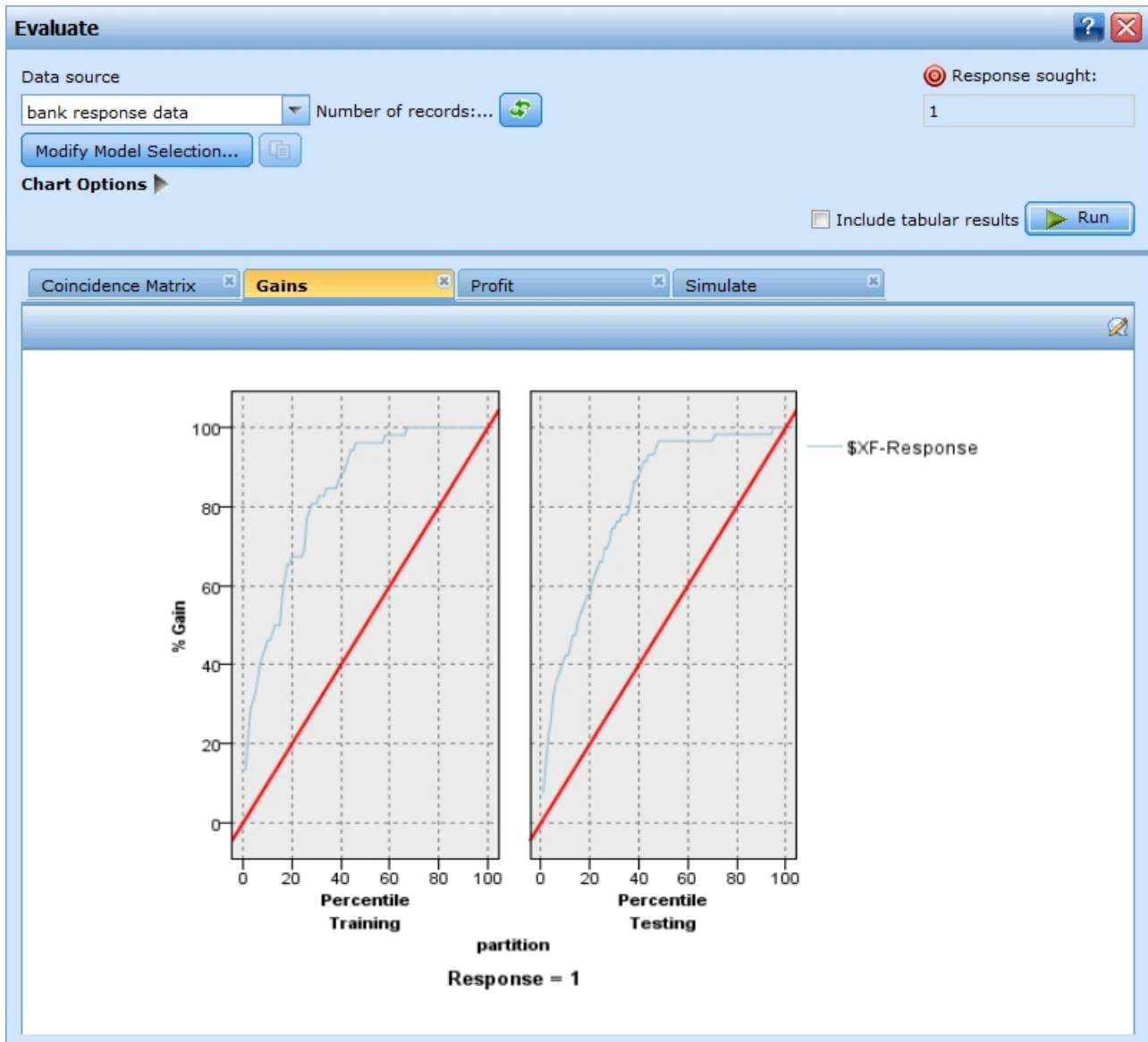


Abbildung 5. Auswerten des Modells

1. Wählen Sie *bank response data* als Datenquelle.
2. Wählen Sie als zu erstellende Diagrammtypen **Gewinndiagramm** und **Profitdiagramm**.
3. Wählen Sie **Maximalen Profit simulieren** und legen Sie als Populationsgröße *3000* fest. Dies entspricht der Zahl aus der Testkampagne.

4. 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.
5. Wählen Sie als **Gesuchte Antwort** den Wert 1 aus.
6. Klicken Sie auf **Ausführen**.

Das Gewinnendiagramm zeigt die Verbesserung (den "Lift"), die durch Einsatz des Modells erreicht werden kann (siehe die vorherige Abbildung). 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.

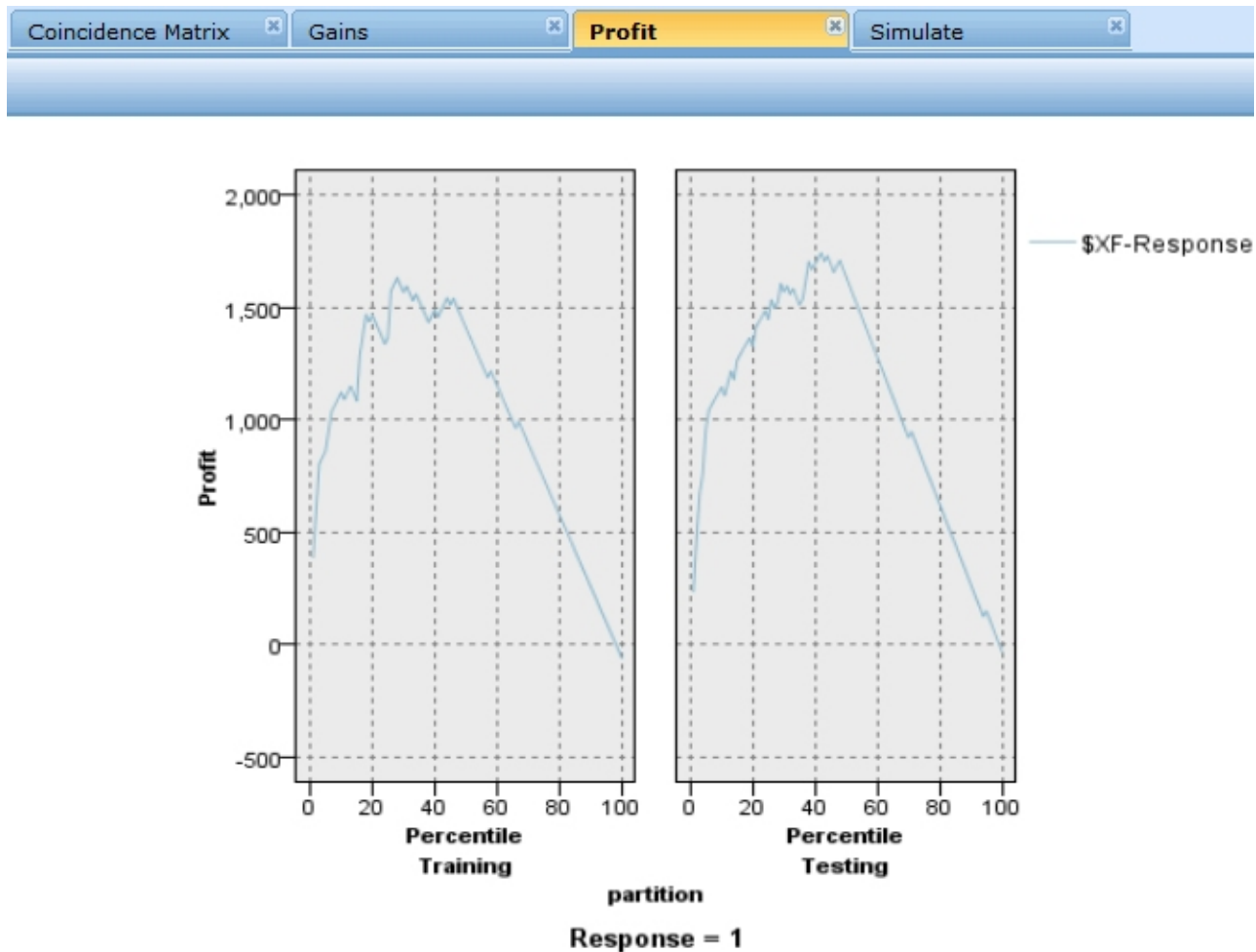


Abbildung 6. Profitdiagramm

Ein konkretes Beispiel erhalten Sie, indem Sie auf der Registerkarte "Simulieren" den erwarteten maximalen Profit für die von Ihnen festgelegte Populationsgröße abrufen.

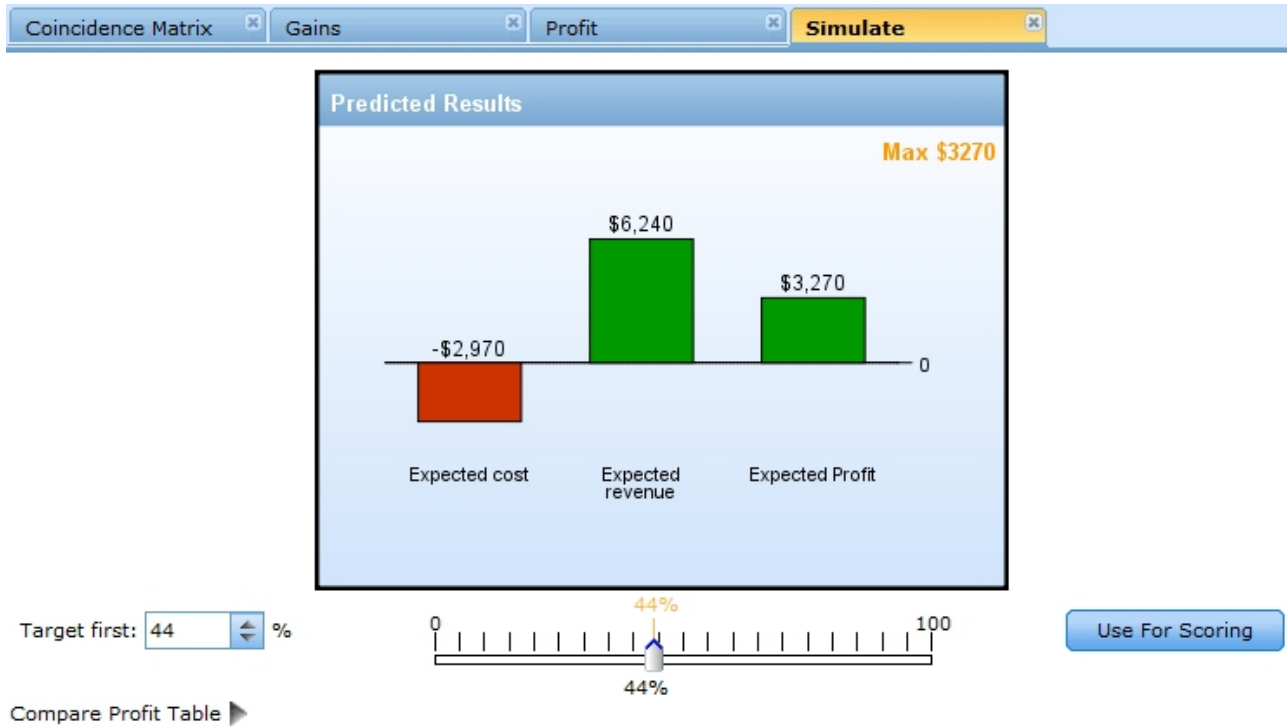


Abbildung 7. Profitsimulation

Der erwartete Profit wird mit der folgenden Formel ermittelt:
 erwarteter Profit = (Antwortwahrscheinlichkeit * Umsatz) - 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 Rücklaufquoten aufgenommen werden. Die Simulation erleichtert Ihnen die Entscheidung, wo die Grenze gezogen werden sollte, um den höchstmöglichen Profit zu realisieren.

Compare Profit Table ▼

Export

	Targeted 44%	Random 44%	Everyone
Offers made	1320	1320	3000
Expected response	104	48.8	111
Expected response(%)	7.9	3.7	3.7
Expected cost	\$2970	\$2970	\$6750
Expected revenue	\$6240	\$2930.4	\$6660
Expected Profit	\$3270	-\$39.6	-\$90

Abbildung 8. Profitvergleich

In der Tabelle "Profitvergleich" wird dieses Ergebnis ausführlicher zusammengefasst.

- **Zielgruppe 44 %.** In dieser Spalte wird das Ergebnis aus der gezielten Ansprache der 44 % Kunden mit der vom Modell vorhergesagten höchsten Rücklaufquote angezeigt. Durch selektive Aufnahme der 1320 Kunden mit der höchsten Antwortwahrscheinlichkeit können Sie Ihren Ertrag maximieren und einen erwarteten Profit von 3270 US-Dollar erzielen.
- **Stichprobe 44 %.** Zeigt das Ergebnis an, wenn Sie eine zufällige Stichprobe derselben Größe ziehen würden. In diesem Fall würden Sie dieselbe Anzahl an Angeboten (1320) versenden, doch die erwartete Rücklaufquote fiel 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 Gesamtrücklaufquote 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 3277,50 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.

7. Wählen Sie nach der Überprüfung der Simulationsergebnisse unter dem Diagramm "Vorhergesagte Ergebnisse" die Option **Zum Scoren verwenden** aus, um die Registerkarte "Scoren" auf der Grundlage des angegebenen Prozentsatzes (in diesem Fall 44 %) zu aktualisieren. Weitere Informationen finden Sie im Thema „Scoren des Modells“ auf Seite 21.
8. Speichern Sie das Projekt bei Bedarf als *bank_response_model.str*. Im nächsten Abschnitt ändern Sie den Modelltyp und erstellen ein neues Projekt.

Vorhersageregemodelle

Vorhersageregemodelle bieten zwar meist die gewünschte Vorhersagekraft, doch in manchen Fällen sind möglicherweise etwas mehr Kontrolle und Einblick in die zugrunde liegende Logik des Modells erforderlich. In diesem Fall können Vorhersageregemodelle (interaktiv) verwendet werden. Mit Vorhersageregemodellen (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.

1. Kehren Sie zur Registerkarte "Modellierung" zurück, klicken Sie auf **Modell ändern** und wählen Sie die Option **Vorhersageregemodell** aus.
2. Wählen Sie *bank response data* als Datenquelle und *Response* als Zielfeld.
3. Legen Sie **1** als die gesuchte Antwort fest, d. h., Sie suchen Kunden, die eine positive Antwort gegeben haben.
4. Wählen Sie **Erstellung beginnen**.
5. 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.
6. Klicken Sie auf das Symbol **Einschließen-** bzw. **Ausschließen**, um festzulegen ob die von den einzelnen Regeln ausgewählten Kunden eingeschlossen werden sollen oder nicht. Das zweite und dritte Regelsegment beispielsweise zeigen eine Rücklaufquote 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:
7. Setzen Sie alle drei Regelsegmente auf **Einschließen** und klicken Sie auf **Auswerten**.
8. Wählen Sie als zu erstellende Diagrammtypen **Gewinndiagramm** und **Profitdiagramm**.

9. 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.
10. Klicken Sie auf **Ausführen** und öffnen Sie die Registerkarte "Simulieren", wenn die Diagramme erstellt wurden.
Der maximale simulierte Profit für dieses Modell beträgt 3178,90 US-Dollar, etwas weniger als für das automatisierte 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 Vorhersageregelmotellen (interaktiv) sind ihre erhöhte Kontrolle und die tieferen Einblicke, die sie bieten.
11. Speichern Sie das Projekt bei Bedarf als *bank_predictive_rule_model.str*.

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.

1. Wählen Sie **Testen** auf der Registerkarte "Modellierung", um eine Vorschau aufzurufen, die zeigt, wie bestimmte Datensätze beim Modellscoring verarbeitet werden. Das Dialogfeld "Testen" öffnet sich.
2. Erhöhen Sie im Bereich **Datensätze** die Rückgabezahl auf 100.
3. Wählen Sie im Bereich **Anzuzeigende Felder auswählen** *Customer ID* aus. Darüber hinaus werden die Felder *\$XF-Response* und *\$XFRP-Response* eingeschlossen, da diese die vorhergesagte Antwort (in diesem Fall 0 für "nein", 1 für "ja") und den Propensity-Score für jeden Kunden angeben.

Sie können wahlweise weitere Felder einschließen, wie *Age* (Alter), *Weeks Since Last Offer* (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.

1. Wechseln Sie zur Registerkarte "Scoren".
2. Stellen Sie bei **Zu scorende Datenquelle** sicher, dass *bank response data* ausgewählt ist.
3. Geben Sie ein Ziel für die Scoring-Ergebnisse an. (Wenn Sie lediglich eine Vorschau der Scores anzeigen möchten, können Sie diesen Schritt überspringen.)
4. Stellen Sie sicher, dass im Bereich **Ausgabefelder** das Feld *Customer ID* ausgewählt ist.
5. Die Einstellungen für **Scoren** sollten bereits so konfiguriert sein, dass das Scoring für denselben Prozentsatz von Datensätzen durchgeführt wird wie in der bei der Auswertung des Modells durchgeführten Profitsimulation. (Wenn Sie bei der Anzeige der Simulationsergebnisse nicht die Schaltfläche **Zum Scoren verwenden** gewählt haben, können Sie die Einstellungen hier aktualisieren.)
6. Klicken Sie unten im Dialogfeld auf **Vorschau**, um ein Muster der Scoring-Ergebnisse anzuzeigen.
7. Klicken Sie auf **Datenübersicht**, um eine Übersicht der Daten anzuzeigen, die bei Ihrem Scoring verwendet werden. Weitere Informationen finden Sie im Thema „Datenübersicht“ auf Seite 45.

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 Mailing-Liste generieren, die auf diesem Ergebnis beruht.

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

Beispiel für Clustermodellierung

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 fertiggestellte Version des Beispiels steht mit der Datei *bank_cluster_model.str* zur Verfügung. Genaueres zur Installation und Beispieldateien können Sie gegebenenfalls von Ihrem Administrator erfahren. Weitere Informationen finden Sie im Thema „Beispieldateien“ auf Seite 11.

1. Kehren Sie zur Anwendungsstartseite zurück und erstellen Sie ein neues IBM SPSS Modeler Advantage-Projekt.
2. Fügen Sie auf der Registerkarte "Daten" im Feld "Projektdatenquellen" eine neue Datenquelle mit der Bezeichnung *bank customer data* (Bankkundendaten) hinzu, wie in „Einrichten von Datenquellen“ auf Seite 13 beschrieben.
3. Klicken Sie auf der Registerkarte "Modellierung" auf **Modell ändern**, wählen Sie die Option **Clustermodell** und klicken Sie auf **Speichern**.
4. Wählen Sie als Datenquelle *bank customer data*.
5. Erweitern Sie den Abschnitt **Optionale Einstellungen**. Bei Clustermodellen stehen zusätzlich optionale Einstellungen unter **Optionen für automatische Cluster** zur Verfügung, um ein Evaluierungsfeld 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 Clustermodell ausgewählt. Diese beiden Einstellungen für "Optionen für automatische Cluster" bieten also gewisse Einflussmöglichkeiten darauf, welches der Modelle tatsächlich ausgewählt wird. Wenn hier ein Evaluierungsfeld angegeben wird, wird das Modell ausgewählt, das die Werte des Evaluierungsfelds am besten differenziert. Wenn der gewünschte Bereich festgelegt wird, werden alle Modelle, die Cluster außerhalb des angegebenen Bereichs finden, verworfen.

6. 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 Clustermodell gewährleistet, das nicht eine zu hohe oder zu geringe Anzahl an Clustern aufweist.
7. 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.
8. 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:

9. 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)
10. Erstellen Sie eine weitere neue Regel namens *young single customers* (junge, alleinstehende Kunden) mit folgenden Ausdrücken.

Age < 35 (Alter)

Marital Status = U (Familienstand)

11. 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 "wertvolle Kunden" und 33 Kunden vom Typ "junge, alleinstehende Kunden" gefunden wurden.
12. 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.
13. 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 Evaluierung und Test des Modells fort, und wenn Sie damit zufrieden sind, führen Sie abschließend ein Scoring durch.

14. Speichern Sie das Projekt bei Bedarf als *bank_cluster_model.str*.

Weitere Details zur Clustermodellierung finden Sie in „Erstellen eines Clustermodells“ auf Seite 63.

Beispiel für Assoziationsmodellierung

Assoziationsmodellierung ist eine wichtige Technik in der Vorhersageanalyse. 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 gegebenenfalls von Ihrem Administrator erfahren. Weitere Informationen finden Sie im Thema „Beispieldateien“ auf Seite 11.

1. Kehren Sie zur Anwendungsstartseite zurück und erstellen Sie ein neues IBM SPSS Modeler Advantage-Projekt.
2. 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*.
3. Klicken Sie auf der Registerkarte "Modellierung" auf **Modell ändern**, wählen Sie die Option **Assoziationsmodell** und klicken Sie auf **Speichern**.
4. Wählen Sie als Datenquelle *retail purchase data* (Einzelhandelsdaten).

5. Wählen Sie als Datenformat die Option **Tabellenform**. In einem Transaktionsdataset würden nur Elemente angezeigt, die eingeschlossen wurden, und alle Elemente würden in einer einzigen Spalte angezeigt. In diesem Dataset 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.

6. 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.

7. Erweitern Sie den Abschnitt "Optionale Einstellungen". Ändern Sie unter "Optionen für Assoziationsregeln" den ersten Schwellenwert von 10 in 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.

8. 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 **Abdeckungsbereich (%)** zeigt an, dass ca. 15 % der Kunden (also etwa 1 von 6) das Smartphone gekauft haben, und die Spalte **Konfidenz (%)** zeigt, dass diese Kunden mit ca. 30 % Wahrscheinlichkeit auch einen Großbildfernseher 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.

9. 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.

10. Klicken Sie beispielsweise auf das Vorschau-Symbol neben dem ersten Datensatz, um zu sehen, dass der erste Kunde einen Großbildfernseher 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 0,3 in der dritten Spalte bedeutet, dass eine Wahrscheinlichkeit von 30 % besteht, dass Kunden, die bereits einen Großbildfernseher und Lautsprecher erworben haben, auch einen Standardfernseher 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.

11. Schließen Sie das Dialogfeld "Testen".

12. Blättern Sie durch die Modellergebnisse, bis Sie Regel Nummer 115 sehen. Beachten Sie, dass Sie bei großen Modellen, die eine große Anzahl Ergebnisseiten aufweisen, auch auf **Regel anhand der ID suchen** klicken und die Regelnummer eingeben können.

13. Klicken Sie auf den Pfeil neben einer der Ergebnisspaltenüberschriften und wählen Sie folgende Optionen:

Spalten > Abdeckungsbereich (N)

Die Spalte **Abdeckungsbereich (N)** 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 leistungsfähige 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 interessiert sind:

14. 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.
15. Wählen Sie die Option **Filter aktivieren** und wählen Sie *Big Screen TV* (Großbildfernseher) als Element aus.
16. Klicken Sie 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 von Kunden zugrunde liegt. Dazu könnten Sie beispielsweise Regeln ausschließen, die auf dem Verhalten von weniger als 10 Kunden beruhen:
17. Klicken Sie auf den Pfeil neben der Spaltenüberschrift "Abdeckungsbereich (N)" und wählen Sie die Option **Sortieren: Aufsteigend** aus.
18. Wählen Sie in der Spalte "Ausschließen?" alle Regeln aus, für die "Abdeckungsbereich (N)" einen Wert kleiner als 10 Instanzen aufweist.
19. 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 für das Modell ein Scoring durchgeführt wird, klicken Sie auf **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* vorhersagen.
20. Um zu sehen, wie viele dieser Regeln angewendet werden, klicken Sie auf **Testen** und dann auf **Ausführen**. Nun wird für weitere Kunden ein Großbildfernseher empfohlen und einige weisen ein höheres Konfidenzniveau auf als andere.
21. Schließen Sie das Dialogfeld "Testen".
22. 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 Abdeckung, der Regelrelevanz, des höchsten Werts für *Lift* (Anstieg der Wahrscheinlichkeit, dass ein bestimmter Artikel gekauft wird) oder der Bereitstellbarkeit auszuwählen. Außerdem ist es sinnvoll, sicherzustellen, dass Kunden, die bereits einen Großbildfernseher gekauft haben, kein weiteres 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. Weitere Informationen finden Sie im Thema „Scoring von Modellen in eine Datenbanktabelle, eine Datei, eine Analytic Server-Instanz oder einen IBM Cognos BI-Server“ auf Seite 75.
23. Speichern Sie das Projekt bei Bedarf als *retail_association_model.str*.

Weitere Informationen zur Assoziationsmodellierung finden Sie in „Erstellen eines Assoziationsmodells“ auf Seite 64.

Kapitel 3. Navigation im Arbeitsbereich

Ändern der Anzeigesprache

Jeder Benutzer kann die Sprache seiner Anwendung ändern, ohne einen Server erneut starten zu müssen. Beispielsweise können Benutzer, die Deutsch, Englisch oder Japanisch sprechen, dieselbe Anwendung in ihrer jeweiligen Sprache verwenden.

Rufen Sie über das Menü **Tools** in Ihrem Browser **Internetoptionen** oder **Optionen** auf, wechseln Sie zur gewünschten Sprache und aktualisieren Sie die Anzeige anschließend. Wenn IBM Analytical Decision Management nicht in Ihrer gewünschten Sprache bereitgestellt wird, wird das Programm standardmäßig in Englisch oder, falls verfügbar, in der zweiten gewünschten Sprache angezeigt.

Die Schnittstelle von IBM Analytical Decision Management unterstützt die folgenden Sprachen. Beachten Sie, dass diese Liste von Release zu Release geändert wird und dass die Dokumentation nicht in allen Sprachen verfügbar ist, in denen die Benutzerschnittstelle zur Verfügung steht.

- Chinesisch (China/vereinfacht)
- Chinesisch (Taiwan/traditionell)
- Deutsch
- Englisch
- Französisch
- Italienisch
- Japanisch
- Koreanisch
- Portugiesisch (Brasilien)
- Russisch
- Spanisch

Starten von Anwendungen

Auf der Anwendungsstartseite 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 der Anwendungsstartseite 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 für diese Seite setzen oder eine Verknüpfung auf Ihrem Desktop einrichten.

Wenn beim Zugreifen auf die URL von IBM Analytical Decision Management oder auf URLs von IBM SPSS Collaboration and Deployment Services Probleme auftreten, wurden sie möglicherweise von einem Administrator angepasst. Wenden Sie sich an Ihren Administrator. Weitere Informationen finden Sie im Kapitel *Kontextstammverzeichnisse der Anwendung* des Installations- und Konfigurationshandbuchs zu IBM SPSS Collaboration and Deployment Services Repository.

- Um ein neues Projekt, ein neues Modell oder eine neue Regel zu erstellen, wählen Sie aus der Dropdown-Liste für die gewünschte Anwendung die Option **Neu** aus und klicken 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. Sie können auch **Durchsuchen** in der Dropdown-Liste auswählen, um ein anderes Objekt bzw. eine andere Version aus-

zuwählen. 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.

Anpassen der Startseite

Mit den Steuerelementen rechts oben auf der Startseite 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 der Anwendungsstartseite die Option **Anwendung hinzufügen** aus. (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 auf der Startseite 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 Löschesymbol (rotes X) in der rechten oberen Ecke des Anwendungsfelds. Alle Verknüpfungen, die auf diese Anwendung verweisen, werden ebenfalls gelöscht. (Das Löschesymbol ist nicht verfügbar, wenn das Anwendungsfeld auf der Seite vom Administrator gesperrt wurde.)
- Ungespeicherte Projekte, Modelle und Regeln werden in der Galerie der Startseite aufgeführt. Diese Galerie enthält Modelle, die sich noch im Aufbau befinden, sowie fertige, aber noch nicht gespeicherte Modelle. Weitere Informationen finden Sie im Thema „Galerie“.

Homepage

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

Daten. Definieren Sie Datasets zur Verwendung durch die Anwendung.

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

Scoren. Scoren Sie das Modell.

Galerie

Die Galerie wird auf der Anwendungsstartseite 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 verwenden, 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 der Registerkarte "Projekte" in der Galerie können Sie über das Dropdown-Menü **Erstellt in** die auf der Registerkarte für jede Anwendung angezeigten Elemente filtern. Wenn Sie z. B. in einer bestimmten Anwendung arbeiten, möchten Sie möglicherweise nicht, dass für andere Anwendungen vorgesehene Objekte ebenfalls angezeigt werden.

Anpassen von Voreinstellungen

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 Aktionen abgewickelt werden müssen. Die von Ihnen festgelegten Voreinstellungen werden in alle Anwendungen übernommen, die Sie verwenden, 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 9. 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 Anwendungsverknüpfungen. Wählen Sie die Anzahl der Objekte aus, die im Dropdown-Menü der Anwendungsstartseite 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 aus. Beachten Sie, dass dadurch vorhandene Daten mit anderer Währung nicht überschrieben werden.

Dezimalstellen. Wählen Sie die gewünschte Anzahl an Dezimalstellen aus.

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

Anzahl der Ergebnisse pro Seite (in Seiten unterteilte Tabellen). Wählen Sie die maximale Anzahl von Ergebnissen aus, die beim Darstellen von Tabellen auf jeder Seite angezeigt werden soll.

Maximale Anzahl der in einem Diagramm angezeigten Elemente. Wenn zu viele Objekte für die Anzeige vorhanden sind, eignet sich die Diagrammform für Ergebnisse unter Umständen nicht besonders. Ge-

ben Sie die maximale Anzahl an Objekten an, die in Diagrammen angezeigt wird. Auf diese Weise wird statt einem Diagramm eine Tabelle angezeigt, wenn die Anzahl der Objekte diesen maximalen Wert überschreitet. Diese Einstellung gilt nur für kategoriale Felder in Datenübersichten und für das Diagramm des Scoring-Zusammenfassungsberichts für Modelle mit einem kategorialen Ziel.

Anzahl der auf Diagramme und Tabellen anzuwendenden Datensegmente (N-Perzentile). Wählen Sie die maximale Anzahl der Datensegmente aus, die innerhalb der Anwendung, soweit zutreffend, für Diagramme und Tabellen verwendet werden sollen.

Modelleinstellungen

Standardmäßig erstellter Modelltyp. Legen Sie fest, ob die Anwendung das Modell automatisch anhand der verfügbaren Daten erstellen soll (Vorhersagemodell), ob ein Modell verwendet werden soll, das Segmente auffindig macht, die in Bezug auf ein bestimmtes Resultat eine höhere oder geringere Wahrscheinlichkeit aufweisen (Vorhersageregelmotiv (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 Dropdown-Menü sind nur Modelltypen verfügbar, die für die Anzeige in der Anwendung konfiguriert wurden.

Daten automatisch partitionieren, um die Modellauswertung von Entwicklungsdatenquellen zu ermöglichen (für weitere Evaluierung und Tests). Bei Auswahl dieser Option werden die Daten zum Trainieren und Testen des Modells in separate Subsets oder Stichproben aufgeteilt. Wenn Sie das Modell auf der Grundlage eines Subsets aufbauen und anhand eines anderen Subsets testen, erhalten Sie eine ungefähre Vorstellung davon, inwieweit es sich auf andere Datensätze verallgemeinern lässt. Sie können außerdem einen Prozentsatz für die zufällig generierte *Trainingspartitionsgröße* und die *Testpartitionsgröße* angeben. Diese Optionen sind für alle Modelltypen mit Ausnahme von Assoziationsmodellen verfügbar.

Daten automatisch bereinigen und zur zuverlässigen Modellerstellung aufbereiten. Wenn Sie diese Option auswählen, 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 auswählen, ob für den Bereinigungsprozess Geschwindigkeit und Genauigkeit gleichermaßen berücksichtigt oder einem der beiden Priorität gegeben werden soll, z. B., wenn Sie mit sehr großen Datensätzen arbeiten oder ein schnelles Ergebnis benötigen.

Anmerkung: Die Einstellungen für die automatische Datenaufbereitung gelten nur beim Erstellen eines neuen Modells. Eine Klassierung erfolgt nicht.

Zeit für Erstellung eines einzelnen Modells beschränken. Legen Sie mit dieser Option die Standardzeit fest, die zur Erstellung eines Modells aufgebracht wird. Dies kann insbesondere bei der Verarbeitung sehr großer Datensätze 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.

Standardtest-/Simulationsdatum. Geben Sie an, ob als Standardsimulationsdatum das aktuelle Datum oder ein anderes Datum verwendet wird.

Öffnen und Speichern von Projekten, Modellen und Regeln

Sie können Ihre Arbeit in IBM SPSS Collaboration and Deployment Services Repository öffnen, speichern oder löschen oder in das lokale Dateisystem herunterladen. Sie können auch IBM SPSS Modeler-Streamdateien aus dem lokalen Dateisystem herunterladen und in IBM SPSS Collaboration and Deployment Services Repository speichern. Die Speicherung in 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 an der letzten bekannten Position in IBM SPSS Collaboration and Deployment Services Repository. Optional können Sie beim Speichern eine neue Beschriftung übernehmen.



Speichert und beschriftet das aktuelle Projekt oder speichert es unter einem neuen Namen. Soweit zutreffend können Sie zwischen der Speicherung in IBM SPSS Collaboration and Deployment Services Repository und dem Herunterladen in das lokale Dateisystem wählen.



Öffnet ein Projekt aus IBM SPSS Collaboration and Deployment Services Repository oder aus dem lokalen Dateisystem. Gegebenenfalls können Sie ein Projekt auch aus dem lokalen Dateisystem hochladen. Der Zugriff auf das Dateisystem wird nicht für alle Anwendungen unterstützt. Wenn Sie einen IBM SPSS Modeler-Stream hochladen, können Sie auswählen, ob die Datei geöffnet oder im Repository gespeichert werden soll.

Öffnen und Speichern über das Repository

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. 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 von 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 gegebenenfalls auf das Vorschau- oder Informationssymbol, um weitere Details zu dieser Datei anzuzeigen.

Um eine Datei zu löschen, klicken Sie auf das Löschsymbold neben der Datei. Dies ist vor allem für Benutzer hilfreich, die keinen Zugriff auf die Client-Software für IBM SPSS Collaboration and Deployment Services Deployment Manager haben.

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 Exportieren von Regeln, die in einer Anwendung verwendet werden, steht auch das Dropdown-Menü **Speichern unter** oder **Exportieren als** zur Verfügung. Wählen Sie im Dropdown-Menü die Option **Einzelnes Regelset** oder **Einzelregeln** aus, um anzugeben, wie die Regel gespeichert werden soll. Wenn nur referenzierte Regeln verfügbar sind, ist das Dropdown-Menü inaktiviert und die Regel wird als einzelnes Regelset 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 beispielsweise *Absatz* und *1-2009* eingeben.

Beschreibung. Geben Sie eine ausführlichere Beschreibung des gespeicherten Objekts ein. Sie könnten beispielsweise *Analyse des Einzelhandelsabsatzes; erstes Quartal 2009* eingeben.

Beschriftung(en). Falls Sie mehrere Versionen desselben gespeicherten Objekts erstellen, können Sie jeder eine eigene, eindeutige Beschriftung geben. Wählen Sie entweder eine vorhandene Beschriftung 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* verwendet werden soll. Wenn eine neuere Version der Regel erstellt wird, ändert sich die in Ihrem Projekt verwendete 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 dringend empfohlen - insbesondere in einer Produktionsumgebung -, da spätere Änderungen an der neuesten Version ansonsten zu nicht erwarteten Konsequenzen führen können. In bestimmten Fällen, z. B. bei der Referenzierung einer Regel innerhalb einer Regel, muss eine beschriftete Version verwendet werden. Weitere Informationen finden Sie im Thema „Exportieren und Wiederverwenden von Regeln“ auf Seite 56.

Wenn keine Version festgelegt wird, wird die jeweils neueste Version 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. Weitere Informationen finden Sie im Thema „Öffnen und Speichern von Projekten, Modellen und Regeln“ auf Seite 31.

Einschränkungen für Beschriftungsnamen. Beim Definieren von Beschriftungen für IBM Analytical Decision Management dürfen die Zeichen **?** **#** **&** und **:** (Doppelpunkt) nicht im Beschriftungsnamen vorkommen, da die Anwendung ansonsten Referenzen auf Objekte, die diese Beschriftungen verwenden, nicht ordnungsgemäß interpretieren könnte. Bei folgendem Beispiel handelt es sich um ungültige Beschriftungsnamen:

Testen & bereitstellen

BEREITSTELLEN:PHASE1

Sperrungen von Projekten

Wenn Sie ein Projekt, ein Modell oder eine Regel öffnen, steht in der rechten oberen Ecke das Kontrollkästchen "Sperrungen" zur Verfügung, mit dem Sie das Objekt im IBM SPSS Collaboration and Deployment Services Repository sperren können, sodass andere Benutzer es nicht bearbeiten können.

- 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 gegebenenfalls zum Speichern der Änderungen aufgefordert. Ungespeicherte Änderungen gehen verloren, wenn das Objekt freigegeben wird.
- Nachdem Sie ein Projekt freigegeben haben, können Sie weiterhin Änderungen an ihm vornehmen, wenn es ein anderer Benutzer nicht zufällig sperrt, während Sie es freigegeben haben (in diesem Fall können Sie keine Änderungen mehr speichern).

Eigenschaften

Das Dialogfeld "Eigenschaften" enthält die Registerkarte **Projekt** und die Registerkarte **Repository**. Zur Anzeige des Dialogfelds "Eigenschaften" klicken Sie auf das Eigenschaftensymbol der Anwendung.



Abbildung 10. Eigenschaftensymbol

Auf der Registerkarte **Projekt** können Sie die maximal zulässige Anzahl an Mitgliedern für *nominale* Felder angeben. Dies wird auf der Registerkarte "Daten" verwendet. Wählen Sie **Maximale Mitglieder für nominale Felder** aus und geben Sie einen Wert an. Der Standardwert ist **250**. Dies ist in Situationen wie den folgenden hilfreich:

1. Ein Benutzer erstellt eine neue Datenquelle und führt einen Datenscan aus.
2. Für eine Reihe von Feldern, die nützliche kategoriale Informationen enthalten, wird im Projektdatenmodell *Ohne Typ* angezeigt, da die Anzahl unterschiedlicher Werte in den Daten den Standardgrenzwert für die maximale Größe, der im Dialogfeld "Eigenschaften" festgelegt wurde, überschreitet.
3. Der Benutzer klickt auf das Eigenschaftensymbol und erhöht den Wert für **Maximale Mitglieder für nominale Felder** auf eine geeigneten Größe für seine Daten.
4. Der Benutzer kehrt zur Registerkarte "Daten" zurück und klickt auf **Datenscan aktualisieren**, für die Felder wird jedoch weiterhin *Ohne Typ* angezeigt, da durch die Aktualisierungsaktion nicht das Maß geändert wird, damit Änderungen, die der Benutzer vorgenommen hat, beibehalten werden. Anders gesagt wird bei einer Aktualisierung das vorhandene Maß für ein Feld beibehalten und es werden nur die Werte neu berechnet, da alle Änderungen, die der Benutzer am Datenmodell vorgenommen hat, beibehalten werden sollten.
5. Damit die Werte in das Projektdatenmodell aufgenommen werden, muss der Benutzer das Maß jedes Feldes vor der Ausführung der Aktualisierung manuell von *Ohne Typ* in *Kategorial* ändern. Beachten Sie, dass es schneller sein kann, die Datenquelle neu zu erstellen, wenn viele Felder betroffen sind.

Beachten Sie, dass die Einstellung "Maximale Mitglieder für nominale Felder" nicht für Datenquellen mit dem Dateityp *.sav* gilt. Viele *SAV*-Dateien enthalten vom Benutzer angepasste Felder und Werte, sodass sie von dieser Einstellung nicht betroffen sind. Der Benutzer muss die Felder manuell ändern.

Auf der Registerkarte **Repository** können Sie einem Projekt, Modell oder Regelset Schlüsselwörter und eine Beschreibung hinzufügen, um das Wiederauffinden zu einem späteren Zeitpunkt zu erleichtern:

- **Schlüsselwörter.** Geben Sie Schlüsselwörter zur Erleichterung der Suche nach dem gespeicherten Objekt ein. Sie könnten beispielsweise *Absatz* und *Q1-2013* eingeben.

- **Beschreibung.** Geben Sie eine ausführlichere Beschreibung des gespeicherten Objekts ein. Sie könnten beispielsweise *Analyse des Einzelhandelsabsatzes; erstes Quartal 2013* eingeben.

Kapitel 4. Verwalten von Datenquellen

Auf der Registerkarte **Daten** können Sie Datenquellen für Analysen, Simulationen und Tests sowie für das Scoring und andere Verwendungszwecke definieren. Datasets können von Ihrem Administrator konfiguriert oder von Ihnen selbst hinzugefügt werden.

Die Registerkarte enthält eine Registerkarte **Quelle** und eine Registerkarte **Abgeleitet**. Die Registerkarte **Quelle** wird in die folgenden Hauptbereiche unterteilt:

- **Projektdatenmodell.** Definiert die von der Anwendung benötigten Felder. Alle anderen Datasets werden relativ zu dieser Quelle abgebildet. Im Abschnitt **Datenquellenfelder** werden die Eingabefelder und -typen aufgelistet, die von Ihrem Projekt verwendet werden. Bei Bedarf können Sie über die Registerkarte **Abgeleitet** Ausdrücke oder Modellausgabe aus unterschiedlichen Datenquellen hinzufügen, um Ihr Projektdatenmodell zu erweitern.
- **Datenquellen.** Im Abschnitt **Projektdatenquellen** werden die Datenquellen aufgelistet, die zur Verwendung mit dem aktuellen Projekt oder Modell oder mit der aktuellen Regel gespeichert wurden. Im Abschnitt **Eigene Datenquellen** werden die von Ihnen definierten oder gespeicherten Datenquellen aufgelistet. Datasets aus dieser Liste können in andere Projekte kopiert oder aus diesen kopiert und anschließend zur Verwendung als Projektdatenmodell hochgestuft 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.

Datensatz-ID. Falls gewünscht, können Sie ein Feld auswählen, das als Entitäts-ID verwendet werden soll, um den SQL-Pushback zu aktivieren. Dies kann die Leistung verbessern, wenn die Daten aus einer Datenbank stammen und wenn das ausgewählte Feld eine ganze Zahl ist. In IBM SPSS Modeler Advantage werden die hier ausgewählten Felder standardmäßig als Ausgabe des Stapelscorings ausgewählt. Als Eingabe für eine Modellerstellung werden sie standardmäßig abgewählt. Dieses Feld ist optional.

Operativ. Wenn dieses Kontrollkästchen ausgewählt ist, sind die Daten im Feld für die Verwendung in Regeln und Ausdrücken sowie für Vorhersagen verfügbar. Entfernen Sie das Kontrollkästchen, wenn das Feld als Zielfeld für die Analyse verwendet werden soll. Wenn Sie beispielsweise eine Abfrage durchführen, um zu beurteilen, wie gut Kunden auf eine Direktmailing-Kampagne ansprechen würden, dann hätte das Ergebnisfeld analytischen Charakter, da Ihnen die Antwortwerte nicht im Voraus bekannt sind.

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

Messung. Zeigt das Messniveau jedes Felds der Datenquelle an. Falls Sie eines der Messniveaus ändern, müssen Sie den Datenscan aktualisieren, damit die Datenkompatibilität gewährleistet bleibt. Weitere Informationen finden Sie im Thema „Messniveaus“ auf Seite 46.

Werte. Listet die Werte für jedes Feld in der Datenquelle auf, z. B. Minimal- und Maximalwerte in einem bestimmten Bereich. Wenn Sie bei Flagfeldern den Mauszeiger über den Wert halten, wird der für dieses Feld definierte "wahre" und "falsche" Wert in einer QuickInfo angezeigt. Sie können auch auf einen Wert klicken, um ihn zu bearbeiten.

Normalerweise definiert das Projektdatenmodell die Eingaben als Set von Feldern, die aus einer einzigen logischen Tabelle stammen. Dies ist bei IBM SPSS Modeler Advantage immer der Fall. Bei anderen Anwendungen kann das Projektdatenmodell jedoch aus einer Primärtabelle mit 0 oder mehr Sekundärtabellen bestehen. Diese Sekundärtabellen werden von den sekundären Datenquellen definiert, die der primären Datenquelle zugeordnet sind.

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, sodass Benutzer sie nicht verändern oder entfernen können, oder alle Datenoptionen zu sperren, sodass 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 verwendeten 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. Klicken Sie auf das Symbol in dieser Spalte, um eine Stichprobe der in der Quelle enthaltenen Daten anzuzeigen. Weitere Informationen finden Sie in „Vorschau von Daten“ auf Seite 45.

Übersicht. Klicken Sie auf das Symbol in dieser Spalte, um eine Übersicht der Datenquelle zu erhalten. Weitere Informationen finden Sie in „Datenübersicht“ auf Seite 45.

Kompatibel. Zeigt entweder einen Hinweis an, dass die Datenquelle als Projektdatenmodell verwendet wird, oder andernfalls einen grünen, orangefarbenen oder roten Kreis, der angibt, wie kompatibel die Datenquelle mit dem Projektdatenmodell ist.

- 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 orangefarbener 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, sodass sie den vom Projektdatenmodell verlangten entsprechen, sowie Abbildungen aufheben. Weitere Informationen finden Sie in „Abbilden von Feldern“ auf Seite 47.

Kopieren. Kopiert die 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 entsprechende Berechtigung erhalten). Die Felder in diesem Bereich der Registerkarte funktionieren genauso wie die des Bereichs **Projektdatenquellen**, mit dem Unterschied, dass es keine **Kompatibilitätsspalte** 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** aus. Weitere Informationen finden Sie in „Erstellen einer neuen Datenquelle“.
- Um das Messniveau oder den Typs eines Felds in der Datenquelle im **Projektdatenmodell** zu ändern, wählen Sie das entsprechende Niveau aus. Weitere Informationen finden Sie in „Messniveau“ auf Seite 46.
- 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.) Weitere Informationen finden Sie in „Abbilden von Feldern“ auf Seite 47.
- Für eine Vorschau der Datenquelle klicken Sie auf das Symbol **Vorschau**. Weitere Informationen finden Sie in „Vorschau von Daten“ auf Seite 45.
- Wenn Sie zur Erweiterung Ihres Projektdatenmodells Zusatzfelder hinzufügen wollen (Ausdrücke, Felder, die Segmentregeln verwenden, oder Modellausgabe aus unterschiedlichen Datenquellen), klicken Sie auf die Registerkarte **Abgeleitet**. Weitere Informationen finden Sie in „Ableiten von Feldern“ auf Seite 48.

Erstellen einer neuen Datenquelle

1. Zum Öffnen des Datenquellen-Editors wählen Sie entweder im Bereich "Projektdatenquellen" oder im Bereich "Eigene Datenquellen" den Befehl **Datenquelle hinzufügen** aus.
2. Geben Sie einen Namen für die Datenquelle ein.
3. Wählen Sie den Typ der Quelle aus. Beachten Sie, dass die Liste der verfügbaren Datenbanken die Datenbanken enthält, die auf dem IBM SPSS Modeler-Computer verfügbar sind.
 - **Datei**. Es werden zahlreiche Dateitypen unterstützt, z. B. Dateien im Textformat (**.txt*, **.csv* usw.), im Excel-Format, im von IBM SPSS Statistics verwendeten **.sav*- und **.zsav*-Format und im SAS-Format. Weitere Informationen finden Sie im Thema „Dateiquelle“ auf Seite 38.
 - **Datenbank**. Alle Datenbanken, die ODBC unterstützen, z. B. Microsoft SQL Server, DB2, Oracle u. a. Weitere Informationen finden Sie im Thema „Datenbankquelle“ auf Seite 41.
 - **Datenansicht**. Erstellt eine Verbindung zur Verwendung mit Regeln von IBM Operational Decision Management. Beachten Sie, dass Administratoren bei komplexen Regeln, die ein hierarchisches Datenmodell verwenden, die Verwendung solcher Regeln einschränken können, da sie zu einer Leistungsverschlechterung führen. Weitere Informationen finden Sie im Thema „Datenansichtsquelle“ auf Seite 41.
 - **IBM Cognos BI-Server**. Jede Datenbank, die IBM Cognos BI-Server unterstützt. Weitere Informationen finden Sie im Thema „IBM Cognos BI-Quelle“ auf Seite 42.
 - **Analytic Server**. Beim Erstellen einer neuen Datenquelle können Sie diesen Quellentyp nicht auswählen. In IBM SPSS Modeler 16 wurde IBM SPSS Analytic Server eingeführt und ermöglichte die Verwendung von HDFS (Hadoop Distributed File System) als Datenquelle. IBM SPSS Modeler-Benutzer können mithilfe des HDFS-Quellenknotens von Analytic Server Streams erstellen und Modelle auf der Grundlage des big-Data-Systems "Hadoop" erstellen. Benutzer von IBM Analytical Decision Management können diese Modellstreams in IBM SPSS Modeler Advantage zum Anzeigen öffnen (sie können keine neuen Projekte mithilfe von Analytic Server-Datenquellen erstellen).
4. 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 die Instanz von 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 zur Zielinstanz von IBM SPSS Modeler Server und wählen Sie diese aus. Bei der Auswahl des Ziels können Sie den Dateinamen bei Bedarf ändern.

Die Option **Lokale Datei hochladen** ist standardmäßig inaktiviert. Administratoren können die browserbasierte Komponente 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 aus, wenn die Namen der einzelnen Spalten in der ersten Zeile der Datenquelle angegeben sind.

Verschlüsselung. Gibt die verwendete Textcodierungsmethode an. Sie haben die Wahl zwischen der Systemstandardeinstellung und UTF-8.

Dezimalzeichen. 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 mehrere Trennzeichen angeben, z. B. ", |" für Datensätze mit mehreren Trennzeichen. Das Standardtrennzeichen ist das Komma.

Hinweis: Wenn das Komma auch als Dezimaltrennzeichen definiert wurde, funktionieren die Standardeinstellungen 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.

Datenformat angeben. Für Datenquellen des Typs **Datei** mit **Felder unterschiedlicher Länge in Nur-Text** können Sie in diesem Abschnitt den Eingabespeichertyp und das Format der einzelnen Felder festlegen, um sicherzustellen, dass die Werte ordnungsgemäß gelesen werden. Dies ist der im Client von IBM SPSS Modeler verfügbaren Funktion vergleichbar. Die Option **Außer Kraft setzen** gibt an, ob der Standardwert außer Kraft gesetzt wird. Durch die Auswahl von **Außer Kraft setzen** werden die Steuerelemente **Speicherung** und **Eingabeformat** aktiviert. Durch die Inaktivierung von **Außer Kraft setzen** werden die Werte auf ihre ursprünglichen Standardwerte zurückgesetzt. **Eingabeformat** gilt nur für die Speichertypen 'rell', 'Datum', 'Zeit' und 'Zeitmarke'.

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

Ungültige Zeichen. Wählen Sie **Verwerfen** aus, um ungültige Zeichen aus der Datenquelle zu entfernen. Wählen Sie **Ersetzen durch** aus, um ungültige Zeichen durch das angegebene Symbol (nur ein Zeichen) zu ersetzen. Ungültige Zeichen sind Nullzeichen (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 **Paare bilden und verwerfen**, um Anführungszeichenpaare zu finden und zu löschen. Kann einem Anführungszeichen kein zweites Anführungszeichen zugeordnet werden, wird eine Fehlernachricht ausgegeben. Sowohl die Option **Verwerfen** als auch **Paare bilden und verwerfen** speichert den Feldwert (ohne Anführungszeichen) als Zeichenfolge.

Anmerkung: Bei Verwendung von **Paare bilden und verwerfen** werden Leerzeichen beibehalten. Bei Verwendung von **Verwerfen** werden nachfolgende Leerzeichen innerhalb und außerhalb von Anführungszeichen entfernt (beispielsweise wird ' " ab c" , "d ef " , " gh i " ' zu 'ab c, d ef, gh i'). Bei Verwendung von **Als Text einschließen** werden Anführungszeichen als normale Zeichen behandelt, sodass führende und nachfolgende Leerzeichen entfernt werden.

Eingabefelder festlegen. Weitere Informationen finden Sie im Thema „Auswählen von Eingabefeldern“ auf Seite 45.

Excel-Datendateien

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

Erste Zeile enthält Spaltennamen. Wählen Sie diese Einstellung aus, 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 inaktiviert.

Wählen Sie ein Arbeitsblatt aus: 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 nicht leeren Zeile oder ab einem explizit angegebenen Zellenbereich importieren.

- **Bereich beginnt bei erster nicht leerer Zeile.** Sucht die erste nicht leere 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. Weitere Informationen finden Sie im Thema „Auswählen von Eingabefeldern“ auf Seite 45.

IBM SPSS Statistics-Datendateien

Wenn Sie eine IBM SPSS Statistics-Datenquelle auswählen (SAV- oder ZSAV-Datei), werden Sie aufgefordert, weitere Details einzugeben. Wenn die Datei kennwortgeschützt ist, werden Sie aufgefordert, das Kennwort einzugeben.

Variablennamen. Wählen Sie eine Methode für die Behandlung von Variablennamen und -beschriftungen nach dem Import aus einer SAV- oder ZSAV-Datei von IBM SPSS Statistics aus.

- **Namen und Beschriftungen lesen.** Wählen Sie diese Option aus, um sowohl Variablennamen 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 aus, um anstelle der kurzen Feldnamen die beschreibenden Variablenbeschriftungen aus der SAV-Datei von IBM SPSS Statistics einzulesen und diese Beschriftungen als Variablennamen zu verwenden.

Werte. Wählen Sie eine Methode für die Behandlung von Werten und Beschriftungen nach dem Import aus einer SAV- oder ZSAV-Datei von IBM SPSS Statistics aus.

- **Daten und Beschriftungen lesen.** Wählen Sie diese Option aus, um sowohl die Werte selbst als auch die Wertbeschriftungen einzulesen; dies ist die Standardeinstellung.
- **Beschriftungen als Daten lesen.** Wählen Sie diese Option aus, wenn Sie anstelle der numerischen oder symbolischen Codes, mit denen die Werte dargestellt werden, die Wertbeschriftungen aus der SAV- oder ZSAV-Datei verwenden möchten. Wenn Sie diese Option beispielsweise für Daten mit dem Feld "Geschlecht" auswählen, dessen Werte 1 und 2 für *männlich* und *weiblich* stehen, wird das Feld in eine Zeichenfolge konvertiert, und *männlich* und *weiblich* werden als tatsächliche Werte importiert.

Vor dem Auswählen dieser Option sollten Sie Ihre IBM SPSS Statistics-Daten unbedingt auf fehlende Werte überprüfen. Wenn ein numerisches Feld beispielsweise Beschriftungen nur für fehlende Werte verwendet (0 = *Keine Antwort*, -99 = *Unbekannt*), werden bei Auswahl der obigen Option nur die Wert-

beschriftungen *Keine Antwort* und *Unbekannt* importiert und das Feld wird in eine Zeichenfolge konvertiert. In diesem Fall sollten Sie die Werte selbst importieren.

Eingabefelder festlegen. Weitere Informationen finden Sie im Thema „Auswählen von Eingabefeldern“ auf Seite 45.

Datenbankquelle

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

Datenbank. Wählen Sie die Datenbank aus, die Sie verwenden möchten.

Verbinden. Wählen Sie aus, ob die Verbindung zur Datenbank über Datenbankberechtigungs-nachweise oder über im IBM SPSS Collaboration and Deployment Services Repository gespeicherte Berechtigungs-nachweise hergestellt werden soll. Geben Sie dann Ihren Datenbanknamen und Ihr Kennwort ein oder wählen Sie einen Berechtigungs-nachweis im Repository aus und klicken Sie auf **Verbinden**, um zu prüfen, ob Sie auf die Datenbank zugreifen können.

Beachten Sie, dass der Datenbankname und das Kennwort nicht gespeichert werden. Daher werden Sie beim nächsten Öffnen und Ausführen des Projekts aufgefordert, sich anzumelden, außer Ihr Administrator hat Single Sign-on (SSO) für IBM SPSS Modeler Server- und ODBC-Datenbankverbindungen konfiguriert. Weitere Informationen finden Sie im Abschnitt *Überlegungen zur Sicherheit* im Installationshandbuch zu IBM Analytical Decision Management.

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

Führende und nachfolgende Leerzeichen löschen. Wählen Sie diese Option aus, um führende und nachfolgende Leerzeichen in Zeichenfolgen beim Importieren zu entfernen. Sie können wahlweise führende, nachfolgende, führende und nachfolgende 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 **Immer** aus, wenn *alle* Tabellen- und Feldnamen in Anführungszeichen gesetzt werden sollen.
- Wählen Sie **Nie** aus, wenn Tabellen- und Feldnamen *nie* in Anführungszeichen gesetzt werden sollen.

Eingabefelder festlegen. Weitere Informationen finden Sie im Thema „Auswählen von Eingabefeldern“ auf Seite 45.

IBM Analytical Decision Management unterstützt dieselben Datenbankdatenquellen wie IBM SPSS Modeler zum Lesen und Schreiben von Daten sowie für Push-Operationen für Daten. Details finden Sie im Produktkompatibilitätsbericht für IBM SPSS Modeler 18. Beachten Sie, dass dies nicht für die Datenbanken gilt, die von IBM SPSS Collaboration and Deployment Services 8 unterstützt werden, da es sich hierbei um die Datenbank handelt, die als Host für das Repository dient.

Datenansichtsquelle

Eine Datenansichtsquelle erstellt eine Verbindung, die mit Regeln von IBM Operational Decision Management verwendet werden soll. Beachten Sie, dass Administratoren bei komplexen Regeln, die ein hierarchisches Datenmodell verwenden, die Verwendung solcher Regeln einschränken können, da sie zu einer Leistungsver schlechterung führen. Details finden Sie im Thema „Sperren der Datenoptionen“ auf Seite 87.

In IBM SPSS Modeler kann ein *Datenansichtsknoten* konfiguriert werden, um Daten, die in einer Analyse-datenansicht von IBM SPSS Collaboration and Deployment Services definiert wurden, in Streams einzufügen. Eine Analysedatenansicht definiert eine Struktur zum Zugreifen auf Daten, die die in Vorhersagemodellen und Geschäftsregeln verwendeten Entitäten definieren. Die Ansicht ordnet die Datenstruktur physischen Datenquellen für die Analyse zu.

Weitere Informationen finden Sie in der Dokumentation zu IBM SPSS Modeler und im Abschnitt *Quellen-, Prozess- und Ausgabeknoten* der Hilfe zu IBM SPSS Modeler.

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

Analytische Datenquelle. Wählen Sie die analytische Datenquelle aus dem Repository aus.

Tabellenname. Wählen Sie die Tabelle aus, deren Daten Sie verwenden möchten. Dies ist eine Tabelle aus dem Datenmodell in der Analysedatenansicht. Jede Tabelle im Datenmodell stellt ein Konzept oder eine Entität dar, das bzw. die am Analyseprozess zur Vorhersage beteiligt ist. Felder für die Tabellen entsprechen Attributen der von den Tabellen dargestellten Entitäten. Wenn Sie z. B. Kundenaufträge analysieren, könnte Ihr Datenmodell eine Tabelle für Kunden und eine Tabelle für Bestellungen enthalten. Die Kundentabelle enthält möglicherweise Felder für die Kunden-ID, das Alter, das Geschlecht, den Familienstand und den Wohnsitz. Die Bestellungstabelle enthält möglicherweise Felder für die Bestellungs-ID, die Anzahl der Artikel in der Bestellung, die Gesamtkosten und die ID für den Kunden, der die Bestellung abgegeben hat. Mithilfe des Felds für die Kunden-ID könnten die Kunden in der Kundentabelle ihren Bestellungen in der Bestellungstabelle zugeordnet werden.

Datenzugriffsplan. Wählen Sie den Datenzugriffsplan aus, dessen Daten Sie verwenden möchten. Ein Datenzugriffsplan ordnet die Datenmodelltabellen in einer Analysedatenansicht physischen Datenquellen zu. Eine Analysedatenansicht enthält normalerweise mehrere Datenzugriffspläne. Wenn Sie den verwendeten Datenzugriffsplan ändern, ändern Sie die Daten, die von Ihrem Projekt verwendet werden. Wenn die Analysedatenansicht beispielsweise einen Datenzugriffsplan zum Trainieren eines Modells und einen Datenzugriffsplan zum Testen eines Modells enthält, können Sie von Trainingsdaten zu Testdaten wechseln, indem Sie den verwendeten Datenzugriffsplan ändern.

Optionale Attribute. Wählen Sie, falls verfügbar und bei Bedarf, optionale Attribute aus, die Sie zur Datenquelle hinzufügen möchten.

Eingabefelder festlegen. Weitere Informationen finden Sie im Thema „Auswählen von Eingabefeldern“ auf Seite 45.

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 die mit IBM Analytical Decision Management verwendete Instanz von IBM SPSS Modeler Server den IBM Cognos BI-Server nicht unterstützt, ist diese Option nicht verfügbar.

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

1. 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.
2. 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 **Berechnungsnachweise festlegen** aus und geben Sie den Namespace, den Benutzernamen und das Kennwort 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 **Kennwort** ein, das zum angegebenen Benutzernamen gehört.

ODER

Navigieren Sie zu einem im IBM SPSS Collaboration and Deployment Services Repository gespeicherten IBM Cognos-Berechtigungs-nachweis und wählen Sie ihn aus.

3. Um sich stattdessen ohne Berechtigungs-nachweise anzumelden (bei dieser Variante können die Felder für Namespace, Benutzernamen und Kennwort nicht ausgefüllt werden), verwenden Sie die Option **Anonyme Verbindung verwenden**.

Hinweis: 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.

1. 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.
2. Sie können auch den Modus **Listenbericht** auswählen und auf **Durchsuchen** klicken, um einen einzelnen Listenbericht auszuwählen. Daraufhin wird ein Dialogfeld mit Paketen und Ordnern angezeigt, aus denen ein Bericht ausgewählt werden kann. *Hinweis:* Es können nur einzelne Listenberichte importiert werden. Das Importieren mehrerer Listen wird nicht unterstützt.
3. 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.
4. Nachdem Sie ein Paket ausgewählt haben, können Sie die zu importierenden Objekte auswählen. Sie können also 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. *Hinweis:* Die Daten müssen im UTF-8-Format vorliegen.
5. 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 mindestens ein Objekt, um die Eingabefelder anzuzeigen, die importiert werden können. Durch die Auswahl eines Objekts werden automatisch alle zugehörigen Eingabefelder ausgewählt.

Parameter prüfen. 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 die Spalte













Wert ein. Mithilfe von Parametern können vor dem Importieren von Daten Einstellungen vorgenommen werden (z. B. zum Ausführen einer parametrisierten Berechnung).

Filter prüfen. 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 prüfen. 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, inaktivieren 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 2. Cognos-Objektsymbole.

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

Analytic Server-Quelle

Wenn Sie eine neue Datenquelle erstellen, können Sie den Analytic Server-Datenquellentyp nicht auswählen. Er wird lediglich angezeigt, wenn Sie mithilfe von IBM SPSS Modeler Advantage ein Modell öffnen, das in IBM SPSS Modeler für eine Analytic Server-Datenquelle erstellt wurde.

IBM SPSS Analytic Server wurde in IBM SPSS Modeler 16 eingeführt und ermöglicht die Verwendung von HDFS (Hadoop Distributed File System) als Datenquelle. IBM SPSS Modeler-Benutzer können mithilfe des HDFS-Quellenknotens von Analytic Server Streams erstellen und Modelle auf der Grundlage des Hadoop-Big-Data-Systems erstellen. Benutzer von IBM Analytical Decision Management können diese Modellstreams in IBM SPSS Modeler Advantage öffnen (Sie können keine neuen Projekte mithilfe von

Analytic Server-Datenquellen erstellen). Beachten Sie, dass Sie bestimmte Aktionen, wie beispielsweise das Ändern des Modelltyps oder des Ziels, nicht ausführen können.

Beim Anzeigen einer Analytic Server-Datenquelle auf der Registerkarte "Daten" stehen die folgenden Informationen zur Verfügung:

Datenquellename. Der Name der Datenquelle in IBM SPSS Modeler. Sie können diesen Namen bei Bedarf ändern. Die übrigen folgenden Felder können nicht bearbeitet werden.

Analytic Server-Quelle. Der Pfad zur Datenquelle, wie er in IBM SPSS Modeler konfiguriert wurde.

Eingabefelder festlegen. Die verwendeten Eingabefelder.

Auswählen von Eingabefeldern

Wenn Sie über große Datenmengen verfügen, können Sie die Menge der aus Ihrer Quelle importierten Daten verringern oder optimieren und beispielsweise nur die für Ihr Interessengebiet relevanten Felder verwenden.

1. Klicken Sie im unteren Bereich des Dialogfelds "Datenquellen-Editor" auf **Eingabefelder festlegen**; im Dialogfeld werden alle Datenfelder Ihrer Quelle eingublendet. Außerdem werden auch alle Ausdrücke angezeigt, die von Ihnen erstellt wurden. Details zum Erstellen von Ausdrücken finden Sie in „Expression Builder“ auf Seite 49.
2. 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. Weitere Informationen finden Sie in „Messniveaus“ auf Seite 46.

Vorschau von Daten

Um zu prüfen, ob Sie die richtige Datenquelle ausgewählt haben, klicken Sie auf das Symbol **Vorschau**, um eine Stichprobe der in der Quelle enthaltenen Daten anzuzeigen.

Datenübersicht

Die Funktion zur Datenübersicht ist auf der Registerkarte "Daten" aller Anwendungen, auf den Registerkarten "Modellierung" und "Scoren" in IBM SPSS Modeler Advantage sowie auf der Registerkarte "Bereitstellen" von Anwendungen verfügbar, die die Funktion "Jetzt scoren" enthalten.

1. Um eine Übersicht der Datenquelle zu erstellen, klicken Sie auf das über die Anwendung verfügbare Symbol "Datenübersicht".



Abbildung 11. Symbol "Datenübersicht"

2. Das Dialogfeld "Datenübersicht" wird angezeigt. Wählen Sie bei Bedarf ein Überlagerungsfeld aus der Dropdown-Liste im Abschnitt "Übersichtsoptionen" aus.
Nach dem Erstellen der Übersicht sind Registerkarten verfügbar, mit denen Sie nur Ergebnisse für das primär ausgewählte Feld anzeigen können oder mit denen Sie das primär ausgewählte Feld mit dem in dieser Dropdown-Liste angegebenen Feld überlagern können. Beispielsweise könnten Sie in den Ergebnissen Informationen zum **Alter** von Kunden anzeigen und diese anschließend mit einem anderen Feld, wie z. B. mit dem Feld **Geschlecht**, überlagern.
3. Wählen Sie bei Bedarf die Option **Partitionierte Daten verwenden** im Abschnitt "Übersichtsoptionen" aus. Diese Option ist verfügbar, wenn Sie auf der Registerkarte "Modellierung" unter "Optionale Einstellungen" die Option **Daten automatisch partitionieren, um die Modellauswertung von Entwicklungsdatenquellen zu ermöglichen (für weitere Evaluierung und Tests)** ausgewählt haben. Diese

Option teilt die Daten in separate Subsets oder Stichproben zum Trainieren und Testen des Modells auf. Wenn Sie das Modell auf der Grundlage eines Subsets erstellen und anhand eines anderen Subsets testen, erhalten Sie eine ungefähre Vorstellung davon, wie es sich auf andere Datasets verallgemeinern lässt.

4. Wählen Sie die Datenfelder aus, die in die Übersicht einbezogen werden sollen, und klicken Sie auf **Übersicht erstellen**. Alle in der Datenquelle verfügbaren Felder werden aufgelistet. Standardmäßig sind alle Modelleingabefelder und das Ziel ausgewählt.

Die Datenübersicht wird erstellt und die Ergebnisse werden angezeigt. Sie können die Informationen sortieren und auswählen, welche Spalten angezeigt werden sollen.

5. Klicken Sie auf ein beliebiges Feld, um die zugehörigen Details anzuzeigen. Für jedes ausgewählte Feld wird eine neue Registerkarte mit Ergebnissen angezeigt, sodass Sie Diagramme und Tabellen anzeigen und Überlagerungsfelder auswählen können, falls verfügbar.

Messniveaus

Mit dem Messniveau können die von den einzelnen Datenfeldern 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* ist. Alternativ können Sie das Messniveau im Dialogfeld **Datenquellen-Editor** ändern, wenn Sie die zu verwendenden Eingabefelder festlegen. Weitere Informationen finden Sie in „Auswählen von Eingabefeldern“ auf Seite 45.

Die folgenden Messniveaus stehen zur Verfügung:

- **Stetig** Wird zur Beschreibung numerischer Werte verwendet, z. B. als Bereich von 0-100 oder 0,75-1,25. Stetige Werte können ganze Zahlen, reelle Zahlen oder Daten/Uhrzeiten sein.
- **Kategorial** Wird für Zeichenfolgewerte verwendet, 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 abhängig von der im Dialogfeld "Projekteigenschaften" angegebenen maximalen Anzahl an Mitgliedern für nominale Felder das Messniveau *Flag*, *Nominal* oder *Ohne Typ* festgelegt.
- **Flag** Wird für Daten mit zwei unterschiedlichen Werten verwendet, die das Vorliegen oder das Nichtvorliegen eines Merkmals angeben, wie beispielsweise wahr und falsch, Ja und Nein oder 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 eine QuickInfo angezeigt, wenn Sie den Mauszeiger über die Werte auf der Registerkarte "Daten" bewegen. Daten können als Text, Ganzzahl, reelle Zahl, Datum, Uhrzeit oder als Zeitmarke dargestellt sein.
- **Nominal** Wird zur Beschreibung von Daten mit mehreren unterschiedlichen Werten verwendet, die jeweils als Mitglied einer Menge behandelt werden, z. B. klein/mittel/groß. Nominale Daten können einen beliebigen Speichertyp aufweisen: "Numerisch", "Zeichenfolge" oder "Datum/Uhrzeit". Beachten Sie, dass die Werte durch die Festlegung des Messniveaus *Nominal* nicht automatisch in den Speichertyp "Zeichenfolge" umgewandelt werden. Weitere Informationen zum Festlegen der maximal zulässigen Anzahl an Mitgliedern für nominale Felder finden Sie in „Eigenschaften“ auf Seite 33.
- **Ordinal** Wird für Daten mit mehreren unterschiedlichen Werten verwendet, die eine implizite Reihenfolge haben. 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 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 verwenden, wo Nominalfelder verwendet werden können. Außerdem können Felder beliebigen Speichertyps (reell, Ganzzahl, Zeichenfolge, Datum, Uhrzeit usw.) als Ordinalfelder definiert werden.

- **Ohne Typ** Wird für Daten verwendet, die keinem der obigen Typen entsprechen, für Felder mit einem einzigen Wert sowie für nominale Daten, bei denen die Menge mehr als die zulässige Anzahl Mitglieder enthält. Dieser Typ ist außerdem in Fällen nützlich, wo das Messniveau andernfalls eine Menge mit sehr vielen Mitgliedern wäre (z. B. Kontonummern). Wenn Sie **Ohne Typ** für ein Feld wählen, wird die Rolle automatisch auf **Keine** gesetzt. **Datensatz-ID** stellt die einzige Alternative dar. Die voreingestellte Maximalgröße für Mengen beträgt 250 verschiedene Werte. Dieser Wert kann im Dialogfeld "Projekteigenschaften", auf das über das Symbolleistensymbol zugegriffen werden kann, angepasst bzw. inaktiviert werden.
- **Sammlung** Wird zur Angabe von nicht räumlichen Daten verwendet, die in einer Liste aufgezeichnet werden. Eine Sammlung ist gewissermaßen ein Listenfeld mit der Tiefe 0, bei dem die Elemente in dieser Liste eines der anderen Messniveaus aufweisen.
- **Georäumlich** Wird zusammen mit dem Speichertyp **Liste** verwendet, um Geodaten anzugeben. Listen können entweder Felder für Ganzzahllisten oder Listen reeller Zahlen sein, deren Listentiefe zwischen null und zwei inklusive beträgt.

Bearbeiten von Datenfeldwerten

In Datentabellen werden in der Spalte **Werte** die Werte für die einzelnen Felder in der Datenquelle aufgelistet (z. B. die Minimal- und Maximalwerte in einem Bereich). Wenn Sie bei Flagfeldern den Mauszeiger über den Wert halten, wird der für dieses Feld definierte "wahre" und "falsche" Wert in einer QuickInfo angezeigt. Sie können auch auf einen Wert klicken, um ihn zu bearbeiten. Ein Dialogfeld wird geöffnet, dessen Inhalt abhängig vom Typ des ausgewählten Datenfelds variieren kann. Zum Beispiel:

- Wenn der Feldtyp **Stetig** ist, können Sie obere und untere Grenzwerte eingeben.
- Wenn der Feldtyp **Flag** ist, können Sie die Werte "wahr" und "falsch" austauschen.
- Wenn der Feldtyp **Nominal** ist, können Sie Werte hinzufügen, entfernen, neu ordnen und umbenennen.

Abbilden 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.

1. 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.
2. Wählen Sie ein Feld in der Spalte **Verfügbare Eingaben** sowie das entsprechende Feld in der Spalte **Pflichteingabefelder** aus und klicken Sie auf **Abbilden**, um die beiden Felder abzubilden. Wiederholen Sie diesen Vorgang, bis alle Felder der Spalte **Verfügbare Eingaben** abgebildet wurden.
3. Wenn Sie fertig sind, klicken Sie auf **OK**.

Sofern möglich, ordnet IBM Analytical Decision Management automatisch kompatible Speichertypen zu, wenn ein Benutzer eine zusätzliche Datenquelle hinzufügt. Wenn der Typ nicht kompatibel ist, darf der Benutzer Typen umwandeln.

In der folgenden Tabelle werden die zulässigen Zuordnungen aufgelistet.

Tabelle 3. Typen, die umgewandelt werden können

In →						
Von ↓	Ganze Zahl	Reelle Zahl	Zeichenfolge	Datum	Zeit	Zeitmarke
Ganze Zahl	nicht zutreffend	Automatisch	Ja	nicht zulässig	nicht zulässig	nicht zulässig
Reelle Zahl	Ja (Dezimalstellen abgeschnitten)	nicht zutreffend	Ja	nicht zulässig	nicht zulässig	nicht zulässig
Zeichenfolge	Ja	Ja	nicht zutreffend	Ja	Ja	Ja
Datum	nicht zulässig	nicht zulässig	Ja	nicht zutreffend	nicht zulässig	Ja (Zeit als 00:00)
Zeit	nicht zulässig	nicht zulässig	Ja	nicht zulässig	nicht zutreffend	Ja (Basisdatum)
Zeitmarke	nicht zulässig	nicht zulässig	Ja	Automatisch (Zeit abgeschnitten)	Automatisch (Datum abgeschnitten)	nicht zutreffend

Ableiten von Feldern

In den meisten Anwendungen von IBM Analytical Decision Management können Sie Ihrem Projektdatenmodell Zusatzfelder hinzufügen, z. B. Ausdrücke, Modellausgabe, über Segmentregeln erstellte Felder oder Felder aus Sekundärtabellen. Auf der Registerkarte **Abgeleitet** der Registerkarte **Daten** werden alle Felder angezeigt, die als Teil der Daten eingeschlossen werden sollen, und die Felder können so angeordnet werden, dass sie kumulativ verwendet werden können. (Ihre Reihenfolge in der Tabelle stimmt mit der Reihenfolge überein, in der sie berechnet werden; beispielsweise kann das zweite abgeleitete Feld das erste abgeleitete Feld als Eingabe verwenden.)

Die Registerkarte **Abgeleitet** ist in der Anwendung IBM SPSS Rules Management nicht enthalten.

- Klicken Sie auf der Registerkarte **Daten** auf die Registerkarte **Abgeleitet**.

So fügen Sie einen Ausdruck hinzu:

- Klicken Sie auf das Symbol **Feld anhand eines Ausdrucks ableiten**. Expression Builder wird angezeigt. Weitere Informationen finden Sie in „Expression Builder“ auf Seite 49.

So fügen Sie eine Modellausgabe hinzu:

1. Klicken Sie auf das Symbol **Feld anhand eines Modells oder einer Regel ableiten** und navigieren Sie dann zu einem Modell und wählen Sie es aus. Das Dialogfeld **Eingaben und Ausgaben** für das ausgewählte Modell wird angezeigt.
2. Wenn die Modelleingabefelder kompatibel sind, wird der Abschnitt **Modelleingabenzuordnung** 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.
3. Im Abschnitt **Modellausgaben** werden alle verfügbaren Modellausgaben für das ausgewählte Modell aufgelistet. Wählen Sie die Modellausgabefelder aus, die Sie zu Ihren Daten hinzufügen möchten. In der Spalte **Modellausgabename** 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 Projektdatenmodellfeld. 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.

4. Unten im Dialogfeld wird der Abschnitt **Modellparameter** aufgenommen, wenn der ausgewählte Stream Scoring-Parameter enthält. Geben Sie Werte für die Modellparameter ein.
5. Klicken Sie auf **Speichern**, wenn Sie die Arbeit in diesem Dialogfeld abgeschlossen haben. Alle ausgewählten Modellausgaben werden dem Abschnitt **Zusatzfelder** auf der Registerkarte **Abgeleitet** hinzugefügt.

So erstellen Sie ein Feld mithilfe von Segmentregeln:

1. Klicken Sie auf das Symbol **Feld anhand von Segmentregeln ableiten**.
2. Erstellen Sie eine neue Segmentregel. Segmente werden durch einen oder mehrere Ausdrücke definiert, die den Wert "wahr" (true) oder "falsch" (false) zurückgeben. Weitere Informationen finden Sie in „Definieren von Segmentregeln“ auf Seite 53.
3. Klicken Sie auf **OK**, wenn Sie dieses Dialogfeld fertig bearbeitet haben. Die neue Segmentregel wird dem Abschnitt **Zusatzfelder** auf der Registerkarte **Abgeleitet** hinzugefügt.

So fügen Sie ein Feld aus einer Sekundärtabelle hinzu:

1. Wenn Sie eine Sekundärtabelle mit einer Eins-zu-eins-Beziehung auf der Registerkarte **Daten** definiert haben, klicken Sie auf das Symbol **Feld aus Sekundärtabelle hinzufügen**.
2. Wählen Sie eine Sekundärtabelle in der Dropdown-Liste und dann mindestens ein Tabellenfeld aus.
3. Klicken Sie auf **Speichern**, wenn Sie die Arbeit in diesem Dialogfeld abgeschlossen haben. Die Felder werden dem Abschnitt **Zusatzfelder** auf der Registerkarte **Abgeleitet** hinzugefügt.

Expression Builder

Mit dem Ausdruckserstellungsprogramm können Sie zusätzliche Felder oder Attribute zur Verwendung durch die 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 genauso wie operative Daten über die Anwendung verfügbar (z. B. in Regeln und Modellprädiktoren).

Wechseln Sie zur Registerkarte **Daten** und klicken Sie auf die Registerkarte **Abgeleitet**, um vorhandene Ausdrücke anzuzeigen oder einen neuen Ausdruck zu erstellen.

Erstellen eines Ausdrucks

1. Geben Sie in das Dialogfeld **Expression Builder** einen eindeutigen Namen für den Ausdruck ein.
2. Geben Sie Daten in das Feld **Ausdrucksgleichung** ein und verwenden Sie dabei die Funktions- und Feldlisten als Referenz
oder
3. Wählen Sie die gewünschten Felder und Funktionen in den Listen aus.
4. Doppelklicken Sie oder klicken Sie auf den blauen Pfeil, um das Feld oder die Funktion dem Feld **Ausdrucksgleichung** hinzuzufügen.
5. Verwenden Sie die Operandenschaltflächen in der Mitte des Dialogfelds, um Operationen in den Ausdruck einzufügen.

6. Geben Sie eine **Beschreibung** ein, aus der hervorgeht, wofür der Ausdruck verwendet wird.
7. Klicken Sie auf **Prüfen**, um den Ausdruck zu validieren, und klicken Sie anschließend auf **Speichern**.

Anmerkung: Sie können Ausdrücke vom IBM SPSS Modeler-Client kopieren und einfügen. Beachten Sie jedoch, dass die Ergebnisse möglicherweise nicht immer wie erwartet ausfallen, wenn der Ausdruck von einem Stream abhängt. (Wenn Sie z. B. eine unterstützte @-Funktion oder Zeilenfunktion einfügen, wird möglicherweise nicht das gewünschte Ergebnis erzeugt, weil der restliche Stream nicht für die entsprechende Unterstützung konfiguriert ist.)

Es werden ausschließlich die in der Benutzerschnittstelle von IBM Analytical Decision Management bereitgestellten Ausdrücke unterstützt.

Auswählen von Funktionen

In der Funktionsliste werden alle verfügbaren -Funktionen und -Operatoren angezeigt. Blättern Sie, um eine Funktion in der Liste auszuwählen, oder zeigen Sie mithilfe der Dropdown-Liste ein Subset von Funktionen oder Operatoren an, um die Suche zu erleichtern. Die verfügbaren Funktionen sind zur leichteren Suche in Kategorien zusammengefasst.

Nachfolgend sind Details zu den Kategorien aufgelistet. Detailliertere Informationen finden Sie in der Dokumentation zu IBM SPSS Modeler.

Tabelle 4. Funktionskategorien

Funktionstyp	Beschreibung
Zuletzt verwendete Funktionen	Enthält eine Liste der während der aktuellen Sitzung verwendeten Funktionen.
Operatoren	Listet alle Operatoren auf, die Sie beim Erstellen von Ausdrücken verwenden können. Operatoren stehen auch über die Schaltflächen in der Mitte des Dialogfelds zur Verfügung.
Informationsfunktionen	Dient dazu, Erkenntnisse über Feldwerte zu gewinnen. Die Funktion <code>is_string</code> beispielsweise gibt "true" für alle Datensätze zurück, die den Typ "Zeichenfolge" haben.
Konvertierungsfunktionen	Wird zum Erstellen neuer Felder oder zum Konvertieren des Speichertyps verwendet. Die Funktion <code>to_timestamp</code> beispielsweise konvertiert das ausgewählte Feld in eine Zeitmarke.
Vergleichsfunktionen	Wird verwendet, um Feldwerte mit anderen Feldwerten oder mit einer angegebenen Zeichenfolge zu vergleichen. Mit <code><=</code> beispielsweise wird verglichen, ob die Werte zweier Felder kleiner oder gleich sind.
Logische Funktionen	Dient zur Durchführung logischer Operationen, z. B. <code>if</code> , <code>then</code> und <code>else</code> .
Numerische Funktionen	Dient zur Durchführung numerischer Berechnungen, z. B. für den natürlichen Logarithmus von Feldwerten.
Zufallsfunktionen	Dient zur zufälligen Auswahl von Elementen oder Generierung von Zahlen.
Zeichenfolgefunktionen	Wird zur Durchführung verschiedener Operationen für Zeichenfolgen verwendet, z. B. <code>stripchar</code> zum Entfernen eines angegebenen Zeichens.
Datums-/Uhrzeitfunktionen	Dient der Durchführung verschiedener Operationen für Datums-, Zeit- und Zeitmarkenfelder.

Tabelle 4. Funktionskategorien (Forts.)

Funktionstyp	Beschreibung
Funktionen für Leerzeichen und Nullen	Wird zum Zugreifen auf oder Markieren von benutzerdefinierten Leerzeichen oder systemdefiniert fehlenden Werten bzw. zum häufigen Füllen dieser Werte verwendet. @BLANK wird nicht unterstützt, da das entsprechende Verhalten von anderen Knoten wie dem Typknoten abhängig ist. @NULL(FIELD) und undef werden unterstützt. @NULL(FIELD) hat einen booleschen Datentyp und gibt "wahr" (true) zurück, wenn der Wert von FIELD der systemdefiniert fehlende Wert \$null\$ ist. Für alle anderen Werte, einschließlich benutzerdefinierter Leerzeichen, wird "falsch" (false) zurückgegeben. Wenn Sie beides prüfen wollen, verwenden Sie @BLANK(FIELD) und @NULL(FIELD). undef weist den Typ "Any" (Alle) auf und wird normalerweise zum Eingeben des Werts \$null\$ verwendet, beispielsweise, um leere Werte durch Nullen im Füllerknoten zu füllen.
Alle Funktionen	Enthält eine vollständige Liste der verfügbaren Funktionen.

Auswählen von Feldern und Werten

In der Feldliste werden alle verfügbaren Felder angezeigt. Blättern Sie, um ein Feld in der Liste auszuwählen. Doppelklicken Sie oder klicken Sie auf den blauen Pfeil, um dem Ausdruck ein Feld hinzuzufügen.

Anzeigen und Auswählen von Werten

Feldwerte können an verschiedenen Stellen in der Anwendung angezeigt werden. Wählen Sie zum Anzeigen von Werten für ein Feld über Expression Builder das gewünschte Feld aus und klicken Sie auf die Schaltfläche **Ausdruckseingaben**, um ein Dialogfeld zu öffnen, in dem die Werte für das ausgewählte Feld aufgelistet werden. Anschließend können Sie einen Wert auswählen und auf **Einfügen** klicken, um den Wert in den aktuellen Ausdruck einzufügen.



Abbildung 12. Schaltfläche "Ausdruckseingaben"

Für Flagfelder und nominale Felder werden alle definierten Werte aufgeführt. Für stetige Felder (numerischer Bereich) werden die Minimal- und Maximalwerte angezeigt.

Überprüfen von Ausdrücken

Klicken Sie in Expression Builder auf **Überprüfen** (unten rechts), um den Ausdruck zu validieren. Ausdrücke, die nicht geprüft wurden, werden rot angezeigt. Wenn Fehler gefunden werden, gibt eine Nachricht die Fehlerursache an.

Folgende Elemente werden überprüft:

- Richtige Anführungszeichen für Werte und Feldnamen
- Richtige Verwendung von Parametern und globalen Variablen
- Richtige Verwendung von Operatoren
- Vorhandensein von referenzierten Feldern
- Vorhandensein und Definition von referenzierten Globalwerten

Wenn Sie Fehler in der Syntax feststellen, versuchen Sie, den Ausdruck mithilfe der Listen- und Operator-schaltflächen zu erstellen, anstatt den Ausdruck manuell einzugeben. Bei dieser Methode werden automatisch die richtigen Anführungszeichen für Felder und Werte hinzugefügt.

Beachten Sie beim Erstellen von Ausdrücken in IBM Analytical Decision Management die folgenden Einschränkungen. Ausdrücke können die folgenden Elemente nicht enthalten:

- Verweis auf einen IBM SPSS Modeler-Datenstromparameter
- Verweis auf einen globalen Datenstrom von IBM SPSS Modeler
- Verweis auf eine Datenbankfunktion
- Verweis auf eine der folgenden θ -Funktionen für Sonderfelder oder Feldwerte:
 - @TARGET
 - @PREDICTED
 - @FIELD
 - @PARTITION_FIELD
 - @TRAINING_PARTITION
 - @TESTING_PARTITION
 - @VALIDATION_PARTITION

Verwenden von Ausdrücken, die einen booleschen Ausdruck zurückgeben

IBM SPSS Modeler und IBM Analytical Decision Management haben keinen booleschen Datentyp. Boolesche Ausdrücke können nicht zum unmittelbaren Erstellen eines neuen Felds verwendet werden. Die Verwendung boolescher Ausdrücke ist auf Ausdrucksstrukturen begrenzt, die einen booleschen Ausdruck erwarten (beispielsweise die Anweisung `if`). Im Folgenden finden Sie drei Beispiele:

```
if ('RFM Score' > 5) then "T" else "F" endif
if ('RFM Score' < 5) then "Low" elseif ('RFM Score' >= 5 and 'RFM Score' < 10) then "Medium"
else "High" endif
hassubstring('Phone type','Coolphone')
```

In den ersten beiden Beispielen oben können die "RFM Score"-Ausdrücke (`'RFM Score' > 5`, `'RFM Score' >= 5` und `'RFM Score' < 10`) als boolesche Ausdrücke aufgefasst werden.

Diese Ausdrücke können nicht allein verwendet werden, um ein Feld abzuleiten. Es würde ein Fehler wie Ungültiger Typ für einen Ausdruck, der keine Bedingung ist: Boolean angezeigt werden. Ausdrücke wie diese können jedoch in bedingten Ausdrücken wie in den vollständigen Beispielen verwendet werden und bedingte Ausdrücke können zum Ableiten von Feldern mit kategorialen Messniveaus, beispielsweise einem Feld des Typs "Flag", verwendet werden. "Flag" ist das Standardmessniveau für ein Feld mit zwei Werten.

Das zweite Beispiel oben könnte ein Feld mit drei unterschiedlichen Werten generieren (solange während des Datenscans jeweils Datensätze in den Kategorien "Low", "Medium" und "High" gefunden werden). Dies würde wahrscheinlich dazu führen, dass das Messniveau des abgeleiteten Felds standardmäßig den Wert "Nominal" annimmt. Es ist unwahrscheinlich, dass das dritte Beispiel oben ein Feld mit zwei unterschiedlichen Werten zurückgibt.

Weitere Informationen zu Messniveaus finden Sie in „Messniveaus“ auf Seite 46.

Kapitel 5. Verwenden von Regeln

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. Alternativ können Sie gemeinsam genutzte Regeln mithilfe von IBM SPSS Rules Management erstellen und bearbeiten.

Erstellen von Regeln

- 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. Weitere Informationen finden Sie im Thema „Exportieren und Wiederverwenden von Regeln“ auf Seite 56.
- Zur Erstellung einer gemeinsam genutzten Regel starten Sie IBM SPSS Rules Management über die Anwendungsstartseite.

Definieren von Segmentregeln






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" (true) oder "falsch" (false) zurückgeben. Beispielsweise könnte eine Regel mit einem einzelnen Segment vorhanden sein, die den Wert "wahr" zurückgibt, wenn der Wert des Felds *Age* kleiner als 18 ist. Segmente können über in den Daten definierte Felder oder Ausdrücke, Neigungswerte oder über die von einem Modell gelieferten Vorhersagen definiert werden. In Anwendungen mit der Registerkarte "Definieren", die so konfiguriert sind, dass mehrere Dimensionshierarchien möglich sind (die als *Elemente* bezeichnet werden), können Sie auch aus anderen Dimensionsmitgliedern eine Auswahl treffen.

Jedes Segment wird durch mindestens einen Ausdruck definiert und gibt den Wert "wahr" zurück, wenn alle Bedingungen erfüllt sind. 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:

1. Klicken Sie in der Anwendung, der Sie die Regel hinzufügen möchten, auf das Symbol **Neue Regel erstellen**.
2. Legen Sie einen Namen für die Regel fest.
3. Klicken Sie auf den Abwärtspfeil, um für die Eingabespalte ein Feld, Modell oder eine Regel anzugeben.

Tabelle 5. Eingabespaltenoptionen.

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

4. Wählen Sie die Funktion aus, die Sie verwenden möchten. In der folgenden Tabelle sind die verfügbaren Optionen aufgelistet.

Tabelle 6. Verfügbare Operatoren.

Werttyp:	Verfügbare Operatoren
Zeitmarke, Datum, Uhrzeit	<ul style="list-style-type: none"> • = (gleich) • != (nicht gleich) • > (größer als) • < (kleiner als) • >= (größer oder gleich) • <= (kleiner oder gleich) • ZWISCHEN (BETWEEN) • IST NULL (ISNIL) • IST NICHT NULL (IS NOT NIL)
A-Bereich	<ul style="list-style-type: none"> • = (gleich) • != (nicht gleich) • > (größer als) • < (kleiner als) • >= (größer oder gleich) • <= (kleiner oder gleich) • ZWISCHEN (BETWEEN) • IST NULL (ISNIL) • IST NICHT NULL (IS NOT NIL) • ELEMENT VON (ONE OF) • KEIN ELEMENT VON (NOT ONE OF)
A-Satz	<ul style="list-style-type: none"> • = (gleich) • != (nicht gleich) • IST NULL (ISNIL) • IST NICHT NULL (IS NOT NIL) • ELEMENT VON (ONE OF) • KEIN ELEMENT VON (NOT ONE OF)

Tabelle 6. Verfügbare Operatoren (Forts.).

Werttyp:	Verfügbare Operatoren
A-Flag	<ul style="list-style-type: none"> • = (gleich) • != (nicht gleich) • IST NULL (ISNIL) • IST NICHT NULL (IS NOT NIL)

- 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. Weitere Informationen finden Sie im Thema „Exportieren und Wiederverwenden von Regeln“ auf Seite 56.

Beachten Sie, dass eine Segmentregel, wie beispielsweise Alter ZWISCHEN 50 und 20, keine Ergebnisse zurückgibt. Sie entspricht dem Ausdruck $x \leq \text{Wert}$ und $\text{Wert} < y$, der niemals wahr sein kann. Dies entspricht dem Verhalten von SQL Server. Alter ZWISCHEN 20 und 50 ist das richtige Format.

Beachten Sie auch, dass eine Segmentregel, beispielsweise Geschlecht = \$null\$, ebenfalls keine Ergebnisse zurückgibt. Wenn Sie die Felder herausfiltern wollen, deren Wert für Geschlecht leer ist, können Sie stattdessen einen Ausdruck wie Geschlecht KEIN ELEMENT VON (Männlich, Weiblich) verwenden.

Segmentregelsets

- Segmentregeln können auch zu Sätzen zusammengefasst werden. Das Segmentregelset 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 Segmentregelsets aus einer beliebigen Anwendung, die Regeln verwendet, oder aus IBM SPSS Rules Management exportieren, falls verfügbar. Weitere Informationen finden Sie im Thema „Exportieren und Wiederverwenden von Regeln“ auf Seite 56.

Definieren von Auswahlregeln

Eine Auswahlregel ist ein Satz aus mindestens einer Segmentregel oder mindestens einem Regelset, denen jeweils das Attribut *Einschließen* oder *Ausschließen* zugewiesen ist. Auswahlregeln werden in Anwendungen oft 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 Auswahlregeln verwendet 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. Weitere Informationen finden Sie im Thema „Definieren von Segmentregeln“ auf Seite 53.
- 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 ändern möchten. Für referenzierte Auswahlregelsets 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. Weitere Informationen finden Sie im Thema „Exportieren und Wiederverwenden von Regeln“ auf Seite 56.

Evaluieren von Auswahlregeln

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

- Wenn mehrere Einschlusssegmente angegeben werden, werden bei der Auswertung der Auswahlregeln 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 Ausschlusssegmente 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.
- Beim Erstellen von Auswahlregelsets in IBM SPSS Rules Management müssen alle Regeln denselben Typ aufweisen; entweder *Einschließen* oder *Ausschließen*. Einschluss- und Ausschlussregeln können nicht in einem Regelset kombiniert werden.

Exportieren und Wiederverwenden von Regeln

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 gemeinsam genutzten Regel und Änderungen an dieser Regel werden automatisch für alle Anwendungen übernommen, die diese Version verwenden.

Exportieren von Regeln

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 Regelset oder als einzelne Regeln exportieren wollen. Im Fall von Einzelregeln wird jedes Segment als eigene Regel exportiert.
- Wenn Sie beim Export mehrerer Regeln **Mit Auswahl speichern** auswählen, wird ein Auswahlregelset 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 Regelsets exportiert werden. Falls **Mit Auswahl speichern** nicht ausgewählt ist, wird ein Segmentregelset exportiert.
- Wenn Sie beim Export von einem Aggregationsregel-Editor **Mit Risikopunkten exportieren** wählen, wird ein Aggregationsregelset erstellt. Falls diese Option nicht ausgewählt ist, wird ein Segmentregelset exportiert.

Verwenden von vorhandenen Regeln

- Klicken Sie in einer Anwendung, in der Sie eine gemeinsam genutzte Regel verwenden 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 aus. 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 nicht erwartete 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(set)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" aus.

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 dringend empfohlen - insbesondere in einer Produktionsumgebung -, da spätere Änderungen an der neuesten Version ansonsten zu nicht erwarteten 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 Regelset werden verschachtelte Referenzen auf Einzelregeln beibehalten.
- Segment- und Auswahlregelsets können in den meisten Fällen wie Einzelregeln referenziert und kombiniert werden. Das kombinierte Regelset kann jedoch erst exportiert werden, nachdem eventuelle verschachtelte Regelsets in Einzelregeln umgewandelt wurden. Ein exportiertes Regelset kann daher Referenzen auf verschachtelte Regeln enthalten, *aber nicht auf verschachtelte Regelsets*.
- Aggregationsregelsets können ebenfalls referenziert, jedoch nicht mit anderen Regeln kombiniert werden. Sie dürfen keine verschachtelten Referenzen auf andere Regeln enthalten. Wenn ein Aggregationsregelset referenziert wird, ersetzt die referenzierte Regel den gesamten Inhalt des Aggregationsregel-Editors. (Falls kein Aggregationsregelset referenziert wird, kann der Aggregationsregel-Editor lokale Regeln, referenzierte Segmentregeln und Segmentregelsets gleichermaßen enthalten.)
- Für IBM SPSS Rules Management gelten die gleichen Einschränkungen wie für exportierte Regeln. Verschachtelte Regeln sind zulässig, verschachtelte Regelsets dagegen nicht. Der Grund ist, dass IBM SPSS Rules Management für die Erstellung und Bearbeitung gemeinsam genutzter Regeln entwickelt wurde (ähnlich dem Export aus einer der anderen Anwendungen).
- Um den Einschränkungen für referenzierte Regelsets auszuweichen, ob in IBM SPSS Rules Management oder anderswo, können Regelsets immer in einzelne lokale Regeln umgewandelt werden.
- Gemeinsam genutzte Regeln können keine Referenzen auf externe Regeln von IBM Operational Decision Management enthalten.

Hinzufügen von Anmerkungen

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** verwendet 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 IBM Operational Decision Management

Regeln, die in einem Business Rule Management System wie IBM Operational Decision Management erstellt wurden, können in Anwendungen von IBM Analytical Decision Management 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 standardmäßig 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.
- Datenquellen der Analysedatenansichten unterstützen die Verwendung komplexer (hierarchischer) externer Regeln. Um eine komplexe externe Regel referenzieren zu können, muss es sich bei dem Projekt-datenmodell um eine *Datenansichtsquelle* handeln und die Funktion muss aktiviert sein (auf der Registerkarte "Daten" muss ein Administrator die entsprechende Option **Verwendung von ODM-Regeln auf der Basis hierarchischer Daten zulassen** auswählen).
- Sie können in IBM Analytical Decision Management keine gemeinsam genutzte Regel erstellen, die eine Referenz auf eine externe Regel enthält.

Kapitel 6. Erstellen von Vorhersagemodellen

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 codieren, 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 Analysedaten 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. potenzielle 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, für die 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.

Hinweis: Wenn **Parameter** zur Anwendung (Streamdatei) in IBM SPSS Modeler hinzugefügt wurden, steht auf den Registerkarten "Modellierung" und "Scoring" der Link **Parameter** zur Verfügung. Wenn Sie auf den Link **Parameter** klicken, werden die definierten Parameter sowie deren aktuelle Werte angezeigt. Sie können die Parameterwerte ändern, um die Modellerstellung oder das Scoring zu beeinflussen.

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 verwendenden 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 Flagfeld mit der Bezeichnung *Antwort* ist, gibt das Modell ein Feld mit der Bezeichnung *XF-Antwort* zurück, das die für dieses Feld vorhergesagten Werte enthält.

1. Geben Sie eine Datenquelle an. Dies kann eine analytische Datenquelle sein, in der das Ergebnis, das Sie vorhersagen möchten, aufgezeichnet ist.
2. 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.
Alle anderen Felder werden automatisch als Eingabe aufgenommen und können für die Vorhersage dieses Werts nützlich sein.
3. Legen Sie die gewünschten optionalen Einstellungen fest. Weitere Informationen finden Sie im Thema „Optionale Modelleinstellungen“ auf Seite 65.
4. Klicken Sie gegebenenfalls auf das Symbol **Datenübersicht**, um eine Übersicht der Daten abzurufen, die zum Erstellen des aktuellen Modells verwendet werden. Weitere Informationen finden Sie im Thema „Datenübersicht“ auf Seite 45.
5. 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. Weitere Informationen finden Sie im Thema „Galerie“ auf Seite 28. Falls die Modellerstellung länger als erwartet dauert, klicken Sie auf die Schaltfläche **Stopp**, um zur zuletzt gespeicherten Version zurückzukehren.

Nach Abschluss der Modellerstellung werden die Ergebnisse angezeigt. Sie können entweder die Option **Kombinierte Modellergebnisse** (Leistungsdiagramme für das soeben erstellte kombinierte Modell) oder die Option **Einzelne Modellergebnisse** (allgemeine Details der einzelnen Modelle, die im kombinierten Modell enthalten sind) auswählen.

Wenn Sie die Option *Einzelne Modellergebnisse* auswählen, verbirgt sich hinter dem Namen der einzelnen Modelle ein Link, der ein neues Dialogfeld öffnet. Dieses enthält Informationen zum Modell, wie z. B. das Zielfeld und die Eingabefelder. Wenn es sich bei dem ausgewählten Modell um ein *Baummodell* handelt, wird zudem eine Strukturbaumanzeige-funktion angezeigt.

6. Optional können Sie mit der Auswertungs- und Testfunktion kontrollieren, welche Wirkung das Modell auf Ihre Testdaten hat. Weitere Informationen finden Sie im Thema „Auswerten von Modellen“ auf Seite 69.

Ergebnisse der Vorhersagemodellierung

Wenn ein Vorhersagemodell fertiggestellt wird, können Sie im Ergebnisabschnitt zwischen der Anzeige für **Kombinierte Modellergebnisse** und **Einzelne Modellergebnisse** wechseln. Die kombinierten Modellergebnisse enthalten Verteilungs- und Variableneinflussdiagramme für das erstellte kombinierte Modell, die eine schnelle visuelle Einschätzung der Leistungsfähigkeit des Modells sowie der Bedeutung der verschiedenen Prädiktoren für das Endergebnis ermöglichen. Die einzelnen Modellergebnisse listen Informationen zu allen einzelnen Modellen auf, aus denen das kombinierte Modell besteht, wobei Sie auf jedes einzelne Modell klicken können, um die entsprechenden Details anzuzeigen.

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 Setziele wird die Verteilung in Form eines Balkendiagramms mit einem Balken für jede Kategorie des Zielfelds angezeigt. Die Höhe eines Balkens gibt an, wie oft das jeweilige Ergebnis vom Modell vorhergesagt wird. Die farbigen Bereiche der einzelnen Balken drücken aus, wie oft 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.

Prädiktoreinflussdiagramme

Normalerweise ist es sinnvoll, sich bei der Modellierung auf die wichtigsten Prädiktorfelder zu konzentrieren und die Felder zu verwerfen bzw. zu ignorieren, die die geringste Bedeutung haben. Das Prädiktoreinflussdiagramm 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 Prädiktoreinflusses kann erheblich länger dauern als die Modellerstellung, insbesondere bei großen Datensets.

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.

Einzelne Modellergebnisse

Einzelne Modellergebnisse zeigen Informationen zu jedem einzelnen Modell an, das im kombinierten Modell enthalten ist. Die Namen der einzelnen Modelle sind als Link definiert, über den ein neues Dialogfeld geöffnet wird. Dieses Dialogfeld enthält vollständige Details zum Modell, wie z. B. das Zielfeld und die Eingabefelder. Wenn es sich bei dem ausgewählten Modell um ein *Baummodell* handelt, wird zudem eine spezielle Strukturbaumanzeigefunktion angezeigt. Die Strukturbaumanzeigefunktion ist interaktiv und stellt die folgende Funktionalität bereit:

- Vergrößern und Verkleinern der Baumstruktur
- Anzeigen der Baumstrukturelemente als Tabellen und/oder Diagramme
- Ändern der Ausrichtung der Baumstruktur (von unten nach oben, von links nach rechts oder von rechts nach links)
- Verwenden des *Baumstrukturfensters* und der Pfeilsymbole für eine einfachere Navigation in großen Baumstrukturen

Erstellen eines Vorhersageregelmodells (interaktiv)

Mit Vorhersageregelmolellen (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 Auswahlregeln mit **Einschluss-** und **Ausschlussattributen** ähnlich, die angeben, wie die verschiedenen Segmente gescort wurden.

Vorhersageregelmolelle (interaktiv) sind Auswahlregeln ähnlich, werden aber auf andere Weise gescort. Im Fall von Auswahlregeln werden ausgeschlossene Datensätze entfernt, sodass 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. Ausgeschlossene Segmente werden als *NULL* gescort, ebenso wie alle Datensätze, die dem Rest zugeordnet werden, da sie keiner Regel entsprechen.

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. Vorhersageregelmuster (interaktiv) sind somit gut zur Erstellung von Mailing-Listen geeignet und im Customer Relationship Management weit verbreitet, unter anderem in Call-Center- und Marketing-Anwendungen.

- Die Vorhersageregelmuster (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 Auswahlregeln 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 Vorhersageregelmuster (interaktiv) ab:

1. Geben Sie eine Datenquelle an. Dies kann eine analytische Quelle sein, in der das Ergebnis, das Sie vorhersagen möchten, erfasst wird.
2. Legen Sie ein Zielfeld fest, welches das Ergebnis erfasst, das Sie vorhersagen wollen.
3. Geben Sie den gesuchten Wert oder die gesuchte Antwort an. Wenn der Name des Zielfelds beispielsweise *Antwort* ist, könnten Sie *Wahr* oder *Falsch* auswählen, um nach Kunden zu suchen, die das Angebot angenommen oder abgelehnt haben.
Beispielsweise kann das Zielfeld angeben, welche Kunden gekündigt haben, auf frühere Angebote eingegangen sind, Schadensmeldungen eingereicht haben oder Ähnliches. Alle anderen Felder werden automatisch als Eingaben aufgenommen, die sich bei der Vorhersage dieses Werts als hilfreich erweisen könnten.
4. Legen Sie die gewünschten optionalen Einstellungen fest. Weitere Informationen finden Sie im Thema „Optionale Modelleinstellungen“ auf Seite 65.
5. Klicken Sie gegebenenfalls auf das Symbol **Datenübersicht**, um eine Übersicht der Daten abzurufen, die zum Erstellen des aktuellen Modells verwendet werden. Weitere Informationen finden Sie im Thema „Datenübersicht“ auf Seite 45.
6. Klicken Sie auf **Erstellung beginnen**.
7. Geben Sie an, ob Sie Segmente mit hoher oder niedriger Wahrscheinlichkeit in Bezug auf das angegebene Ziel suchen möchten. Sie könnten z. B. nach Gruppen oder Kunden suchen, die mit der höchsten oder geringsten Wahrscheinlichkeit kündigen, und sie dementsprechend einschließen.
8. Geben Sie die maximale Anzahl zu suchender Segmente an. Üblicherweise wird aus Gründen der Einfachheit und Leistung eine möglichst niedrige Zahl gewählt.
9. Klicken Sie auf **Modell erweitern**, um eine Liste mit ein oder mehreren Segmentregeln zu erhalten, die Gruppen mit höherer oder geringerer Wahrscheinlichkeit aufzeigen.
10. Klicken Sie auf das Symbol **Einschließen** bzw. **Ausschließen**, 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.
11. Optional können Sie Segmente nach Bedarf manuell hinzufügen und bearbeiten. Weitere Informationen finden Sie im Thema „Definieren von Segmentregeln“ auf Seite 53.

- Optional können Sie mit der Auswertungs- und Testfunktion kontrollieren, welche Wirkung das Modell auf Ihre Testdaten hat.

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 Clustering-Modellen können Datensätze in eine bestimmte Anzahl von Clustern unterteilt werden. So können Sie natürliche Gruppen in Ihren Daten bestimmen.

Clustering-Modelle konzentrieren sich auf die Ermittlung ähnlicher Datensätze und auf die 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 von Clustering-Modellen zu anderen Techniken des Maschinenlernens: Es gibt kein vordefiniertes Ausgabe- oder 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.

Clustering-Methoden 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.

Clustering-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:

1. Geben Sie eine Datenquelle an. Dies kann eine analytische Quelle sein, in der das Ergebnis, das Sie vorhersagen möchten, erfasst wird.
2. Legen Sie die gewünschten optionalen Einstellungen fest. Weitere Informationen finden Sie im Thema „Optionale Modelleinstellungen“ auf Seite 65.
3. Klicken Sie gegebenenfalls auf das Symbol **Datenübersicht**, um eine Übersicht der Daten abzurufen, die zum Erstellen des aktuellen Modells verwendet werden. Weitere Informationen finden Sie im Thema „Datenübersicht“ auf Seite 45.
4. Klicken Sie auf **Cluster suchen**.
5. Optional können Sie manuelle Cluster hinzufügen. Weitere Informationen finden Sie im Thema „Verwenden von manuellen Clustern“ auf Seite 64.
6. Optional können Sie mit der Auswertungs- und Testfunktion kontrollieren, welche Wirkung das Modell auf Ihre Testdaten hat.

Verwenden von manuellen Clustern

Bei Bedarf können Cluster manuell definiert werden. Im Anschluss an die Verwendung der manuellen Cluster wird dann die automatisierte Clusterbildung ausgeführt, um nach Clustern von Datensätzen zu suchen, die von keinem der manuell definierten Cluster erkannt wurden. Erstellen Sie manuelle Cluster über ein vom ersten Treffer ausgehendes segmentbasiertes Zuordnungsregelset. 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.

1. 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.
2. Ändern Sie bei Bedarf die Beschriftungsnamen der Cluster.
3. 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 <= -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 Standardentscheidungs bäume (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 Dataset verglichen werden. Die "guten" Regeln werden gespeichert. Anschließend werden alle Regeln, nachdem sie verschiedenen Beschränkungen unterworfen wurden, weiter spezialisiert. **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 Indizierungsschemas verwendet, um den potenziell großen Suchbereich zu reduzieren.

So erstellen Sie ein Assoziationsmodell:

1. Wählen Sie eine Datenquelle aus. Dies kann eine analytische Quelle sein, in der das Ergebnis, das Sie vorhersagen möchten, erfasst wird.
2. Geben Sie ein Datenformat an. Weitere Informationen finden Sie im Thema „Vergleich von Tabellen- und Transaktionsdaten“ auf Seite 65.
3. Legen Sie die gewünschten optionalen Einstellungen fest. Weitere Informationen finden Sie im Thema „Optionale Modelleinstellungen“ auf Seite 65.

4. Klicken Sie gegebenenfalls auf das Symbol **Datenübersicht**, um eine Übersicht der Daten abzurufen, die zum Erstellen des aktuellen Modells verwendet werden. Weitere Informationen finden Sie im Thema „Datenübersicht“ auf Seite 45.
5. Klicken Sie auf **Modell erstellen**.
6. Optional können Sie mit der Testfunktion prüfen, welche Wirkung das Modell auf Ihren Testdaten hat. Beachten Sie, dass für Assoziationsmodelle keine Evaluierungsfunktion verfügbar ist.

Bei anderen Modelltypen können Sie zwischen verschiedenen Modellierungsverfahren wählen. Für Assoziationsmodelle wird der Modelltyp *Apriori* verwendet.

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 Dropdown-Menüs 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 **Kassenrollenformat** bezeichnet.

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

Tabellendaten

In Tabellendaten (auch als **Warenkorb-** oder **Wahrheitstabellendaten** bezeichnet) werden die Elemente durch gesonderte Flags dargestellt, wobei jedes Flagfeld für das Vorliegen bzw. die Abwesenheit eines bestimmten Elements steht. Jeder Datensatz steht für ein komplettes Set zugehöriger Elemente. Prinzipiell können Flagfelder 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. Beachten Sie, dass je nach Modelltyp und Ziel möglicherweise nicht jede dieser Einstellungen angezeigt wird. Änderungen dieser Einstellungen treten bei einem vorhandenen Modell erst in Kraft, nachdem es neu erstellt wurde.

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 inaktiviert werden, u. a. für solche, deren Datenvorbereitung von einem erfahrenen Anwender speziell angepasst wurde.

Anmerkung: Die Einstellungen für die automatische Datenaufbereitung gelten nur beim Erstellen eines neuen Modells. Eine Klassierung erfolgt nicht.

Daten automatisch partitionieren, um die Modellauswertung von Entwicklungsdatenquellen zu ermöglichen (für weitere Evaluierung und Tests). Bei Auswahl dieser Option werden die Daten in separate Subsets oder Stichproben zum Trainieren und Testen des Modells aufgeteilt. Wenn Sie das Modell auf der Grundlage eines Subsets aufbauen und anhand eines anderen Subsets testen, erhalten Sie eine ungefähre Vorstellung davon, inwieweit es sich auf andere Datasets verallgemeinern lässt. Sie können außerdem einen Prozentsatz für die zufällig generierte *Trainingspartitionsgröße* und die *Testpartitionsgröße* angeben.

Die minimale Trainingspartitionsgröße ist 1, die maximale 100. Die minimale Testpartitionsgröße ist 0, die maximale 100. Die minimale Validierungspartitionsgröße ist 0, die maximale 100. Der Gesamtprozentsatz darf maximal 100 sein, damit das Modell erfolgreich erstellt werden kann.

Diese Partitionierungsoptionen sind für alle Modelltypen mit Ausnahme von Assoziationsmodellen verfügbar.

Verfahren für Modellerstellung auswählen. Klicken Sie bei Bedarf auf **Auswählen**, um die Modellierungsverfahren auszuwählen, die beim Erstellen des Modells einbezogen werden sollen. Alle verfügbaren Modellierungsverfahren werden basierend auf dem ausgewählten Ziel aufgelistet. Die verfügbaren Modellierungsverfahren (und die standardmäßig ausgewählten Verfahren) können je nach ausgewähltem Ziel unterschiedlich sein. Beispielsweise wird der Algorithmus *Entscheidungsliste* nur für *Flagziele* angezeigt. Diese Funktion ist nur für Vorhersagemodelle verfügbar.

Ausführliche Details zu den Modellierungsverfahren finden Sie im *IBM SPSS Modeler Algorithms Guide* und in der weiteren Dokumentation, die zum Lieferumfang des IBM SPSS Modeler-Produkts gehört und im Internet verfügbar ist.

Maximale Anzahl der zu kombinierenden Modelle. Ermöglicht es Ihnen, die maximale Anzahl von Modellen festzulegen, die gespeichert und kombiniert werden sollen. Wenn Sie für diese Option 1 festlegen, wird nur ein einziges Modell erstellt und gespeichert. Wenn der Wert der erstellten Modelle geringer ist als der hier eingegebene Wert, werden alle erstellten Modelle gespeichert und kombiniert. Diese Option ist nur für Vorhersagemodelle verfügbar.

Profitkriterien (zum Einstufen von Modellen mit binären Zielen). Ermöglicht es Ihnen, Profitkriterien vor dem Erstellen des Modells festzulegen. Die hier festgelegten Werte werden als Standardwerte für die Modellauswertung festgelegt. Diese Option ist nur für Vorhersagemodelle mit einem ausgewählten Flagziel verfügbar.

Zu verwendende Eingaben festlegen. Ermöglicht es Ihnen, die Felder auszuwählen, die Sie verwenden möchten. Meist werden dies die Felder sein, die einen praktischen Bezug zum Gegenstand der Prognose haben, z. B. Alter oder Einkommen. Wenn Sie über große Datasets 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 verknüpfted Eingabefeld (einen Ausdruck) klicken, wird der Ausdrucksvierer für diesen Ausdruck angezeigt. Zum Bearbeiten eines Ausdrucks verwenden Sie die Registerkarte "Daten". Weitere Informationen finden Sie im Thema „Expression Builder“ auf Seite 49.

Zu verwendende Auswahlen festlegen. 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. Weitere Informationen finden Sie im Thema „Definieren von Auswahlregeln“ auf Seite 55.

Wenn globale Auswahlen 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 Auswahlregeln verwendet werden sollen. Es ist nicht möglich, nur ein Subset davon auszuwählen.

Segmentierungsoptionen

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

Minimale Segmentgröße. 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 ist 0, die zulässige Maximaleinstellung 99,9.
- **Absolutwert.** Legt die minimale Gruppengröße als absolute Anzahl der Datensätze fest. Die zulässige Minimaleinstellung ist 1, eine Maximaleinstellung ist nicht vorhanden.

Maximale Anzahl an Attributen. Legt die maximale Anzahl von Bedingungen pro Segmentregel fest. Die zulässige Minimaleinstellung ist 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 Regel "Anzahl der Bedingungen pro Segment". 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.

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

Gewünschten Bereich für die Anzahl der gefundenen Cluster festlegen. 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.
- **Daten sind bereits nach ID sortiert (nur Transaktionsformat)**. Wählen Sie diese Option aus, um die standardmäßige Sortierung beim Start der Modellerstellung zu inaktivieren, 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 Vorhersagen**. 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 Elementset auszugeben. Folgende Kriterien stehen zur Verfügung:
 - **Konfidenz** zeigt das Verhältnis von Regelunterstützung zu Regelbedingungsunterstützung an. Es wird also der Anteil der Datensätze mit der/den festgelegten Regelbedingung(en) angegeben, bei der/denen auch die Regelprognose(n) wahr ist/sind. Wenn beispielsweise 50 % der Trainingsdaten "Brot" (Regelbedingungsunterstützung), aber nur 20 % sowohl "Brot" als auch "Käse" (Regelunterstützung) enthalten, ist die Vorhersage für die Regel Brot -> Käse Regelunterstützung / Regelbedingungsunterstützung, in diesem Fall 40 %.
 - **Abdeckungsbereich (%)** zeigt die Regelbedingungsunterstützung 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 Regelbedingungsunterstützung 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. Wenn beispielsweise 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 % vorhersagt, dass die Leute Brot kaufen, weist die Regel einen Lift von annähernd 1 auf. Dies 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 Bereitstellbarkeitsstatistik ist folgendermaßen definiert: $([\text{Regelbedingungsunterstützung in Anzahl der Datensätze} - \text{Regelunterstützung in Anzahl der Datensätze}] / \text{Anzahl der Datensätze}) * 100$. Dabei steht Regelbedingungsunterstützung für die Anzahl der Datensätze, bei denen die Regel-

bedingungen wahr sind und Regelunterstützung für 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 vorhandene 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.
- **Wiederholte Vorhersagen 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 & Käse -> Wein
Käse & Obst -> Wein
Inaktivieren Sie diese Option, um beim Scoring Wiederholungsprognosen auszuschließen.
- **Elemente ohne Entsprechung ignorieren.** Wählen Sie diese Option aus, um das Vorliegen zusätzlicher Elemente im Elementset zu ignorieren. Wenn diese Option beispielsweise für einen Warenkorb ausgewählt wird, der [Zelt & Schlafsack & Wasserkessel] enthält, gilt die Regel Zelt & Schlafsack -> Gaskocher trotz des zusätzlichen Elements (Wasserkessel) im Warenkorb.

Auswerten von Modellen

Klicken Sie nach dem Erstellen oder Öffnen eines abgeschlossenen Modells auf **Auswerten**, um Verteilungs-, Fehlerverteilungs-, Fehlerklassifizierungstabellen-, Gewinn-, Lift-, Treffer-, Profit- und ROI-Diagramme abzurufen und den maximal möglichen Profit zu simulieren. Welche Diagrammtypen und welche anderen Optionen verfügbar sind, ist abhängig vom ausgewählten Modelltyp. Das Dialogfeld "Auswerten" steht nur zur Verfügung, wenn das ausgewählte Modell abgeschlossen ist. Beachten Sie, dass für Assoziationsmodelle keine Evaluierungsfunktion verfügbar ist.

1. Wählen Sie im oberen Bereich ein Dataset für die Auswertung aus. Wie bei der Modellerstellung muss dies ein analytisches Dataset sein, dessen tatsächliche Antworten bekannt sind. (Andernfalls wäre es nicht möglich, tatsächliche Antworten mit den vom Modell vorhergesagten Antworten zu vergleichen.)
2. Wenn Sie die Modelle so ändern möchten, dass sie in der Auswertung enthalten sind, klicken Sie auf **Modellauswahl ändern**. Nachdem Sie Ihre Auswahl getroffen haben, können Sie bei Bedarf auch auf die folgende Schaltfläche klicken: **Aktuelle Modellauswahl wieder in die Tabelle mit den einzelnen Modellergebnissen kopieren**. Diese Optionen sind nur für den Modelltyp "Vorhersage" verfügbar.



Abbildung 13. Aktuelle Modellauswahl wieder in die Tabelle mit den einzelnen Modellergebnissen kopieren

3. Wählen Sie die Diagramme aus, die Sie erstellen möchten.
4. Wenn Sie den Profit simulieren, wählen Sie das Feld **Maximalen Profit simulieren** aus und geben die Population ein, auf der die Profitberechnung, das Budget und der Mindestwert basieren sollen. Weitere Informationen finden Sie im Thema „Simulieren des maximalen Profits“ auf Seite 71.

5. 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.
6. 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.
7. Wenn Sie ein Leistungsdiagramm anfordern, können Sie auch die Option **Einzelne Modellevaluierungen für Leistungsdiagramme einschließen** auswählen.
8. Wählen Sie bei Bedarf **Tabellarische Ergebnisse aufnehmen** aus. Dann wird mehr Zeit für die Rückgabe der Auswertungsergebnisse benötigt.
9. 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 Datasets als das Testdataset. Für Bereichsziele ist das Verteilungsdiagramm der einzige verfügbare Diagrammtyp.

Sie können die folgenden weiteren Diagramme auswählen:

Fehlerverteilungsdiagramme

Das Fehlerverteilungsdiagramm zeigt ein Histogramm der Fehler in den Modellergebnissen an, wobei der Fehler den Unterschied zwischen dem vorhergesagten und dem tatsächlichen Wert darstellt. Ein positiver Fehler entsteht, wenn der vorhergesagte Wert über dem tatsächlichen Wert liegt. Ein negativer Fehler entsteht, wenn er unter dem tatsächlichen Wert liegt. Dieses Diagramm ist nur für stetige Ziele mit dem Modelltyp "Vorhersage" verfügbar.

Fehlerklassifizierungstabellendiagramme

In einer Fehlerklassifizierungstabelle wird die Anzahl der *vorhergesagten* Ergebnisse im Vergleich zu den *tatsächlichen* Ergebnissen dargestellt. Bei einem numerischen Ziel ist der Bereich in die Anzahl der angegebenen Klassen aufgeteilt.

Gewinndiagramme

Gewinndiagramme werden verwendet, um den durch den Einsatz des Modells erhaltenen "Gewinn" oder "Lift" darzustellen. Der "Gewinn" 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 Rücklaufquote 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 Gewinn.

Liftdiagramme

Das Liftdiagramm vergleicht den Prozentsatz der in jedem Inkrement enthaltenen Datensätze, bei denen es sich um Treffer handelt, mit dem Prozentsatz aller Treffer im Trainingsdataset.

Trefferdiagramme

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

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. Weitere Informationen finden Sie im Thema „Simulieren des maximalen Profits“.

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.

Simulieren des maximalen Profits

Auf Grundlage der vom Modell vorhergesagten Rücklaufquote 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.

So sagen Sie den maximalen Profit vorher:

1. Geben Sie die Population oder Gesamtzahl der Datensätze an, auf die sich die Vorhersage stützen soll.
Hinweis: Beim Scoring eines Modells wird automatisch die Gesamtzahl der Datensätze in der Datenquelle oder im Subset festgelegt.
2. Geben Sie einen Wert für das Budget an, wenn der Budgetparameter in die Profitsimulation einbezogen werden soll. Geben Sie auch einen Mindestwert an.
3. 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.
4. Geben Sie den erwarteten Umsatz an, den eine positive Antwort generieren würde. Beispiel: Der Einzelhandelspreis eines Artikels einer Kampagne.
5. Wählen Sie die Antwort aus, an der Sie interessiert sind.
6. 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.

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 hierzu auf **Zum Scoring verwenden**. Der Prozentsatz der Datensätze, die in die Zielgruppe aufgenommen werden sollen, wird im Bereich "Neigungsoptionen" der Registerkarte "Scoring" im Feld **Obere** % angezeigt.

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 Rücklaufquote, 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.

Hinweis: Dieses Diagramm wird bei der ersten Ausführung der Auswertung erstellt. Bei Änderungen, die Sie über den Schieberegler im Profitsimulationsbereich vornehmen, wird es nicht automatisch geändert.

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. Im Folgenden finden Sie einige Beispiele für Ergebnistypen, die je nach Anwendungstyp angezeigt werden können:

- Sie können beispielsweise einige ausgewählte Kunden betrachten und ablesen, welche Angebote ihnen jeweils zugeordnet sind.
 - Sie können ausgewählte Kundendatensätze betrachten und ablesen, welche Kunden mit hoher Wahrscheinlichkeit auf einen Werbebrief reagieren würden.
 - Sie könnten ausgewählte Versicherungsschäden betrachten und ablesen, welche Aktionen ihnen jeweils zugeordnet sind.
 - Sie können ausgewählte Fälle betrachten und ablesen, wie die einzelnen Regeln ausgewertet werden.
1. Klicken Sie auf der Registerkarte "Modellierung" auf **Testen**. Das Dialogfeld "Testen" wird angezeigt.
 2. Wählen Sie die Datenquelle, die Sie verwenden möchten, oder stattdessen im Abschnitt "Datensätze" die Option **Benutzerdefinierte Daten auswählen** aus, wenn Sie die Testdatensätze manuell eingeben wollen.
 3. Legen Sie gegebenenfalls fest, welche Auswahlen 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. Weitere Informationen finden Sie im Thema „Definieren von Auswahlregeln“ auf Seite 55.
 4. Wählen Sie bei Verwendung einer Datenquelle die Anzahl der Datensätze aus, die zurückgegeben werden sollen. Die ersten N bis maximal 100 Datensätze können zurückgegeben werden. Diese werden entweder vom Anfang der Daten oder aus einer Zufallsstichprobe von Datensätzen gelesen. Wenn Sie ein Assoziationsmodell mithilfe von Transaktionsdaten testen, wird die Stichprobe der Eingabedaten auf der Transaktionsebene nicht auf der Datensatzebene erstellt.
 5. Wenn Sie benutzerdefinierte Daten verwenden, wählen Sie **Benutzerdefinierte Daten auswählen** aus 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.
 6. Wählen Sie unter **Anzuzeigende Felder auswählen** die Felder aus, die in der Testausgabe angezeigt werden sollen.
 7. Klicken Sie auf **Ausführen**. In der Testzusammenfassung wird eine Liste mit erfolgreichen Zuordnungen für jeden Datensatz angezeigt.
 8. Über das Informationssymbol am Anfang einer Zeile erhalten Sie Details zur jeweiligen Zeile.

Verwenden von Modellscores in Anwendungen

Beim Scoring eines Modells wird das Modell auf bestimmte Daten oder eine bestimmte Population 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 Datenbank-

tabelle oder eine Flatfile geschrieben oder als Eingaben für die Segment-, Auswahl und Zuordnungsregeln verwendet werden, die die Entscheidungen in einer Anwendung steuern.

Modellscores können auf folgende Weisen verwendet werden:

- Als Eingabe für Regeln. Weitere Informationen finden Sie im Thema „Erstellen von Regeln“ auf Seite 53.
- Für Anwendungen, die Stapelscoring unterstützen, können Scores in eine Datenbanktabelle oder eine Flatfile geschrieben werden. Weitere Informationen finden Sie im Thema „Scoring von Modellen in eine Datenbanktabelle, eine Datei, eine Analytic Server-Instanz oder einen IBM Cognos BI-Server“ auf Seite 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. Weitere Informationen finden Sie im Thema „Messniveaus“ auf Seite 46.

Flagziele. Automatisierte Modelle mit Flagzielen 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>**. Propensity-Score; gibt die Wahrscheinlichkeit des Wahr-Werts für jeden Datensatz an. Propensity-Scores unterscheiden sich von Konfidenzscore, 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 Propensity-Score, da dieser eine hohe Wahrscheinlichkeit dafür angibt, dass *keine* Antwort erfolgt.

Kategoriale Ziele. Automatisierte 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. Automatisierte Modelle mit stetigen numerischen Zielen geben die folgenden Felder aus:

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

Vorhersageregelmuster (interaktiv). Vorhersageregelmuster (nur Flagziele 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. Weitere Informationen finden Sie im Thema „Erstellen eines Vorhersageregelmusters (interaktiv)“ auf Seite 61.
- **\$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>**. Propensity-Score; gibt die Wahrscheinlichkeit der gewünschten Antwort an. Nur für Flagziele verfügbar.

Kapitel 7. Scoren von Modellen

Scoring von Modellen in eine Datenbanktabelle, eine Datei, eine Analytic Server-Instanz oder einen IBM 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 zum Erstellen eines Modells verwenden, das vorhersagt, welche Kunden am wahrscheinlichsten auf eine Kampagne reagieren. Anschließend könnten Sie mit dem Modell Datensätze für Personen scoren, die derzeit nicht zu Ihren Kunden gehören, die aber möglicherweise für ein Werbemailing in Frage kommen.

Folgende grundlegende Arbeitsschritte müssen zum Scoren eines Modells durchgeführt werden:

1. Wählen Sie die Datenquelle sowie alle Subsets dieser Daten aus, die gescort werden sollen.
2. Wählen Sie das Ziel für die gescorten Daten aus.
3. Wählen Sie die auszugebenden Informationen sowie die Vorgehensweise beim Scoren der Informationen aus.
4. Geben Sie an, welche Datensätze gescort werden sollen. Wählen Sie in der Scoring-Dropdown-Liste eine der folgenden Optionen zum Scoren aus: **Alle Datensätze**, die **Oberen N %**, die **Oberen N Datensätze**, **Min./Max. Neigung** (nur für *Flagziele* verfügbar) oder **Minimalscore** (nur für *stetige* Ziele verfügbar).
5. Rufen Sie bei Bedarf eine Vorschau der erwarteten Ergebnisse auf und erstellen Sie eine Datenübersicht.
6. Klicken Sie auf **Jetzt scoren**, um das Modell gemäß Ihren Einstellungen zu scoren. Prüfen Sie nach Abschluss des Vorgangs den Abschnitt "Scoring-Zusammenfassungsbericht". Weitere Informationen finden Sie im Thema „Scoring-Zusammenfassungsbericht“ auf Seite 82.

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*. Weitere Informationen finden Sie im Thema „Auswählen von Ausgabefeldern“ auf Seite 81.

Anmerkung: Wenn die Quelldaten aus einer Datenbank stammen, wird die Ausführung an die Datenbank zurückgegeben, wann immer dies möglich ist (in einem Prozess, der als "SQL-Pushback" bezeichnet wird). Im Allgemeinen unterstützt IBM Analytical Decision Management dieselbe SQL-Pushback-Stufe wie IBM SPSS Modeler, jedoch mit den folgenden Einschränkungen:

- Wenn das Projekt zusätzliche Felder oder Auswahlregeln enthält, die Referenzen auf Modelle oder externe Regeln aufweisen, werden diese Referenzen nicht an die Datenbank zurückgegeben und dadurch kann verhindert werden, dass der für den Projektstream als Ganzes erreichbare Pushback-Grad erreicht wird.
- In anderen Anwendungen als IBM SPSS Modeler Advantage wird der Pushback nur bis zu Auswahlregeln (globalen Auswahlen und Mitgliedsauswahlen in der ersten Dimension) aktiviert. Der Pushback-Grad kann erhöht werden, wenn Sie ein Feld für ganze Zahlen als Entitäts-ID im Projektdatenmodell auswählen.

IBM SPSS Modeler Advantage

Bei **IBM SPSS Modeler Advantage** enthält der gespeicherte Stream zwei Verzweigungen: die Buildverzweigung und die Scoring-Verzweigung.

Der Quellenknoten der Buildverzweigung basiert immer auf der Datenquelle des Projektdatenmodells. Er enthält Felder, die entweder als betriebsbereit oder als nicht betriebsbereit markiert sind.

Der Quellenknoten für die Scoring-Verzweigung basiert ebenfalls immer auf der Datenquelle des Projektdatenmodells. Er enthält jedoch nur Felder, die als betriebsbereit markiert sind. Felder, die als nicht betriebsbereit markiert sind, werden nicht einbezogen, da sie für das Scoring nicht erforderlich sind. Das Zielfeld eines Vorhersagemodells ist z. B. normalerweise als nicht betriebsbereit markiert und ist nicht als Scoring-Eingabe erforderlich.

Die Scoring-Verzweigung des gespeicherten Streams weist standardmäßig einen Endknoten **Tabelle** auf. Dieser Tabellenknoten wird verwendet, um die Scoring-Verzweigung in den Bereitstellungseinstellungen des Streams anzugeben. Diese Informationen werden vom Scoring-Service von IBM SPSS Collaboration and Deployment Services zur Ermittlung des Endknotens verwendet, der bei der Ausführung des Echtzeitscorings verwendet werden soll.

Die Registerkarte **Scoren** (sofern konfiguriert) enthält einen Abschnitt für **Stapelscoring**. Unter **Stapelscoring** befindet sich ein Abschnitt für **Optionen zum Scoren**. Diese Einstellungen gelten nur, wenn das Scoring innerhalb der Benutzerschnittstelle von IBM Analytical Decision Management ausgeführt wird, indem auf die Schaltfläche **Jetzt scoren** geklickt wird. Sie werden im generierten Stream, der gespeichert wird, nicht widergespiegelt. Daher gelten sie nicht beim Scoren über den IBM SPSS Modeler-Client oder beim Planen eines Jobs mithilfe des Clients von IBM SPSS Collaboration and Deployment Services Deployment Manager.

Das Stapelscoring verfügt auch über Steuerelemente zum Auswählen der Ziel- und Ausgabefelder. Beim Speichern des Streams wird ein zusätzlicher Endknoten für die Scoring-Verzweigung erstellt. Die Einstellungen für das Ziel- und Ausgabefeld werden zum Erstellen eines Endknotens **Exportieren** für die Scoring-Verzweigung verwendet. Diese Einstellungen gelten beim Ausführen des Streams über den Client von IBM SPSS Collaboration and Deployment Services Deployment Manager oder beim Planen eines Jobs mithilfe des Clients von IBM SPSS Collaboration and Deployment Services Deployment Manager.

Andere Projekte als Projekte von IBM SPSS Modeler Advantage

Bei Projekten, die **keine Projekte von IBM SPSS Modeler Advantage sind**, weist der gespeicherte Stream nur die Scoring-Verzweigung auf. Wie bei der Scoring-Verzweigung von IBM SPSS Modeler Advantage basiert der Quellenknoten für die Scoring-Verzweigung immer auf der Datenquelle des Projektdatenmodells. Er enthält ebenfalls nur Felder, die als betriebsbereit markiert sind. Bei der Verwendung von Sekundärtabellen werden zusätzliche Quellenknoten erstellt. Diese sekundären Quellenknoten werden von den sekundären Datenquellen definiert, die der primären Datenquelle zugeordnet sind.

Die Registerkarte **Bereitstellen** (sofern konfiguriert) kann Abschnitte für **Echtzeitscoring** und **Stapelscoring** enthalten. Diese Einstellungen sind unabhängig voneinander.

Beim Speichern des Streams werden die Einstellungen für **Echtzeitscoring** verwendet, um einen Endknoten **Tabelle** für die Scoring-Verzweigung zu erstellen. Dieser Tabellenknoten wird verwendet, um die Scoring-Verzweigung in den Bereitstellungseinstellungen des gespeicherten Streams anzugeben. Diese Informationen werden vom Scoring-Service von IBM SPSS Collaboration and Deployment Services zur Ermittlung des Endknotens verwendet, der bei der Ausführung des Echtzeitscorings verwendet werden soll.

Unter **Stapelscoring** gibt es einen Abschnitt für **Optionen zum Scoren**. Diese Einstellungen gelten nur, wenn das Scoring innerhalb der Benutzerschnittstelle von IBM Analytical Decision Management ausgeführt wird, indem auf die Schaltfläche **Jetzt scoren** geklickt wird. Sie werden im generierten Stream, der gespeichert wird, nicht widergespiegelt. Daher gelten sie nicht beim Scoren über den Scoring-Service von IBM SPSS Collaboration and Deployment Services oder beim Planen eines Jobs mithilfe des Clients von IBM SPSS Collaboration and Deployment Services Deployment Manager.

Das Stapelscoring verfügt auch über Steuerelemente zum Auswählen der Ziel- und Ausgabefelder. Beim Speichern des Streams werden diese Einstellungen zum Erstellen eines Endknotens **Exportieren** für die Scoring-Verzweigung verwendet. Diese Einstellungen gelten beim Planen eines Jobs mithilfe des Clients von IBM SPSS Collaboration and Deployment Services Deployment Manager oder beim Ausführen des Streams über den IBM SPSS Modeler-Client.

Alle Projekte

Sowohl Projekte von IBM SPSS Modeler Advantage als auch andere Projekte als Projekte von IBM SPSS Modeler Advantage können unterschiedliche Endknoten für das Echtzeitscoring und Stapelscoring erstellen. Stellen Sie beim Planen von Jobs mithilfe des Clients von IBM SPSS Collaboration and Deployment Services Deployment Manager sicher, dass Sie die Jobschritte so konfigurieren, dass nur ein bestimmter Exportknoten für das Stapelscoring ausgeführt wird.

Auswählen von Daten und Datensatzgruppen

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. Weitere Informationen finden Sie im Thema „Definieren von Auswahlregeln“ auf Seite 55.

Auswählen von Datensätzen für Scoring

Bei Modellen, die Propensity-Scores 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 Mailing-Liste erstellen. Propensity-Scores geben die Wahrscheinlichkeit des Wahr-Werts für das Ziel an und sind für Vorhersagemodelle mit Flagzielen sowie für Vorhersageregelmole (interaktiv) verfügbar. Bei Vorhersageregelmole (interaktiv) ist die ausgewählte Antwort der Wahr-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. Weitere Informationen finden Sie im Thema „Simulieren des maximalen Profits“ auf Seite 71.
- **Die Anzahl der Datensätze mit den höchsten Neigungswerten.** Standardmäßig werden die höchsten 1000 Datensätze ausgewählt (soweit verfügbar).
- **Durch Angabe der minimalen und maximalen Neigungswerte.** Zum Beispiel alle Datensätze mit einer Antwortwahrscheinlichkeit zwischen 75 % und 100 %.

Auswählen des Scoring-Ziels

Sie können die Scoring-Daten wahlweise in eine Datenbank, in eine Datei, an eine Analytic Server-Instanz oder an einen Cognos BI-Server ausgeben. Zu den unterstützten Dateitypen gehören das Textformat (txt, *.csv), das Excel-Format, das von IBM SPSS Statistics verwendete *.sav-Format, das SAS-Format sowie die von Data Collection unterstützten Formate. Weitere Informationen finden Sie im Thema „Dateiziel“ auf Seite 78.

Datenbankziel

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

1. Geben Sie in das Feld **Datenbankname** den Datenbanknamen ein oder wählen Sie die gewünschte Datenbank aus.

2. 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.
3. Ändern Sie bei Bedarf die Standardzeichenfolgröße.

Wenn Sie die Daten an eine vorhandene Datenbank anhängen oder diese überschreiben, können Sie die Scoring-Felder auf Felder des vorhandenen Ziels abbilden. Weitere Informationen finden Sie im Thema „Abilden von Scoring-Feldern“ auf Seite 80.

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. Weitere Informationen finden Sie im Thema „Abilden von Scoring-Feldern“ auf Seite 80.

Excel-Datei

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

Feldnamen aufheben. Gibt an, ob Feldnamen in die erste Zeile des Arbeitsblatts eingefügt werden sollen.

Flatfile

Wenn Sie angeben, dass die Scoring-Ausgabe an eine Flatfile gesendet werden soll, können Sie weitere Details eingeben, um den Dateiinhalte 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 aufheben.** 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** auswählen, 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** auswählen, geben Sie das oder die gewünschten Anführungszeichen in das Textfeld ein.

Verschlüsselung. Gibt die verwendete Textcodierungsmethode an. Sie haben die Wahl zwischen der Systemstandardeinstellung und UTF-8.

Dezimalzeichen. 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.

IBM SPSS Statistics-Datendatei

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

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

SAS-Datendatei

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) oder 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 aus, um sowohl Feldnamen als auch Felddbeschriftungen zu exportieren. Die Namen werden als SAS-Variablenamen exportiert, die Beschriftungen entsprechend als SAS-Variablenbeschriftungen.
- **Namen als Variablenbeschriftungen.** Wählen Sie diese Option aus, 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** aus.

Analytic Server-Ziel

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

1. Geben Sie die **Datenbankquelle** an.
2. Geben Sie an, wie die Scores gespeichert werden sollen. Sie können 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.
3. Wählen Sie Ausgabefelder aus, die in die Scoring-Ausgabe eingeschlossen werden sollen.
4. Geben Sie an, welche Datensätze gesort werden sollen. Wählen Sie in der Scoring-Dropdown-Liste eine der folgenden Optionen zum Sortieren aus: **Alle Datensätze**, die **Oberen N %**, die **Oberen N Datensätze**, **Min./Max. Neigung** (nur für *Flagziele* verfügbar) oder **Minimalscore** (nur für *stetige* Ziele verfügbar).

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.

Hinweis: Sie können nur relationale Daten, keine OLAP-Daten exportieren.

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 Kennwort 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.

1. Wenn Sie als **Ziel** den Cognos BI-Server auswählen und auf **Verbinden** klicken, werden Sie aufgefordert, die Details für die Serververbindung einzugeben.
 - a. 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.
 - b. 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 **Berechtigungsnaehweise festlegen** aus und geben Sie den Namespace, den Benutzernamen und das Kennwort 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 **Kennwort** ein, das zum angegebenen Benutzernamen gehört.

ODER

Navigieren Sie zu einem im IBM SPSS Collaboration and Deployment Services Repository gespeicherten IBM Cognos-Berechtigungsnaehweise und wählen Sie ihn aus.

- c. Um sich stattdessen ohne Berechtigungsnaehweise anzumelden (bei dieser Variante können die Felder für Namespace, Benutzernamen und Kennwort nicht ausgefüllt werden), verwenden Sie die Option **Anonyme Verbindung verwenden**.

Hinweis: Bei einigen Serververbindungen sind anonyme Verbindungen nicht zulässig.

2. Wählen Sie die **Datenquelle** aus.
3. 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.
4. 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 vorhandenes Paket ist nicht möglich. Weitere Informationen finden Sie im Thema „Auswählen von Cognos-Objektdetails“ auf Seite 43.
5. Geben Sie zur Auswahl des ODBC-Ziels den Datenbanknamen in das Feld **Datenbankname** ein oder wählen Sie die gewünschte Datenbank aus.
6. 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. Weitere Informationen finden Sie im Thema „Abbilden von Scoring-Feldern“.

Abbilden von Scoring-Feldern

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.
1. 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.
 2. Wählen Sie aus den verfügbaren Ausgabezielen und klicken Sie auf **Abbilden**, um sie auf die entsprechenden Scoring-Felder abzubilden.
 3. Wenn Sie fertig sind, klicken Sie auf **Speichern**.

Auswählen von Ausgabefeldern

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. Weitere Informationen finden Sie im Thema „Messniveaus“ auf Seite 46.

Flagziele. Automatisierte Modelle mit Flagzielen 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>**. Propensity-Score; gibt die Wahrscheinlichkeit des Wahr-Werts für jeden Datensatz an. Propensity-Scores unterscheiden sich von Konfidenzscore, 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 Propensity-Score, da dieser eine hohe Wahrscheinlichkeit dafür angibt, dass *keine* Antwort erfolgt.

Kategoriale Ziele. Automatisierte 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. Automatisierte 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 Flagziele 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. Weitere Informationen finden Sie im Thema „Erstellen eines Vorhersageregelmotells (interaktiv)“ auf Seite 61.
- **\$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>**. Propensity-Score; gibt die Wahrscheinlichkeit der gewünschten Antwort an. Nur für Flagziele verfügbar.

Scoring von Modellen

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

Hinweis: Wenn die Daten ungültig oder unvollständig sind, wird beim Versuch, die Vorschau aufzurufen, eine Warnung angezeigt.

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 Symbol zum Aktualisieren. 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.

Anmerkungen:

Wenn die Daten ungültig oder unvollständig sind, wird beim Versuch, sie zu scoren, eine Warnung angezeigt.

Beim Scoring in eine Datenbank wird die Ausführung an die Datenbank zurückgegeben, wann immer dies möglich ist (wird oft als "SQL-Pushback" bezeichnet). Beachten Sie, dass die Ausführung nicht an die Datenbank zurückgegeben wird, wenn das Projekt eine Regel enthält, die eine Referenz auf ein Modell enthält.

Sobald die Daten erfolgreich gescort wurden, erhalten Sie eine Nachricht, in der die Speicherposition der Ergebnisse (in der von Ihnen angegebenen Datenbank oder Datei) angegeben ist. Der Abschnitt "Scoring-Zusammenfassungsbericht" enthält eine Zusammenfassung. Weitere Informationen finden Sie im Thema „Scoring-Zusammenfassungsbericht“. Wenn Sie die Scoring-Ausgabe an eine Flatfile senden möchten und wenn ein Administrator die Funktion aktiviert hat, können Sie bei Bedarf eine lokale Kopie der Datei herunterladen.

Scoring-Zusammenfassungsbericht

Nach dem Scoring des Projekts werden im Abschnitt "Scoring-Zusammenfassungsbericht" die folgenden Details bereitgestellt: die Position, an die die Daten gescort wurden, die Dateigröße (falls verfügbar), der Datenbank- und Tabellename (falls verfügbar), der Modell- oder Projektname, die Anzahl der gescorten Datensätze und in einigen Fällen ein Diagramm und/oder eine Tabelle mit einer Zusammenfassung des

Modells. Der Abschnitt "Scoring-Zusammenfassungsbericht" ist in IBM SPSS Modeler Advantage sowie auf der Registerkarte "Bereitstellen" von Anwendungen verfügbar, die die Funktion "Jetzt scoren" umfassen.

Wenn Sie die Scoring-Ausgabe an eine Flatfile senden möchten, können Sie bei Bedarf auch eine lokale Kopie der Datei herunterladen. Administratoren können die browserbasierte Komponente IBM SPSS Collaboration and Deployment Services Deployment Manager verwenden, um die Funktion zu aktivieren bzw. zu inaktivieren. Ferner können Sie damit die maximal zulässige Dateigröße erhöhen oder verringern. Wenn die Schaltfläche **Datei herunterladen** inaktiv ist, überschreitet die Dateigröße die vom Administrator festgelegte maximale Größe oder der Administrator hat die Funktion inaktiviert. Die Funktion ist standardmäßig inaktiviert.

Kapitel 8. Verwalten von Anwendungen

Administratoren können Anwendungen von IBM Analytical Decision Management für Fachanwender vorkonfigurieren. Diese Konfigurierbarkeit erleichtert Fachanwendern 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.

1. Öffnen Sie IBM SPSS Collaboration and Deployment Services Deployment Manager.
2. Wählen Sie im Menü "Extras" die Option **Serververwaltung** aus.
3. Wählen Sie die Registerkarte "Serververwaltung" aus und stellen Sie die Verbindung zu Ihrem Server her.
4. Öffnen Sie unter **Benutzer und Gruppen** das lokale Benutzerrepository.
5. Klicken Sie auf **Neue Gruppe**, um eine neue Gruppe für Administratoren und eine neue für Fachanwender 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 Anwendungen von IBM Analytical Decision Management in Ihrem Unternehmen berechtigt.

6. Wechseln Sie zu "Rollen" und klicken Sie auf **Neue Rolle**, um eine neue administrative Rolle und eine neue Fachanwenderrolle einzurichten. Zum Beispiel:
 - Administrator von IBM Analytical Decision Management
 - Benutzer von IBM Analytical Decision Management

Weisen Sie den neuen Rollen zumindest die folgenden Aktionen zu.

Tabelle 7. Rollenaktionen.

Rolle	Erforderliche Aktionen
Decision Management-Administrator	Auf Inhalte und Verzeichnisse zugreifen
	Decision Management-Verwaltung
	Jobbearbeitung
	Bericht dynamisch ausführen
	Modell scoren
	Arbeit senden
	Alle Versionen anzeigen
Decision Management-Benutzer	Auf Inhalte und Verzeichnisse zugreifen
	Jobbearbeitung
	Bericht dynamisch ausführen
	Modell scoren
	Arbeit senden
	Neueste anzeigen ODER Alle Versionen anzeigen*

**Hinweis:* Die Aktion *Neueste anzeigen* oder die Aktion *Alle Versionen anzeigen* ist erforderlich, wenn Benutzer auf Projektversionen zugreifen müssen, die sie nicht erstellt haben. Andernfalls werden die Projekte im schreibgeschützten Modus geöffnet, wenn der aktuelle Benutzer keinen Zugriff auf die neueste Version hat.

7. Wählen Sie die Rolle "Decision Management-Administrator" aus und klicken Sie auf **Benutzer und Gruppen bearbeiten**. Fügen Sie die soeben erstellte Gruppe "Decision Management-Administratoren" hinzu.
8. Wählen Sie die Rolle "Decision Management-Benutzer" aus und klicken Sie 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 Benutzerschnittstelle von IBM Analytical Decision Management zur Verfügung. Diese Optionen ermöglichen es Administratoren, Registerkarten für Fachanwender zu sperren oder auszublenden.

- Oben auf jeder Registerkarte befindet sich die Option **Registerkarte <RK-Name> für Nichtadministratoren ausblenden**, wobei <RK-Name> der Name der Registerkarte ist. Dadurch wird die gesamte Registerkarte für Fachanwender 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 Nichtadministratoren ausblenden** aus.
- 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, sodass Fachanwender nur zum Lesen berechtigt sind. Wenn Sie z. B. Fachanwendern 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 Fachanwender voreinstellen wollen.
- Administratoren können auch einzelne Einstellungen einer Registerkarte durch Klicken auf das Schlosssymbol sperren, wo verfügbar.



Bei Anwendungen mit der Registerkarte "Definieren" können die Administratoren beispielsweise den Dimensionshierarchiebaum sperren, sodass dieser nicht von Fachanwendern geändert werden kann. Durch Klicken auf das Schlosssymbol wird das zugehörige Steuerelement ge- bzw. entsperrt. Fachanwender sehen ein verriegeltes Schlosssymbol, wenn der Administrator ein Steuerelement gesperrt hat. Das entriegelte Schlosssymbol 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 Anwendungsverknüpfungen

Der Anwendungsstartseite gibt Fachanwendern 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 auf der Startseite aller Benutzer eingeblendet werden. Melden Sie sich mit Administratorrechten an und gehen Sie wie folgt vor:

- Wenn Sie auf das Schlosssymbol einer Anwendungsverknüpfung klicken, können Fachanwender diese nicht aus ihrer Startseite entfernen. Beachten Sie jedoch, dass Fachanwender immer noch andere Anwendungsverknüpfungen hinzufügen und entfernen können. Der Benutzer sieht bei gesperrten Verknüpfungen anstelle des Schließsymbols ein verriegeltes Schlosssymbol. Weitere Informationen finden Sie im Thema „Starten von Anwendungen“ auf Seite 27.

Ausblenden der Homepage

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

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

1. Wählen Sie auf der Homepage **Anwendungshomepage ausblenden** aus.
2. Legen Sie im Dropdown-Menü fest, welche Registerkarte zuerst eingeblendet werden soll, wenn ein Benutzer die Anwendung öffnet.

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

Sperrungen der Datenoptionen

Daten

Die Registerkarte "Daten" definiert die von der Anwendung für Analysen, Simulationen und Tests sowie für das Scoring und andere operative Einsätze genutzten Datensets. Diese Datenquellen werden relativ zu Ihrer Serverinstallation definiert. Daher sind alle Pfadangaben zu Datendateien, ODBC-Quellen usw. relativ zum Server zu verstehen.

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

- Bei der Erstellung einer neuen Datenquelle **Datenquelle sperren** auswählen, um nach der Erstellung eine Sperrung zu bewirken.
- Einzelne Datenquellen durch Klicken auf das Schlosssymbol neben der Datenquelle in der Tabelle oder im Dialogfenster "Datenquellen-Editor" sperren und entsperren.
- Durch Auswahl eines Datenmodells im Dropdown-Menü **Projektmodell** und Klicken auf das Schlosssymbol das in der Anwendung verwendete Datenmodell einstellen und sperren.

Beachten Sie, dass Fachanwender eine gesperrte Datenquelle aus den Projektdatensets in den Bereich "Eigene Datensets" kopieren können, sie dann aber in beiden Sets gesperrt bleibt. Weitere Informationen finden Sie in Kapitel 4, „Verwalten von Datenquellen“, auf Seite 35.

- Wenn es sich bei dem Projektmodell um eine *Datenansicht* handelt, verwenden Sie die Dropdown-Liste zum Zulassen von auf hierarchischen Daten basierenden ODM-Regeln (ODM - IBM Operational Decision Management), um die Verwendung von ODM-Regeln zu steuern, die auf einem hierarchischen Datenmodell basieren. Die Verwendung solcher Regeln kann zu einer Leistungsverminderung führen. Die folgenden Optionen stehen in der Dropdown-Liste zur Verfügung:
 - **Nicht zulassen** wird für die beste Leistung empfohlen. Lässt nicht zu, dass ODM-Regeln mit einer komplexen Datenstruktur im Projekt verwendet werden. Wenn ein Benutzer eine komplexe Regel in der Anwendung auswählt, wird ein Fehler angezeigt.
 - **Nur auf Registerkarte 'Modellierung'** lässt das Referenzieren von komplexen Regeln beim Erstellen von Modellen auf der Registerkarte "Modellierung" zu. Dadurch wird die Flexibilität erweitert, allerdings wird die Leistung beeinträchtigt. Diese Option gilt nur für IBM SPSS Modeler Advantage.

Sperren 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 Fachanwendern 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 vorhandene Tabelle einer bestimmten Datenbank zulassen, anstatt Fachanwendern die Wahl zwischen Überschreiben vorhandener Daten, Hinzufügung zu einer vorhandenen Tabelle, Erstellung einer neuen Tabelle oder Speicherung in einer Datei zu geben.

- Zum Sperren aller Optionen für das Scoring-Ziel klicken Sie auf das Schlosssymbol neben dem Dropdown-Menü "Ziel" auf der Registerkarte "Bereitstellen". Fachanwender können die Zieloptionen dann einsehen, aber nicht modifizieren.

Herunterladen von Projektmetadaten zur Verwendung beim Erstellen externer Regeln

Regeln, die in einem Business Rules Management System wie IBM Operational Decision Management erstellt wurden, können referenziert und in Anwendungen von IBM Analytical Decision Management verwendet werden, vorausgesetzt, sie wurden so entwickelt, dass sie dasselbe Datenmodell unterstützen, das im aktuellen Projekt von IBM Analytical Decision Management verwendet wird. Diese Regeln können auch als Web-Service für die Anwendung bereitgestellt werden. Dafür muss ein Administrator von IBM Analytical Decision Management die relevanten Metadaten herunterladen. Es gibt zwei Möglichkeiten, um die Metadaten herunterzuladen und die Komplexität des Datenmodells sowie die potenzielle daraus resultierende Regelkomplexität sind dementsprechend verschieden:

Herunterladen von Projektmetadaten aus IBM Analytical Decision Management

Ein Administrator kann auf das Symbol **Metadaten herunterladen** in IBM Analytical Decision Management klicken, um eine ZIP-Datei herunterzuladen, die Metadaten für das aktuelle Projekt enthält. Die heruntergeladene Datei wird lokal auf dem System gespeichert, das vom Administrator zum Auslösen des Downloads verwendet wird.



Abbildung 14. Symbol "Metadaten herunterladen"

Die auf diese Weise heruntergeladenen Metadaten unterstützen die Erstellung von Regeln auf der Basis des Projektdatenmodells. Das Projektdatenmodell ist unstrukturiert (nicht hierarchisch), sodass Regeln, die auf der Basis dieser Metadaten erstellt wurden, nur ein unstrukturiertes Datenmodell verwenden können. Eine Regel, die auf der Basis dieses unstrukturierten Datenmodells erstellt wurde, kann in einem Projekt von IBM Analytical Decision Management unabhängig vom Typ der Datenquelle des Projektdatenmodells referenziert werden.

Die IP-Datei enthält folgende Dateien:

- **XML-Schemadefinition (*.XSD)**. Enthält Definitionen der Feldtypen im Datenmodell für das aktuelle Projekt von IBM Analytical Decision Management. Diese Datei kann in ein externes Entwicklungstool wie IBM Operational Decision Management Rule Designer importiert werden, um Regeln zur Verwendung mit dem aktuellen Projekt zu entwickeln.
- **Web Service Description Language-Vorlage (*.WSDL)**. Unterstützt die Entwicklung von Web-Services, die das Datenmodell verwenden. Die WSDL-Datei wird als Vorlage bereitgestellt, mit der Web-Services

erstellt werden können, die für die Vermittlung von Nachrichten zwischen der Anwendung von IBM Analytical Decision Management und einem externen Regelservice zuständig sind. Details zum vermittelnden Web-Service können manuell vom Integrator eingegeben werden.

Herunterladen von Datenmodellmetadaten für eine Analysedatenansicht von IBM SPSS Collaboration and Deployment Services

In IBM SPSS Collaboration and Deployment Services Deployment Manager können Sie Metadaten für eine bestimmte Analysedatenansicht exportieren.

Die auf diese Weise heruntergeladenen Metadaten unterstützen die Erstellung von Regeln auf der Basis des logischen Datenmodells einer Analysedatenansicht. Das logische Datenmodell einer Analysedatenansicht kann hierarchisch sein. Eine Datenansichtsquelle von IBM Analytical Decision Management unterstützt die Generierung einer XML-Darstellung der hierarchischen Daten für jeden Datensatz, die dann während der Ausführung an externe Regeln übergeben werden können. Regeln, die auf der Basis dieser Metadaten erstellt werden, können komplex sein und das hierarchische Datenmodell verwenden.

In IBM Analytical Decision Management kann eine externe Regel, die auf einem hierarchischen Datenmodell basiert, nur in einem Projekt referenziert werden, in dem das Projektdatenmodell eine Datenansichtsquelle verwendet. Um die Kompatibilität sicherzustellen, muss die Datenansichtsquelle und die externe Regel auf demselben logischen Datenmodell der Analysedatenansicht basieren. Nur die Datenansichtsquelle kann die hierarchischen Daten auf der Basis der Analysedatenansichtsinstanz generieren, die sie referenziert. Anders gesagt muss das Projektdatenmodell im Projekt von IBM Analytical Decision Management dem Datenmodell entsprechen, das zur Erstellung der Regel verwendet wird, und die Analysedatenansicht ermöglicht dies.

So laden Sie eine ZIP-Datei mit Metadaten für das logische Datenmodell der Analysedatenansicht herunter:

1. Öffnen Sie in IBM SPSS Collaboration and Deployment Services Deployment Manager eine Analysedatenansicht aus dem Repository und wechseln Sie zur Registerkarte **Logisch**.
2. Klicken Sie mit der rechten Maustaste auf eine Tabelle und exportieren Sie die XOM-Metadaten.
3. Wählen Sie einen Speicherort für die Datei aus und klicken Sie auf **Speichern**.

Die ZIP-Datei enthält folgende Dateien:

XML-Schemadefinitionen (*.XSD). Es sind mindestens zwei XML-Schemadefinitionen enthalten. Eine enthält eine Definition, die für einen Regelantworttyp hilfreich sein kann. Die anderen Schemas enthalten die Definitionen der Tabellen, einschließlich Details zur hierarchischen Beziehung und Feldtypen im logischen Datenmodell der Analysedatenansicht.

Sie können diese Schemadateien in ein externes Entwicklungstool wie IBM Operational Decision Management Rule Designer importieren, um Regeln zur Verwendung mit einem Projekt zu entwickeln, dessen Projektdatenmodell eine Datenansichtsquelle auf der Basis derselben Analysedatenansichtsinstanz verwendet.

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

Beachten Sie, dass Administratoren die Verwendung dieser komplexen externen Regeln einschränken können, da sie zu einer Leistungsver schlechterung führen. Details finden Sie im Thema „Sperren der Datenoptionen“ auf Seite 87.

Bearbeiten der XML-Schemadefinition

Die von IBM Analytical Decision Management heruntergeladene XSD-Datei enthält zwei komplexe Datentypen: einen für das Projektdatenmodell von IBM Analytical Decision Management, das für den Eingab-

beparameter für Regelprojekte von IBM Operational Decision Management verwendet werden soll, und einen für einen Regelrückgabety, der für den Ausgabeparameter für Regelprojekte von IBM Operational Decision Management verwendet werden soll.

Die Metadaten der Analysedatenansicht von IBM SPSS Collaboration and Deployment Services Deployment Manager enthalten mindestens zwei *XSD*-Dateien: eine, die sich auf das logische Datenmodell einer Analysedatenansicht bezieht und die für den Eingabeparameter für Regelprojekte von IBM Operational Decision Management verwendet werden soll, und mindestens eine, die einen komplexen Typ enthält, der einen Regelrückgabety darstellt, und die für den Ausgabeparameter für Regelprojekte von IBM Operational Decision Management verwendet werden soll.

Erfahrene Benutzer können die erstellte *XSD*-Schemadefinition bearbeiten und für die entsprechende(n) Regel(n), die sie erstellen möchten, anpassen. Unter bestimmten Umständen kann es vorkommen, dass der Benutzer vor dem Erstellen seiner Regeln die in der zugehörigen *XSD*-Schemadefinition definierten Eingabe- und/oder Antworttypen bearbeiten möchte. Der erfahrene Benutzer muss beispielsweise möglicherweise das auf den Eingabeparameter bezogene Schema bearbeiten, damit seine erstellten Regeln gegebenenfalls Eingabefelder mit Nullwerten verarbeiten können. Der erfahrene Benutzer kann diesen Vorgang ausführen, indem er über die *XSD*-Schemadefinition angibt, welche Eingabeparameterfelder/-elemente erforderlich sind und welche optional sind.

Bearbeiten des Eingabeparametertyps

Die Anzahl der möglichen Vorkommen der einzelnen feldbezogenen Elemente ist im *XSD*-Schema nicht explizit festgelegt (d. h., die Attribute *minOccurs* und *maxOccurs* sind nicht festgelegt und der Standardwert für beide Attribute ist 1). Somit sind alle Felder im Eingabeparametertyp für die erstellten Regeln erforderlich, wenn das Schema nicht bearbeitet wird, bevor es zum Generieren von Regeln für IBM Operational Decision Management verwendet wird.

Bearbeiten des Eingabeparametertyps für das Projektdatenmodell

Wenn das Schema, das den Eingabeparametertyp definiert, durch Herunterladen der Metadaten eines Projekts von IBM Analytical Decision Management generiert wurde, ist das Datenmodell des Eingabeparameters unstrukturiert. In diesem Fall erstellt IBM Analytical Decision Management die vom Regelservice für die Ausführung benötigten Nutzdaten direkt. Dadurch erhält der erfahrene Benutzer beim Bearbeiten des Eingabeparametertyps mehr Freiheit und kann trotzdem die Kompatibilität aufrechterhalten (d. h., der erfahrene Benutzer kann Felder/Elemente entfernen, die er für das Regelauthoring nicht benötigt, oder er kann angeben, dass Felder/Elemente optional sind).

Bearbeiten des Eingabeparametertyps für das logische Datenmodell der Analysedatenansicht

Wenn das Schema, das den Eingabeparametertyp definiert, durch Exportieren des logischen Datenmodells einer Analysedatenansicht aus dem Client von IBM SPSS Collaboration and Deployment Services Deployment Manager generiert wurde, kann das Datenmodell des Eingabeparameters hierarchisch sein. Wenn es hierarchisch ist, erstellt IBM Analytical Decision Management die vom Regelservice für die Ausführung benötigten Nutzdaten *nicht* direkt. Der Hauptteil der Nutzdaten wird durch die Analysedatenansicht generiert. Dadurch wird eingeschränkt, wie der erfahrene Benutzer den Eingabeparametertyp bearbeiten kann und die Kompatibilität aufrechterhalten kann (der erfahrene Benutzer kann z. B. nur angeben, dass Felder/Elemente optional sind).

Allgemeine Informationen

Vor dem Aufrufen einer Regel, die in einem Business Rules Management System wie IBM Operational Decision Management erstellt wurde, fragt IBM Analytical Decision Management den Regelservice, ob sein Eingabeparameterdatenmodell hierarchisch oder unstrukturiert ist. Wenn der Eingabeparameter unstrukturiert ist, wird auch überprüft, welche Eingaben vom Regelservice erforderlich sind.

Wenn das Datenmodell des Eingabeparameters unstrukturiert ist und der Datensatz, der gescort werden soll, einen Nullwert für eines der Felder enthält, die für den Regelservice erforderlich sind, ruft IBM Analytical Decision Management die Regel nicht auf und die Ausgabe-/Antwortfelder der Regel werden mit Standardwerten ausgefüllt. Der Wert hängt folgendermaßen vom Typ des Ausgabe-/Antwortfelds ab:

Tabelle 8. Werte für Antwortfeldtypen.

Antwortfeldtyp	Wert
xsd:string	""
xsd:boolean	""
xsd:int	0
xsd:long	0
xsd:byte	0
xsd:double	0.0
xsd:float	0.0
xsd:date	0.0
xsd:time	0.0
xsd:datetime	0.0

Dies kann sich auf die Ergebnisse beim Scoring von Datensätzen über einen Regelservice auswirken. Beim Bearbeiten von Eingabeparametern im XSD-Schema sollten Sie daher vorsichtig vorgehen.

Wenn das Datenmodell des Eingabeparameters hierarchisch ist und der Datensatz, der gescort werden soll, einen Nullwert für eines der Felder enthält, die für den Regelservice erforderlich sind, ruft IBM Analytical Decision Management die Regel auf, das Feld oder die Felder mit einem Nullwert wird bzw. werden jedoch in der vom Datenansichtsquellenknoten erstellten Nutzlast übergangen. Wenn das übergangene Feld ein erforderliches Feld war, führt dies dazu, dass der Regelaufruf fehlschlägt. Daher sollten Sie beim Bearbeiten des Eingabeparameters im XSD-Schema wieder vorsichtig vorgehen.

Für das Bearbeiten des Schemas sollten Sie sich mit den Regelservices auskennen, die erstellt werden. Wenn ein bestimmtes Eingabefeld nicht für die Regel erforderlich ist und das Eingabeparameterdatenmodell der Regel **unstrukturiert** ist, kann das feldbezogene Element entweder aus dem Eingabeparameter im Schema **entfernt** oder **als optional markiert** werden. Wenn ein bestimmtes Feld nicht für die Regel erforderlich ist und der Eingabeparameter der Regel **hierarchisch** ist, kann das feldbezogene Element **nicht entfernt** werden, es kann jedoch **als optional markiert** werden. Dies ist darauf zurückzuführen, dass IBM Analytical Decision Management die Nutzdaten für Regeln nicht direkt auf der Basis eines hierarchische Datenmodells generiert.

Wenn Sie ein Eingabeparameterfeld entfernen, wird der Feldwert beim Aufrufen der Regel ignoriert und die Regel kann keine Beziehung zum Feld herstellen. Wenn Sie ein Eingabeparameterfeld als optional festlegen, ist ein Nullwert zulässig. Um dies zu prüfen, muss die Regel zudem defensiv geschrieben werden. Wenn das Eingabeparameterdatenmodell der Regel beispielsweise unstrukturiert ist und Sie wissen, dass das Feld Age nicht von der Regel verwendet wird, können Sie das zugehörige Element entfernen, um anzugeben, dass dieses Feld nicht erforderlich ist. Im folgenden Beispiel eines XSD-Schemas ist dargestellt, dass der gesamte Abschnitt `<xsd:element name="Age">` entfernt werden kann, wenn das Feld Age nicht erforderlich ist:

```
<xsd:complexType name="drugs">
  <xsd:annotation>
    <xsd:appinfo>
      <dmname>drugs</dmname>
    </xsd:appinfo>
  </xsd:annotation>
<xsd:sequence>
  <xsd:element name="Age">
```

```

<xsd:annotation>
<xsd:appinfo>
  <dmname>Age</dmname>
</xsd:appinfo>
</xsd:annotation>
<xsd:simpleType>
  <xsd:restriction base="xsd:long">
    <xsd:minInclusive value="15"/>
    <xsd:maxInclusive value="74"/>
  </xsd:restriction>
</xsd:simpleType>
</xsd:element>
<xsd:element name="Sex">
  <xsd:annotation>
  <xsd:appinfo>
    <dmname>Sex</dmname>
  </xsd:appinfo>
  </xsd:annotation>
  <xsd:simpleType>
    <xsd:restriction base="xsd:string">
      <xsd:enumeration value="F"/>
      <xsd:enumeration value="M"/>
    </xsd:restriction>
  </xsd:simpleType>
</xsd:element>
...
</xsd:sequence>
</xsd:complexType>

```

Wenn das Eingabeparameterdatenmodell der Regel unstrukturiert oder hierarchisch ist, könnten Sie auch ein Eingabefeldelement als optional festlegen, wenn es für die Regel(n) nicht erforderlich ist, indem Sie den Wert für das Attribut `minOccurs` des Elements auf 0 setzen:

```

<xsd:complexType name="drugs">
  <xsd:annotation>
  <xsd:appinfo>
    <dmname>drugs</dmname>
  </xsd:appinfo>
  </xsd:annotation>
<xsd:sequence>
  <xsd:element name="Age" minOccurs="0">
    <xsd:annotation>
    <xsd:appinfo>
      <dmname>Age</dmname>
    </xsd:appinfo>
    </xsd:annotation>
    <xsd:simpleType>
      <xsd:restriction base="xsd:long">
        <xsd:minInclusive value="15"/>
        <xsd:maxInclusive value="74"/>
      </xsd:restriction>
    </xsd:simpleType>
  </xsd:element>
  <xsd:element name="Sex">
    <xsd:annotation>
    <xsd:appinfo>
      <dmname>Sex</dmname>
    </xsd:appinfo>
    </xsd:annotation>
    <xsd:simpleType>
      <xsd:restriction base="xsd:string">
        <xsd:enumeration value="F"/>
        <xsd:enumeration value="M"/>
      </xsd:restriction>
    </xsd:simpleType>
  </xsd:element>
  ...
</xsd:sequence>
</xsd:complexType>

```

Beispiel für ein hierarchisches Datenmodell:

```

<xsd:complexType name="Customer">
  <xsd:sequence>
    <xsd:element name="FirstName" type="xsd:string" minOccurs="0"/>
    <xsd:element name="Surname" type="xsd:string" minOccurs="0"/>
    <xsd:element name="Age" type="xsd:long"/>
    <xsd:element name="Sex" type="xsd:string"/>
    <xsd:element name="CustID" type="xsd:string"/>
    <xsd:element minOccurs="unbounded" minOccurs="0" name="Orders" type="xom:Order"/>
  </xsd:sequence>
</xsd:complexType>
<xsd:complexType name="Order">
  <xsd:sequence>
    <xsd:element name="CustID" type="xsd:string"/>
    <xsd:element name="NoOfItems" type="xsd:long"/>
    <xsd:element name="TotalCost" type="xsd:long" minOccurs="0"/>
    <xsd:element name="Date" type="xsd:date"/>
  </xsd:sequence>
</xsd:complexType>

```

Die Regeln können auch defensiv geschrieben werden, um Nullwerte zu bearbeiten. In diesem Fall kann das Feld im Schema als optional markiert werden. Im folgenden Beispiel ist die Regel *currentOfferIsReplacePhone* von IBM Operational Decision Management dargestellt, die zum Bearbeiten mit Nullwerten geschrieben wurde:

```

if
  the CURRENT OFFER of 'the customer' is not null
  and the CURRENT OFFER of 'the customer' is "Replace Phone"
then
  make it true that 'the response' is value ;
else
  make it false that 'the response' is value ;

```

Bearbeiten des Antwortparametertyps

Hinweis: Es ist möglich, Regeln für IBM Operational Decision Management zu erstellen und zu verwenden, die verschiedene Typen an IBM Analytical Decision Management zurückgeben (*boolean* (boolesch) zur Verwendung bei Auswahlen, *numeric* (numerisch) beispielsweise zur Verwendung in Aggregationsregeln und *string* (Zeichenfolge)). Um jedoch einen anderen Typ als *string* bei der Datenantwortregel zu erstellen, muss die XML-Schemadefinition (.XSD), die den Antworttyp enthält, 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 Datenantwortvariable auf "numeric" gesetzt wurde, eine weitere für "string", eine für "boolean" usw.).

Die von IBM Analytical Decision Management heruntergeladene XSD-Datei enthält zwei komplexe Datentypen: einen für das Projektdatenmodell von IBM Analytical Decision Management, das für den Eingabeparameter für Projekte von IBM Operational Decision Management verwendet werden soll, und einen für einen Regelrückgabetyt, der für den Ausgabeparameter von Regelprojekten von IBM Operational Decision Management verwendet werden soll.

Die antwortbezogene XSD-Datei, die von IBM SPSS Collaboration and Deployment Services Deployment Manager heruntergeladen wurde, enthält die Antworttypdefinition und enthält nur den einen komplexen Datentyp.

In beiden oben stehenden Fällen definiert der Rückgabetyt (dessen Name immer ein Suffix "response" enthält) ein value-Element, das standardmäßig den Typ *string* aufweist. Wenn Sie ein Regelprojekt mit einem anderen Rückgabewert erstellen möchten, müssen Sie die heruntergeladene XSD-Datei vor dem Importieren des Regelprojekts von IBM Operational Decision Management manuell bearbeiten, damit das Element value des Rückgabetyps 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>

```

```

<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>

```

Kapitel 9. Aufrufen von Hilfe

Es gibt drei Stufen der Benutzerhilfe:

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



- Infotext, der 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 angezeigt werden, indem Sie auf die kleinen Infotextsymbole der Benutzerschnittstelle klicken:



- Standard-QuickInfos, 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 die zugehörige QuickInfo anzuzeigen.

Verfügbare Dokumentation

Die Dokumentation zu IBM Analytical Decision Management ist in zwei Formaten verfügbar. Wenn Sie in der Anwendung auf die Hilfe zugreifen, wird die Onlineversion des IBM Knowledge Center gestartet und das relevante Hilfethema wird angezeigt. Wenn Sie mit dem Starten der Hilfe Probleme haben, wenden Sie sich an Ihren Administrator, falls Sie keine Zugriffsberechtigung für die URL des externen Knowledge Center haben.

Das Knowledge Center ist eine zentrale Onlinesammlung der gesamten Dokumentation. Es ist immer verfügbar, auch wenn das Produkt nicht aktiv ist. Es wird empfohlen, in Ihrem Browser ein entsprechendes Lesezeichen zu setzen. Es enthält außerdem Dokumentation für andere IBM Produkte wie z. B. IBM SPSS Modeler, IBM SPSS Collaboration and Deployment Services und Statistics.

Für den Fall, dass Sie keine Internetverbindung haben oder das Knowledge Center aufgrund von Wartungsarbeiten vorübergehend nicht verfügbar ist, steht die gesamte Dokumentation auch in Portable Document Format (*.pdf) zur Verfügung. Die PDF-Dateien sind in den heruntergeladenen Installationsdateien verfügbar und werden außerdem im Installationsverzeichnis von IBM SPSS Collaboration and Deployment Services installiert (z. B. *C:\Programme\IBM\SPSS\Deployment\8.0\Server\documentation\DecisionManagement*).

Die Dokumentation ist außerdem im Internet unter der Adresse <http://www-01.ibm.com/support/docview.wss?uid=swg27047202> im PDF-Format verfügbar. Releaseinformationen finden Sie unter <http://www-01.ibm.com/support/docview.wss?uid=swg27047203>.

Anhang A. Gemeinsame Nutzung von Streams

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

Modelle, die dem in diesem Anhang beschriebenen Format entsprechen, können mithilfe von IBM SPSS Modeler Advantage angezeigt, bearbeitet, gescort und aktualisiert werden. Andere Modelle können ebenfalls in IBM SPSS Modeler Advantage gescort werden. Sie werden jedoch möglicherweise nicht richtig angezeigt und können nicht aktualisiert werden.

Öffnen von Streams in IBM Analytical Decision Management

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

- Wenn 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 oder IBM Analytical Decision Management unterstützt wird. Anders ausgedrückt werden alle von IBM SPSS Modeler unterstützten Quellentypen auch von IBM SPSS Modeler Advantage und IBM Analytical Decision Management unterstützt, solange nur ein einziger Quellenknoten im Stream vorhanden ist (keine zusammengeführten Datasets).
- 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 oder IBM Analytical Decision Management unterstützt wird. Zudem muss der ausgewählte Modellierungsknoten entweder ein Knoten für die automatisierte Modellierung (unterstützt für Vorhersagemodelle), ein Entscheidungslistenknoten (unterstützt für Vorhersageregelmmodellierung), ein Clustering-Knoten (unterstützt für Clustermodellierung) oder ein Assoziationsregelknoten (unterstützt für Assoziationsmodellierung) sein. Außerdem muss ein Typknoten vorhanden sein, der das Zielfeld sowie andere bei der Modellierung verwendete Informationen definiert.

Angeben von Bereitstellungseinstellungen in IBM SPSS Modeler

In IBM SPSS Modeler werden Bereitstellungseinstellungen im Dialogfeld "Streameigenschaften" wie folgt angegeben:

1. Wählen Sie in den IBM SPSS Modeler-Menüs Folgendes aus:
Extras > Streameigenschaften > Bereitstellung
2. Wählen Sie in der Liste **Bereitstellungstyp** die Option **Nur Scoring** bzw. **Modellaktualisierung** aus.
3. Wählen Sie nach Bedarf einen **Scoring-Knoten**, **Modellierungsknoten** und **Modellknoten** aus.
4. Wählen Sie **Überprüfen** aus, 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 IBM SPSS Modeler zu bearbeiten, als sie ohne Vorlage mit IBM 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 lokalen Regelknoten, die eine Ausdrucksregel unmittelbar unterhalb jedes Quellenknotens enthalten, werden die zugehörigen Felder in das Dialogfeld "Zusatzfelder hinzufügen/bearbeiten" aufgenommen. Beachten Sie, dass lokale Regeln nur in IBM Analytical Decision Management oder IBM SPSS Modeler Advantage erstellt werden können und nicht von IBM SPSS Modeler.
- Wenn es oberhalb des Modellierungsknotens (und unterhalb des Typknotens) einen Partitionsknoten gibt, ist **Daten automatisch partitionieren, um die Modellauswertung von Entwicklungsdatenquellen zu ermöglichen (für weitere Evaluierung und Tests)** in IBM SPSS Modeler Advantage aktiviert und der Status des Kontrollkästchens wird danach festgelegt, ob der Partitionsknoten aktiviert ist oder nicht.
- Wenn sich unterhalb des Typknotens ein ADP-Knoten (Automatisierte Datenvorbereitung) befindet, ist **Daten automatisch bereinigen und zur zuverlässigen Modellerstellung aufbereiten** aktiviert und der Status 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 Auswahl festlegen** mit Daten zu füllen.
- Wenn Streamparameter definiert wurden (Registerkarte "Parameter" im Dialogfeld "Streameigenschaften") 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 über IBM SPSS Modeler

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

Anhang B. Behindertengerechte Bedienung

Funktionen zu behindertengerechter Bedienung 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 Navigationsstasten von Web-Browsern verwendet.

Da IBM Analytical Decision Management eine browserbasierte Anwendung ist, werden Funktionen zur behindertengerechten Bedienung für Menschen mit Behinderung, z. B. Sehbehinderte, über die Web-Browser-Einstellungen konfiguriert. Weitere Informationen zu Benutzerschnittstellenaktionen wie z. B. Direktaufrufen über die Tastatur, Sprachausgabeprogramm o. Ä. finden Sie in der Dokumentation Ihres Web-Browsers.

Software anderer Anbieter

IBM Analytical Decision Management kann die Verwendung von Software anderer Anbieter erforderlich machen, die nicht von der IBM Lizenzvereinbarung abgedeckt wird. IBM gibt keine Gewährleistung ab hinsichtlich der behindertengerechten Bedienung dieser Produkte. Wenden Sie sich an den Hersteller, um Informationen zur behindertengerechten Bedienung seiner Produkte zu erhalten.

IBM und behindertengerechte Bedienung

Im IBM Human Ability and Accessibility Center finden Sie weitere Informationen zum Engagement von IBM im Bereich der behindertengerechten Bedienung.

Bemerkungen

Die vorliegenden Informationen wurden für Produkte und Services entwickelt, die auf dem deutschen Markt angeboten werden. IBM stellt dieses Material möglicherweise auch in anderen Sprachen zur Verfügung. Für den Zugriff auf das Material in einer anderen Sprache kann eine Kopie des Produkts oder der Produktversion in der jeweiligen Sprache erforderlich sein.

Möglicherweise bietet IBM die in dieser Dokumentation beschriebenen Produkte, Services oder Funktionen in anderen Ländern nicht an. Informationen über die gegenwärtig im jeweiligen Land verfügbaren Produkte und Services sind beim zuständigen IBM Ansprechpartner erhältlich. Hinweise auf IBM Lizenzprogramme oder andere IBM Produkte bedeuten nicht, dass nur Programme, Produkte oder Services von IBM verwendet werden können. Anstelle der IBM Produkte, Programme oder Services können auch andere, ihnen äquivalente Produkte, Programme oder Services verwendet werden, solange diese keine gewerblichen oder anderen Schutzrechte von IBM verletzen. Die Verantwortung für den Betrieb von Produkten, Programmen und Services anderer Anbieter liegt beim Kunden.

Für in diesem Handbuch beschriebene Erzeugnisse und Verfahren kann es IBM Patente oder Patentanmeldungen geben. Mit der Auslieferung dieses Handbuchs ist keine Lizenzierung dieser Patente verbunden. Lizenzanforderungen sind schriftlich an folgende Adresse zu richten (Anfragen an diese Adresse müssen auf Englisch formuliert werden):

*IBM Director of Licensing
IBM Europe, Middle East & Africa
Tour Descartes
2, avenue Gambetta
92066 Paris La Defense
France*

Diese Informationen können technische Ungenauigkeiten oder typografische Fehler enthalten. Die hier enthaltenen Informationen werden in regelmäßigen Zeitabständen aktualisiert und als Neuausgabe veröffentlicht. IBM kann ohne weitere Mitteilung jederzeit Verbesserungen und/oder Änderungen an den in dieser Veröffentlichung beschriebenen Produkten und/oder Programmen vornehmen.

Verweise in diesen Informationen auf Websites anderer Anbieter werden lediglich als Service für den Kunden bereitgestellt und stellen keinerlei Billigung des Inhalts dieser Websites dar. Das über diese Websites verfügbare Material ist nicht Bestandteil des Materials für dieses IBM Produkt. Die Verwendung dieser Websites geschieht auf eigene Verantwortung.

Werden an IBM Informationen eingesandt, können diese beliebig verwendet werden, ohne dass eine Verpflichtung gegenüber dem Einsender entsteht.

Lizenznehmer des Programms, die Informationen zu diesem Produkt wünschen mit der Zielsetzung: (i) den Austausch von Informationen zwischen unabhängig voneinander erstellten Programmen und anderen Programmen (einschließlich des vorliegenden Programms) sowie (ii) die gemeinsame Nutzung der ausgetauschten Informationen zu ermöglichen, wenden sich an folgende Adresse:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
USA*

Die Bereitstellung dieser Informationen kann unter Umständen von bestimmten Bedingungen - in einigen Fällen auch von der Zahlung einer Gebühr - abhängig sein.

Die Lieferung des in diesem Dokument beschriebenen Lizenzprogramms sowie des zugehörigen Lizenzmaterials erfolgt auf der Basis der IBM Rahmenvereinbarung bzw. der Allgemeinen Geschäftsbedingungen von IBM, der IBM Internationalen Nutzungsbedingungen für Programmpakete oder einer äquivalenten Vereinbarung.

Die angeführten Leistungsdaten und Kundenbeispiele dienen nur zur Illustration. Die tatsächlichen Ergebnisse beim Leistungsverhalten sind abhängig von der jeweiligen Konfiguration und den Betriebsbedingungen.

Alle Informationen zu Produkten anderer Anbieter stammen von den Anbietern der aufgeführten Produkte, deren veröffentlichten Ankündigungen oder anderen allgemein verfügbaren Quellen. IBM hat diese Produkte nicht getestet und kann daher keine Aussagen zu Leistung, Kompatibilität oder anderen Merkmalen machen. Fragen zu den Leistungsmerkmalen von Produkten anderer Anbieter sind an den jeweiligen Anbieter zu richten.

Aussagen über Pläne und Absichten von IBM unterliegen Änderungen oder können zurückgenommen werden und repräsentieren nur die Ziele von IBM.

Diese Veröffentlichung enthält Beispiele für Daten und Berichte des alltäglichen Geschäftsablaufs. Sie sollen nur die Funktionen des Lizenzprogramms illustrieren und können Namen von Personen, Firmen, Marken oder Produkten enthalten. Alle diese Namen sind frei erfunden; Ähnlichkeiten mit tatsächlichen Namen und Adressen sind rein zufällig.

Marken

IBM, das IBM Logo und ibm.com sind Marken oder eingetragene Marken der IBM Corporation in den USA und/oder anderen Ländern. Weitere Produkt- und Servicennamen können Marken von IBM oder anderen Unternehmen sein. Eine aktuelle Liste der IBM Marken finden Sie auf der Webseite "Copyright and trademark information" unter www.ibm.com/legal/copytrade.shtml.

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

Intel, das Intel-Logo, Intel Inside, das Intel Inside-Logo, Intel Centrino, das Intel Centrino-Logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium und Pentium sind Marken oder eingetragene Marken der Intel Corporation oder ihrer Tochtergesellschaften in den USA oder anderen Ländern.

Linux ist eine eingetragene Marke von Linus Torvalds in den USA und/oder anderen Ländern.

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

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

Java und alle auf Java basierenden Marken und Logos sind Marken oder eingetragene Marken der Oracle Corporation und/oder ihrer verbundenen Unternehmen.

Bedingungen für Produktdokumentation

Die Berechtigungen zur Nutzung dieser Veröffentlichungen werden Ihnen auf der Basis der folgenden Bedingungen gewährt.

Anwendbarkeit

Diese Bedingungen sind eine Ergänzung der Nutzungsbedingungen auf der IBM Website.

Persönliche Nutzung

Sie dürfen diese Veröffentlichungen für Ihre persönliche, nicht kommerzielle Nutzung unter der Voraussetzung vervielfältigen, dass alle Eigentumsvermerke erhalten bleiben. Sie dürfen diese Veröffentlichungen oder Teile der Veröffentlichungen ohne ausdrückliche Genehmigung von IBM weder weitergeben oder anzeigen noch abgeleitete Werke davon erstellen.

Kommerzielle Nutzung

Sie dürfen diese Veröffentlichungen nur innerhalb Ihres Unternehmens und unter der Voraussetzung, dass alle Eigentumsvermerke erhalten bleiben, vervielfältigen, weitergeben und anzeigen. Sie dürfen diese Veröffentlichungen oder Teile der Veröffentlichungen ohne ausdrückliche Genehmigung von IBM außerhalb Ihres Unternehmens weder vervielfältigen, weitergeben oder anzeigen noch abgeleitete Werke davon erstellen.

Berechtigungen

Abgesehen von den hier gewährten Berechtigungen werden keine weiteren Berechtigungen, Lizenzen oder Rechte (veröffentlicht oder stillschweigend) in Bezug auf die Veröffentlichungen oder darin enthaltene Informationen, Daten, Software oder geistiges Eigentum gewährt.

IBM behält sich das Recht vor, die hierin gewährten Berechtigungen nach eigenem Ermessen zurückzuziehen, wenn sich die Nutzung der Veröffentlichungen für IBM als nachteilig erweist oder wenn die obigen Nutzungsbestimmungen nicht genau befolgt werden.

Sie dürfen diese Informationen nur in Übereinstimmung mit allen anwendbaren Gesetzen und Vorschriften, einschließlich aller US-amerikanischen Exportgesetze und Verordnungen, herunterladen und exportieren.

IBM übernimmt keine Gewährleistung für den Inhalt dieser Veröffentlichungen. Diese Veröffentlichungen werden auf der Grundlage des gegenwärtigen Zustands (auf "as-is"-Basis) und ohne eine ausdrückliche oder stillschweigende Gewährleistung für die Handelsüblichkeit, die Verwendungsfähigkeit für einen bestimmten Zweck oder die Freiheit von Rechten Dritter zur Verfügung gestellt.

Index

A

- Abbilden von Feldern 47
- Abgeleitet (Registerkarte) 48
- Abgeleitete Felder hinzufügen/bearbeiten 48
- Ableiten von Feldern 49
- Allgemeine Optionen 86
- Analytic Server-Daten 44
- Analytic Server-Ziel in Modellen 79
- Ändern der Browsersprache 27
- Ändern der Sprache 27
- Anmerkungen
 - zu Regeln hinzufügen 57
- Anwendungen
 - starten 13
 - Test 72
- Anwendungs-URL 27
- Anwendungshomepage 28
- Anwendungsstartseite 27, 86
 - Anpassung 27
 - Anwendungen hinzufügen 27
- Anwendungsverknüpfungen 27, 86
- Artefakte für Beispielanwendungen 11
- Assoziationsmodelle 64, 68
 - Tabellendaten 65
 - Tabellendaten im Vergleich zu Transaktionsdaten 65
 - Transaktionsdaten 65
- Ausblenden der Homepage 87
- Ausdrücke, die einen booleschen Ausdruck zurückgeben 52
- Ausdrücke hinzufügen 48
- Ausdrücke überprüfen 51
- Ausdruckseditor 49
- Ausgabefelder in Modellen 81
- Auswahlregeln 55
- Auswerten von Modellen 17
 - Clustermodelle 22, 23
 - interaktive Listenmodelle 20
 - Vorhersageregelmolelle 20
- Auswertung von Modellen 69

B

- Bearbeiten von abgeleiteten Feldern 48
- Bearbeiten von Zusatzfeldern 48
- Behindertengerechte Bedienung 99
- Beispiel für Assoziationsmodellierung 23
- Beispielanwendungen 12
- Beispieldateien 11
- Beispieldaten 12
- Benutzervoreinstellungen 29
- Beschreibung 33
- Beschriften von Versionen 32
- Boolescher Ausdruck 52

C

- Clustering-Modelle 63, 67
- Clustermodellierung, Beispiel 22
- Cognos BI-Quellen 42
- Cognos BI-Server als Ziel in Modellen 79

D

- Dateien für Beispielanwendungen 11, 12
- Dateien speichern 31, 32
- Dateiziel in Modellen 78
- Daten für Beispielanwendungen 12
- Datenansicht 41
- Datenauswahl in Modellen 77
- Datenbankquellen 41
- Datenbankziel in Modellen 77
- Datendateien 12
- Datenoptionen 87
- Datenquellen
 - Abgeleitet (Registerkarte) 48
 - Ableiten von Feldern 48
 - Analytic Server 44
 - Ausdrücke hinzufügen 48
 - Dateiquellen 38
 - Datenansicht 41
 - Datenbank 41
 - Datenübersicht 45
 - Definition 35, 37
 - Eingabefelder 45
 - Excel 38
 - Expression Builder 49, 52
 - Felder abbilden 47
 - Felder hinzufügen 48
 - Felder mithilfe von Segmentregeln erstellen 48
 - Feldwerte bearbeiten 47
 - Hadoop 44
 - HDFS 44
 - Hinzufügen von Modellausgaben 48
 - IBM Cognos BI 42
 - IBM SPSS Statistics 38
 - Messniveaus 46
 - textbasiert 38
 - Vorschau 45
- Datensatzauswahl in Modellen 77
- Datentypen 46
- Datenübersicht 45
- Datumsformate 29
- Demoartefakte 11
- Diagramm für die Bedeutsamkeit der Variablen 60
- Diagramme in Modellen 60
- Direktwerbung, Beispiel 11

E

- Eigene Datenquellen 35
- Eigenschaften
 - Beschreibung 33

Eigenschaften (*Forts.*)

- maximale Anzahl Mitglieder für nominale Felder 33
- Projekt 33
- Repository 33
- Schlüsselwörter 33
- Eingabefelder
 - abbilden 47
 - Definition 45
- Erstellen von Modellen
 - Zielfeld 15
- Excel-Datendateien 38
- Exportieren von Regeln 56
- Expression Builder 48, 49, 52
- Externe Regeln
 - in Anwendungen verwenden 58

F

- Fehlerklassifizierungsdiagramme 69
- Fehlerverteilungsdiagramme 69
- Felder
 - abbilden 47
 - ableiten 49
 - in Ausdrücken 51
- Felder abbilden 80
- Felder hinzufügen 48
- Felder mithilfe von Segmentregeln erstellen 48
- Feldwerte bearbeiten 47
- Funktionen
 - in Ausdrücken 50

G

- Galerie
 - lang laufende Aufgaben 28
 - nicht fertiggestellte Aufgaben 28
- Gemeinsam genutzte Regeln 56
- Geschäftsregeln 53
- Gewinndiagramme 17, 69
- Globale Werte
 - in Ausdrücken 51

H

- Hadoop-Daten 44
- HDFS-Daten 44
- Hilfe 95
- Hilfe aufrufen 95
 - verfügbare Dokumentation 95
- Hinzufügen von Feldern mit Segmentregeln 48
- Hinzufügen von Modellausgaben 48
- Hinzufügen von Zusatzfeldern 48
- Hochladen von Streamdateien 31
- Hochladen von Streams 31
- Homepage 28, 87

I

IBM Cognos BI-Quellen 42
IBM Cognos BI-Quellenknoten
Symbole 44
IBM Operational Decision Management,
Regeln
in Anwendungen verwenden 58
IBM SPSS Collaboration and Deployment
Services Repository
speichern in 31, 32
IBM SPSS Modeler Advantage
Beispiel 11
IBM SPSS Statistics-Datendateien 38
Infotext 95
Interaktive Listenmodelle 61, 67

L

Liftdiagramme 69

M

Manueller Cluster 64
Maximale Anzahl Mitglieder für nomina-
le Felder 33
Messniveaus 46
Modell testen 72
Modelle
Analytic Server-Ziel auswählen 79
Assoziation 64
Ausgabefeld auswählen 81
auswerten 17
Clustering 63
manueller Cluster 64
Cognos BI-Server als Ziel auswäh-
len 79
Dateiziel auswählen 78
Datenauswahl 77
Datenbankziel auswählen 77
Datensatzauswahl 77
Diagramme 60
Ergebnisse 60
erstellen 15, 59
Evaluierung 69
in Regeln verwenden 53
interaktive Liste 61
optionale Einstellungen 65, 67, 68
Profit simulieren 71
Scoring 21, 75, 82
Scoring-Felder abbilden 80
Scoring-Ziel auswählen 77
Scoring-Zusammenfassungsbe-
richt 82
Test 21, 72
Voreinstellungen anpassen 29
Vorhersage 59
Vorhersageregeln 61
Zeit für die Erstellung beschrän-
ken 29
Zielfeld 15, 59

N

Neue Funktionen 1, 2, 3
Neuerungen 1, 2, 3

Nicht fertiggestellte Objekte
Galerie 28

O

Öffnen von Dateien 31
Onlinehilfe 95
Operational Decision Management
Regeln 58
Operatoren
in Ausdrücken 50
Optionen für das Scoring-Ziel 88

P

Parameter
in Ausdrücken 51
Prädiktoreinflussdiagramm 60
Profit
maximieren 11
simulieren 17
Profittediagramme 17, 69
Profitsimulation in Modellen 71
Projektdatenmodell 35
Projektdatenquellen 35
Projekte sperren 33

Q

QuickInfos 95

R

Regeln 53
Anmerkungen 57
Auswahl 55
Auswahllogik 55
Erstellung 53
Export 56
externe 58
gemeinsame 56, 58
IBM Operational Decision Manage-
ment 58
Modelle verwenden mit 53
Segment 53
wiederverwenden 56, 58
Regeln ausschließen 55
Regeln einschließen 55
ROI-Diagramme 69

S

Schlüsselwörter 31, 33
Scoring-Felder abbilden 80
Scoring von Modellen 21, 75, 82
Analytic Server-Ziel 79
Ausgabefelder auswählen 81
Cognos BI-Server als Ziel 79
Dateiziel 78
Daten auswählen 77
Datenbankziel 77
Felder abbilden 80
Scoring-Ziel auswählen 77
Scoring-Ziel in Modellen 77
Scoring-Zusammenfassungsbericht 82

Segmentregeln
Modelle verwenden mit 53
Simulation
Standarddatum 29
Simulation des Profits in Modellen 71
Simulationsdatum 29
Sperren der Datenoptionen 87
Sperren der Optionen zum Scoring-
Ziel 88
Sprache 27
Startseite 13, 27, 86
Anpassung 27
Anwendungen hinzufügen 27
Steuerung
allgemeine Optionen 86
Streams
hochladen 31
Streams zwischen IBM Analytical Deci-
sion Management und IBM SPSS Modeler
gemeinsam nutzen 97
Symbole, IBM Cognos BI 44

T

Testen von Anwendungen 72
Testen von Modellen 21
Textbasierte Datendateien 38
Trefferdiagramme 69

U

Überblick
Verwaltung 85
URL
Anwendung 27

V

Verfügbare Dokumentation 95
Verknüpfungen
Anwendung 27
Versionsbeschriftungen 32
Verteilungsdiagramm 60
Verteilungsdiagramme 69
Verwaltung
allgemeine Optionen 86
Anwendungsverknüpfungen 86
Datenoptionen 87
Homepage 87
Optionen für das Scoring-Ziel 88
Überblick 85
Verwaltung, Überblick 85
Vordefinierte Anwendungen 12
Voreinstellungen 29
Vorhersagemodelle 59
Vorhersageregelnmodelle 61, 67
Vorkonfiguration von Anwendungsver-
knüpfungen 86
Vorschau von Daten 45

W

Währungsformate 29

Z

Zeitformate 29

Zielfeld

 Modelle erstellen 15, 59

Zusatzfelder hinzufügen/bearbeiten 48

