

IBM SPSS Analytic Server
Version 1

Benutzerhandbuch

IBM

Hinweis

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen“ auf Seite 19 gelesen werden.

Produktinformation

Diese Ausgabe bezieht sich auf Version 1 Release 0, Modifikation 0 von IBM SPSS Analytic Server und alle nachfolgenden Releases und Modifikationen, bis dieser Hinweis in einer Neuausgabe geändert wird.

Diese Veröffentlichung ist eine Übersetzung des Handbuchs
IBM SPSS Analytic Server, Version 1, User's Guide,
herausgegeben von International Business Machines Corporation, USA

© Copyright International Business Machines Corporation 2013

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
April 2013

Inhaltsverzeichnis

Kapitel 1. Übersicht.	1	Datenquellen	9
Architektur	2	Einstellungen (Dateidatenquellen)	13
		Vorschau und Metadaten (Datenquellen)	15
		Projekte	15
Kapitel 2. SPSS Modeler-Integration.	3	Bemerkungen.	19
Analytic Server-Quelle	3	Marken.	21
Analytic Server - Datenquelle auswählen	3		
Analytic Server-Export	4	Index	23
Analytic Server-Streameigenschaften	4		
Unterstützte Knoten	5		
Kapitel 3. Analytic Server-Konsole	9		

Kapitel 1. Übersicht

IBM® SPSS Analytic Server ist eine Lösung für die Big Data-Analyse, bei der IBM SPSS-Technologie mit Big Data-Systemen kombiniert wird und die Ihnen die Arbeit mit vertrauten IBM SPSS-Benutzeroberflächen ermöglicht, um bisher unlösbare Probleme zu lösen.

Warum Big Data-Analyse von Bedeutung ist

Die Datenvolumen, die von Organisationen erfasst werden, nehmen exponentiell zu. Bei Finanz- und Einzelhandelsunternehmen sind dies z. B. alle Kundentransaktionen in einem Jahr (bzw. zwei oder zehn Jahren), bei Telekommunikations Providern sind dies Anrufrufdatensätze und Gerätesensormesswerte und bei Internetunternehmen sind dies die Ergebnisse von Websuchen.

Eine Big Data-Analyse ist erforderlich, wenn Folgendes vorhanden ist:

- Ein großes Datenvolumen (Terabyte, Petabyte, Exabyte), vor allem, wenn es sich um eine Mischung aus strukturierten und unstrukturierten Daten handelt
- Sich schnell ändernde/summierende Daten

Eine Big Data-Analyse ist außerdem hilfreich, wenn Folgendes gilt:

- Es wird eine große Anzahl (Tausende) von Modellen erstellt
- Modelle werden häufig erstellt/aktualisiert

Herausforderungen

Dieselben Organisationen, die große Datenvolumen erfassen, haben aus den folgenden Gründen häufig Schwierigkeiten bei der eigentlichen Nutzung dieser Daten:

- Die Architektur konventioneller Analyseprodukte ist nicht für die verteilte Verarbeitung geeignet.
- Vorhandene Statistikalgorithmen sind nicht für die Arbeit mit großen Datenmengen und großer Datenvielfalt (Big Data) bestimmt (bei diesen Algorithmen wird erwartet, dass ihnen die Daten zugeführt werden, das Übertragen von Big Data ist jedoch zu aufwendig). Daher gilt Folgendes:
- Für die Durchführung einer technologisch ausgereiften Analyse von Big Data sind neue Kenntnisse und Detailwissen in Bezug auf Big Data-Systeme erforderlich. Sehr wenige Analysten verfügen über diese Kenntnisse.
- Speicherinterne Lösungen können für Probleme mittlerer Größe verwendet werden, sind für wirklich große Datenmengen und -vielfalt aber nicht gut geeignet.

Lösung

Analytic Server stellt Folgendes bereit:

- Eine datenorientierte Architektur, die Big Data-Systeme nutzt, z. B. Hadoop Map/Reduce mit Daten in HDFS.
- Eine definierte Schnittstelle, die neue statistische Algorithmen integriert, die so entwickelt sind, dass sie sich zu den Daten hinbewegen.
- Vertraute IBM SPSS-Benutzeroberflächen, die die Details der Big Data-Umgebungen ausblenden, sodass sich die Analysten auf die Analyse der Daten konzentrieren können.
- Eine Lösung, die für ein Problem beliebiger Größe skalierbar ist.

Architektur

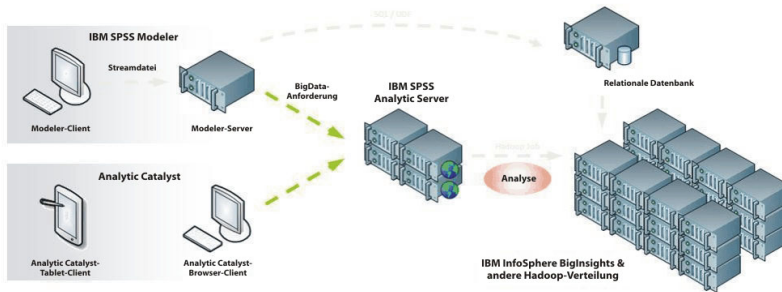


Abbildung 1. Architektur

Analytic Server befindet sich zwischen einer Clientanwendung und einer Hadoop-Cloud. Vorausgesetzt, die Daten befinden sich in der Cloud, gilt für das Arbeiten mit Analytic Server das folgende Schema:

1. Definieren Sie Analytic Server-Datenquellen über die Daten in der Cloud.
2. Definieren Sie die Analyse, die Sie in der Clientanwendung ausführen möchten. Für das aktuelle Release handelt es sich bei den Clientanwendungen um IBM SPSS Modeler und IBM SPSS Analytic Catalyst.
3. Wenn Sie die Analyse ausführen, übergibt die Clientanwendung eine Analytic Server-Ausführungsanforderung.
4. Analytic Server koordiniert den Job zur Ausführung in der Hadoop-Cloud und meldet die Ergebnisse an die Clientanwendung.
5. Sie können die Ergebnisse zum Definieren weiterer Analysen verwenden und so wiederholt sich der Zyklus.

Kapitel 2. SPSS Modeler-Integration

SPSS Modeler ist eine Data-Mining-Workbench, die über einen visuellen Ansatz für die Analyse verfügt. Jede einzelne Aktion in einem Job, vom Zugriff auf eine Datenquelle über die Zusammenführung von Datensätzen bis zur Ausgabe einer neuen Datei oder zur Erstellung eines Modells, wird durch einen Knoten im Erstellungsbereich dargestellt. Diese Aktionen werden miteinander verknüpft, damit ein analytischer Stream gebildet wird.

Zum Erstellen eines SPSS Modeler-Streams, der in HDFS ausgeführt werden kann, beginnen Sie mit einem Knoten für die „Analytic Server-Quelle“. SPSS Modeler überträgt einen möglichst großen Teil des Streams mit einer Pushback-Operation zurück zu Analytic Server und führt dann, falls erforderlich, eine Pull-Operation für ein Subset der Datensätze von HDFS durch, um die Ausführung des Streams "lokal" auf dem SPSS Modeler-Server zu beenden. Sie können die maximale Anzahl an Datensätzen, die SPSS Modeler herunterlädt, in den „Analytic Server-Streameigenschaften“ auf Seite 4 festlegen.

Wenn Ihre Analyse beendet ist und die Datensätze wieder in HDFS geschrieben wurden, beenden Sie den Stream mit einem Knoten für den „Analytic Server-Export“ auf Seite 4.

Analytic Server-Quelle

Mithilfe der Analytic Server-Quelle können Sie einen Stream in HDFS (Hadoop Distributed File System) ausführen. Die Informationen in einer Analytic Server-Datenquelle können aus verschiedenen Quellen stammen. Dazu gehören:

- Textdateien in HDFS
- Datenbanken
- HCatalog

Normalerweise wird ein Stream mit einer Analytic Server-Quelle in HDFS ausgeführt. Wenn ein Stream jedoch einen Knoten enthält, der nicht für die Ausführung in HDFS unterstützt wird, wird ein möglichst großer Teil des Streams mit einer Pushback-Operation zurück zu Analytic Server übertragen und SPSS Modeler Server versucht anschließend, den restlichen Stream zu verarbeiten. Sie müssen für sehr große Datasets eine Teilstichprobe durchführen, indem Sie z. B. einen Stichprobenknoten im Stream platzieren.

Data source. Wenn Ihr SPSS Modeler Server-Administrator eine Verbindung hergestellt hat, wählen Sie eine Datenquelle aus, die die Daten enthält, die Sie verwenden wollen. Eine Datenquelle enthält die zu dieser Quelle zugehörigen Dateien und Metadaten. Klicken Sie auf **Select**, um eine Liste der verfügbaren Datenquellen anzuzeigen. Weitere Informationen finden Sie im Abschnitt „Analytic Server - Datenquelle auswählen“.

Wenn Sie eine neue Datenquelle erstellen müssen oder eine vorhandene Datenquelle bearbeiten müssen, klicken Sie auf **Launch Data Source Editor...** Weitere Informationen zum Erstellen und Bearbeiten von Datenquellen finden Sie in „Datenquellen“ auf Seite 9.

Analytic Server - Datenquelle auswählen

In der Tabelle **Data Sources** wird eine Liste der verfügbaren Datenquellen angezeigt. Wählen Sie die Quelle aus, die Sie verwenden möchten, und klicken Sie auf **OK**.

Klicken Sie auf **Show Owner**, um den Datenquelleneigner anzuzeigen.

Mit **Filter by** können Sie die Datenquellenliste nach Schlüsselwort (**Keyword**) filtern, wodurch die Filterkriterien mit dem Datenquellennamen und der Datenquellenbeschreibung oder dem Eigner (**Owner**) ver-

glichen werden. Sie können eine Kombination aus Zeichenfolgewart und numerischem Wert oder das Platzhalterzeichen (%) als Filterkriterien eingeben. Bei dem Suchbegriff muss die Groß-/Kleinschreibung beachtet werden. Klicken Sie auf **Refresh**, um die Tabelle **Data Sources** zu aktualisieren.

Analytic Server-Export

Mit dem Analytic Server-Export können Sie Daten aus Ihrer Analyse in eine vorhandene Analytic Server-Datenquelle schreiben. Dies können z. B. Textdateien in HDFS (Hadoop Distributed File System) oder eine Datenbank sein.

Normalerweise beginnt ein Stream mit einem Analytic Server-Exportknoten auch mit Analytic Server-Quellenknoten und er wird an Analytic Server übergeben und in HDFS ausgeführt. Alternativ kann ein Stream mit "lokalen" Datenquellen mit einem Analytic Server-Exportknoten enden, damit relativ kleine Datasets (nicht mehr als 100.000 Datensätze) für die Verwendung mit Analytic Server hochgeladen werden können.

Data source. Wählen Sie eine Datenquelle aus, die die Daten enthält, die Sie verwenden wollen. Eine Datenquelle enthält die zu dieser Quelle zugehörigen Dateien und Metadaten. Klicken Sie auf **Select**, um eine Liste der verfügbaren Datenquellen anzuzeigen. Weitere Informationen finden Sie im Abschnitt „Analytic Server - Datenquelle auswählen“ auf Seite 3.

Wenn Sie eine neue Datenquelle erstellen müssen oder eine vorhandene Datenquelle bearbeiten müssen, klicken Sie auf **Launch Data Source Editor...** Weitere Informationen zum Erstellen und Bearbeiten von Datenquellen finden Sie in „Datenquellen“ auf Seite 9.

Mode. Wählen Sie **Append** aus, um der vorhandenen Datenquelle Daten hinzuzufügen, oder **Override**, um den Inhalt der Datenquelle zu ersetzen.

Generate an Import node for this data. Wählen Sie diese Option aus, um einen Quellenknoten für die Daten zu generieren, die in die angegebene Datenquelle exportiert werden. Dieser Knoten wird dem Streamerstellungsbereich hinzugefügt.

Analytic Server-Streameigenschaften

Diese Einstellungen stellen eine Reihe von Optionen für die Arbeit mit Analytic Server bereit.

Maximum number of records to process outside of Analytic Server

Geben Sie die maximale Anzahl der Datensätze an, die von einer Analytic Server-Datenquelle auf den SPSS Modeler-Server importiert werden sollen.

Notification when a node can't be processed in Analytic Server

Diese Einstellung legt fest, was geschieht, wenn ein Stream, der an Analytic Server übergeben würde, einen Knoten enthält, der nicht in Analytic Server verarbeitet werden kann. Geben Sie an, ob eine Warnung ausgegeben werden soll und die Verarbeitung des Streams fortgesetzt werden soll oder ob ein Fehler ausgelöst werden soll und die Verarbeitung gestoppt werden soll.

Split Model Storage Settings

Store split models by reference on Analytic Server when model size (MB) exceeds

Modellnuggets werden normalerweise als Teil des Streams gespeichert. Aufgeteilte Modelle mit vielen Aufteilungen können große Nuggets erzeugen und wenn das Nugget zwischen dem Stream und Analytic Server hin- und hergeschoben wird, kann sich dies auf die Leistung auswirken. Wenn ein aufgeteiltes Modell die angegebene Größe überschreitet, wird es als Lösung in Analytic Server gespeichert und das Nugget in SPSS Modeler enthält eine Referenz auf das Modell.

Default folder to store models by reference on Analytic Server once execution is complete

Geben Sie den Standardpfad an, in dem aufgeteilte Modelle in Analytic Server gespeichert werden sollen. Der Pfad sollte mit einem gültigen Analytic Server-Projektnamen beginnen.

Folder to store promoted models

Geben Sie den Standardpfad an, in dem "hochgestufte" Modelle gespeichert werden sollen. Ein hochgestuftes Modell wird nicht gelöscht, wenn die SPSS Modeler-Sitzung vorbei ist.

Unterstützte Knoten

Viele SPSS Modeler-Knoten werden für die Ausführung in HDFS unterstützt, bei der Ausführung bestimmter Knoten gibt es jedoch möglicherweise einige Unterschiede und einige Knoten werden zurzeit nicht unterstützt. In diesem Thema wird die aktuelle Unterstützungsstufe detailliert beschrieben.

Allgemein

- Einige Zeichen, die normalerweise in einem Modeler-Feldnamen in Anführungszeichen zulässig sind, werden von Analytic Server nicht akzeptiert.
- Damit ein Modeler-Stream in Analytic Server ausgeführt werden kann, muss er mit mindestens einem Analytic Server-Quellenknoten beginnen und mit einem einzelnen Modellierungsknoten oder Analytic Server-Exportknoten enden. Zusammenführungen sind zulässig, eine Verzweigung jedoch nicht.
- Es wird empfohlen, den Speicher von stetigen Zielen als Speicher für reelle Zahlen und nicht als Speicher für ganze Zahlen festzulegen. Scoring-Modelle schreiben immer reelle Werte in die Ausgabedatendateien für stetige Ziele, während das Ausgabedatenmodell für die Scores dem Speicher des Ziels folgt. Wenn ein stetiges Ziel über einen Speicher für ganze Zahlen verfügt, gibt es daher eine Diskrepanz zwischen den geschriebenen Werten und dem Datenmodell für die Scores und diese Diskrepanz führt zu Fehlern, wenn Sie versuchen, die gescorten Daten zu lesen.

Quelle

- Ein Stream, der mit etwas anderem als einem Analytic Server-Quellenknoten beginnt, wird lokal ausgeführt.

Datensatzoperationen

Es werden alle Datensatzoperationen unterstützt. Weitere Hinweise zur Funktion dieser Knoten folgen.

Select

- Unterstützt dieselbe Funktionsgruppe, die auch vom Knoten **Derive** unterstützt wird (siehe unten).

Sample

- Stichprobenziehung auf Blockebene wird nicht unterstützt.
- Komplexe Methoden der Stichprobenziehung werden nicht unterstützt.

Aggregate

- Zusammenhängende Schlüssel werden nicht unterstützt.
- Reihenfolgestatistiken (Median, 1st Quartile, 3rd Quartile) werden nicht unterstützt.

Sort

- Die Registerkarte **Optimization** wird nicht unterstützt.

Merge

- **Merge by Order** wird nicht unterstützt.
- **Merge by Condition** wird nicht unterstützt.

- Die Registerkarte **Optimization** wird nicht unterstützt.
- Das Platzieren eines Stichprobenknotens oder eines Modellnuggets zwischen einem Analytic Server-Quellenknoten und einem Zusammenführungsknoten wird zurzeit nicht unterstützt. Normalerweise ist es möglich, einen Auswahlknoten anzugeben, um die Funktion des Stichprobenknotens zu ersetzen.
- Analytic Server führt bei Schlüsseln für leere Zeichenfolgen keinen Join durch. Wenn also einer der Schlüssel, mit dem Sie die Zusammenführung durchführen, leere Zeichenfolgen enthält, werden alle Datensätze, die die leere Zeichenfolge enthalten, aus der zusammengeführten Ausgabe gelöscht.
- Zusammenführungsoperationen sind relativ langsam. Wenn in HDFS Speicherplatz verfügbar ist, ist es unter Umständen weniger zeitintensiv, wenn Sie Ihre Datenquellen einmal zusammenzuführen und die zusammengeführte Quelle in den folgenden Streams zu verwenden, anstatt die Datenquellen in jedem Stream zusammenzuführen.

Feldoperationen

Die Knoten **Auto Data Prep**, **Type**, **Filter**, **Derive**, **Ensemble**, **Filler**, **Reclassify**, **Binning**, **RFM Analysis**, **Partition**, **SetToFlag**, **Restructure** und **Field Reorder** werden unterstützt. Weitere Hinweise zur Funktion dieser Knoten folgen.

Auto Data Prep

- Das Trainieren des Knotens wird nicht unterstützt. Die Anwendung der Transformationen in einem trainierten Knoten des Typs **Auto Data Prep** auf neue Daten wird unterstützt.

Type

- Die Prüfspalte wird nicht unterstützt.
- Die Registerkarte **Format** wird nicht unterstützt.

Derive

- Alle Ableitungsfunktionen werden unterstützt, mit Ausnahme von Sequenzfunktionen.
- Aufteilungsfelder können nicht in demselben Stream abgeleitet werden, der sie als Aufteilungen verwendet. Sie müssen zwei Streams erstellen: einen, der das Aufteilungsfeld ableitet, und einen, der das Feld als Aufteilungen verwendet.
- Ein Flagfeld kann nicht allein in einem Vergleich verwendet werden. Das heißt, dass `if (flagField) then ... endif` einen Fehler verursacht. Als Fehlerumgehung kann `if (flagField=trueValue) then ... endif` verwendet werden.
- Wenn der Operator ****** verwendet wird, wird empfohlen, den Exponenten als reelle Zahl anzugeben, z. B. $x^{2,0}$ anstelle von x^2 , damit die Ergebnisse mit den Ergebnissen in Modeler übereinstimmen.

Filler

- Unterstützt dieselbe Funktionsgruppe, die auch vom Ableitungsknoten unterstützt wird (siehe oben).

RFM Analysis

- Die Option "Keep in current" für die Handhabung von Bindungen wird nicht unterstützt. RFM-Aktualitäts-, Häufigkeits- und Geldwertscores stimmen nicht immer mit denen überein, die von Modeler aus denselben Daten berechnet werden. Die Scorebereiche sind identisch, Scorezuweisungen (Klassennummern) können sich jedoch um 1 unterscheiden.

Graphs

Alle Diagrammknoten werden unterstützt.

Modeling

Es wird eine begrenzte Anzahl an Modellierungsknoten unterstützt. Beispiele für unterstützte Knoten: Linear, Neural Net, C&RT, Chaid, Quest. Weitere Hinweise zur Funktion dieser Knoten folgen.

Linear

- Fortlaufendes Training vorhandener PSM-Modelle wird nicht unterstützt.
- Das Modellerstellungsziel **Standard** wird nur empfohlen, wenn Aufteilungsfelder so definiert sind, dass die Anzahl an Datensätzen in den einzelnen Aufteilungen nicht "zu groß" ist, wobei die Definition von "zu groß" von der Leistungsstärke einzelner Knoten in Ihrem Hadoop-Cluster abhängt. Im Gegensatz dazu müssen Sie auch darauf bedacht sein, sicherzustellen, dass Aufteilungen nicht so fein definiert sind, dass zu wenige Datensätze für die Erstellung eines Modells vorhanden sind.
- Das Ziel **Boosting** wird nicht unterstützt.
- Das Ziel **Bagging** wird nicht unterstützt.
- Das Ziel **Very large datasets** wird nicht empfohlen, wenn wenige Datensätze vorhanden sind. Oft wird dann entweder kein Modell oder ein vermindertes Modell erstellt. Sie haben möglicherweise auch Probleme, wenn die Eingabedatensätze nach einer Systematik geordnet werden, die gegen die Zufälligkeitsvoraussetzungen hinter den verwendeten Erstellungsalgorithmen für Ensemblemodelle verstößt.
- **Automatic Data Preparation** wird nicht unterstützt. Dies kann Probleme verursachen, wenn versucht wird, anhand von Daten mit vielen fehlenden Werten ein Modell zu erstellen. Normalerweise würden diese als Teil der automatischen Datenaufbereitung imputiert. Als Problemumgehung kann ein Baummodell oder ein neuronales Netz mit der Einstellung **Advanced** verwendet werden, um fehlende ausgewählte Werte zu imputieren.
- Die Genauigkeitsstatistik wird für aufgeteilte Modelle nicht berechnet.

Neural Net

- Fortlaufendes Training vorhandener Standard- oder PSM-Modelle wird nicht unterstützt.
- Das Modellerstellungsziel **Standard** wird nur empfohlen, wenn Aufteilungsfelder so definiert sind, dass die Anzahl an Datensätzen in den einzelnen Aufteilungen nicht "zu groß" ist, wobei die Definition von "zu groß" von der Leistungsstärke einzelner Knoten in Ihrem Hadoop-Cluster abhängt. Im Gegensatz dazu müssen Sie auch darauf bedacht sein, sicherzustellen, dass Aufteilungen nicht so fein definiert sind, dass zu wenige Datensätze für die Erstellung eines Modells vorhanden sind.
- Das Ziel **Boosting** wird nicht unterstützt.
- Das Ziel **Bagging** wird nicht unterstützt.
- Das Ziel **Very large datasets** wird nicht empfohlen, wenn wenige Datensätze vorhanden sind. Oft wird dann entweder kein Modell oder ein vermindertes Modell erstellt. Sie haben möglicherweise auch Probleme, wenn die Eingabedatensätze nach einer Systematik geordnet werden, die gegen die Zufälligkeitsvoraussetzungen hinter den verwendeten Erstellungsalgorithmen für Ensemblemodelle verstößt.
- Wenn in den Daten viele Werte fehlen, verwenden Sie die Einstellung **Advanced**, um fehlende Werte zu imputieren.
- Die Genauigkeitsstatistik wird für aufgeteilte Modelle nicht berechnet.

'C&R Tree', 'CHAID' und 'Quest'

- Fortlaufendes Training vorhandener PSM-Modelle wird nicht unterstützt.
- Das Modellerstellungsziel **Standard** wird nur empfohlen, wenn Aufteilungsfelder so definiert sind, dass die Anzahl an Datensätzen in den einzelnen Aufteilungen nicht "zu groß" ist, wobei die Definition von "zu groß" von der Leistungsstärke einzelner Knoten in Ihrem Hadoop-Cluster abhängt. Im Gegensatz dazu müssen Sie auch darauf bedacht sein, sicherzustellen, dass Aufteilungen nicht so fein definiert sind, dass zu wenige Datensätze für die Erstellung eines Modells vorhanden sind.
- Das Ziel **Boosting** wird nicht unterstützt.
- Das Ziel **Bagging** wird nicht unterstützt.
- Das Ziel **Very large datasets** wird nicht empfohlen, wenn wenige Datensätze vorhanden sind. Oft wird dann entweder kein Modell oder ein vermindertes Modell erstellt. Sie haben möglicherweise auch Probleme, wenn die Eingabedatensätze nach einer Systematik geordnet werden, die gegen die Zufälligkeitsvoraussetzungen hinter den verwendeten Erstellungsalgorithmen für Ensemblemodelle verstößt.
- Interaktive Sitzungen werden nicht unterstützt.
- Die Genauigkeitsstatistik wird für aufgeteilte Modelle nicht berechnet.

Modellscoring

Die folgenden Modellnuggets werden für das Scoring unterstützt: Linear, Neural Net, C&RT, Chaid, Quest, Logistic, Regression, Genlin, TwoStep Cluster, C5, Bayesian Networks, SVM, R, Text Mining.

- Raw Propensity und Adjusted Propensity werden nicht gescort. Als Problemumgehung können Sie denselben Effekt erzielen, indem Sie die Raw Propensity mithilfe eines Ableitungsknotens mit dem folgenden Ausdruck berechnen: `if 'predicted-value' == 'value-of-interest' then 'prob-of-that-value' else 1-'prob-of-that-value' endif`
- Beim Scoring eines Modells überprüft Analytic Server nicht, ob alle im Modell verwendeten Felder im Dataset vorhanden sind. Stellen Sie daher vor der Ausführung in Analytic Server sicher, dass dies der Fall ist.

Ausgabe

Die Knoten **Matrix**, **Analysis**, **Data Audit**, **Transform**, **Statistics** und **Means** werden unterstützt.

Export Ein Stream kann mit einem Analytic Server-Quellenknoten beginnen und mit einem anderen Exportknoten als dem Analytic Server-Exportknoten enden, die Daten werden jedoch von HDFS in SPSS Modeler Server und schließlich an die Exportposition verschoben.

Kapitel 3. Analytic Server-Konsole

Analytic Server stellt eine Thin-Client-Schnittstelle zum Verwalten von Datenquellen und Projekten bereit, auf die auf zwei Arten zugegriffen werden kann:

- Über den Quellen- oder Exportknoten von Analytic Server, der die Konsole in einem Browser öffnet
- Direkt über einen Browser
 1. Geben Sie die URL von Analytic Server in die Adressleiste Ihres Browsers ein. Sie können diese von Ihrem Serveradministrator erhalten.
 2. Geben Sie den Benutzernamen ein, mit dem die Anmeldung am Server erfolgen soll.
 3. Geben Sie das Kennwort ein, das dem angegebenen Benutzernamen zugeordnet ist.

Nach der Anmeldung ist das Datenquellenakkordeon das standardmäßige Akkordeon.

Navigieren in der Konsole

Die Analytic Server-Konsole besteht aus vier Komponenten:

- In der Kopfzeile werden der Produktname und der Name des zurzeit angemeldeten Benutzers angezeigt.
- In der linken Spalte werden die verfügbaren Akkordeons oder funktionalen Gruppierungen angezeigt. Das ausgewählte Akkordeon bestimmt, was im Inhaltsbereich angezeigt wird.
- Inhaltsbereich. Zeigt die Steuerelemente an, die dem zurzeit ausgewählten Akkordeon zugeordnet sind. Die folgenden Abschnitte enthalten Details zum Inhalt der Akkordeons.
- Die Fußzeile enthält den Link zum Abmelden und zeigt die installierte Version von Analytic Server.

Datenquellen

Eine Datenquelle besteht aus einer Sammlung von Datensätzen und einem Datenmodell, die ein Dataset für die Analyse definieren. Die Quelle von Datensätzen kann eine Datei (Text mit Trennzeichen, Text mit fester Breite, Excel) in HDFS, eine Datenbank oder HCatalog sein. Das Datenmodell definiert alle Metadaten (Feldnamen, Speicher, Messniveau usw.), die für die Analyse von Daten erforderlich sind. Datenquelleneigner können Zugriff auf Datenquellen erteilen oder einschränken.

Linke Spalte

In der linken Spalte werden die vorhandenen Datenquellen unter der Akkordeonüberschrift angezeigt.

- Wählen Sie eine Datenquelle aus, um ihre Details im Inhaltsbereich anzuzeigen und ihre Eigenschaften zu bearbeiten. Bei der Eingabe eines Suchbegriffs in den Suchbereich wird die Liste gefiltert, sodass nur Datenquellen angezeigt werden, deren Name den Suchbegriff enthält.
- Klicken Sie auf die Schaltfläche **New data source**, um eine neue Datenquelle mit dem Namen und dem Inhaltstyp zu erstellen, den Sie im Dialogfeld **Add New Data Source** angeben.
 - Bei Datenquellennamen muss die Groß-/Kleinschreibung beachtet werden. Führende und folgende Leerzeichen werden ignoriert. Bestimmte Namen werden abgelehnt, um einen Schutz vor SQL-Injection bereitzustellen.
 - Die verfügbaren Inhaltstypen sind **File**, **Database** und **HCatalog**.

Anmerkung: Wenn der Inhaltstyp ausgewählt wurde, kann er nicht mehr bearbeitet werden.

- Klicken Sie auf die Schaltfläche **Delete data source**, um die Datenquelle zu entfernen. Bei dieser Aktion bleiben alle Dateien, die der Datenquelle zugeordnet sind, intakt.

Inhaltsbereich

Der Inhaltsbereich ist in mehrere Abschnitte unterteilt, die vom Inhaltstyp der Datenquelle abhängen können. Nachdem Sie die Einstellungen für die Datenquelle angegeben haben, klicken Sie auf **Preview and Metadata**, um die Datenquelle fertigzustellen.

Data Source Properties

Einstellungen, die für alle Inhaltstypen gleich sind.

Name Ein bearbeitbares Textfeld, in dem der Name der Datenquelle angezeigt wird. Bei Datenquellennamen muss die Groß-/Kleinschreibung beachtet werden. Führende und folgende Leerzeichen werden ignoriert. Bestimmte Namen werden möglicherweise abgelehnt, um einen Schutz vor SQL-Injection bereitzustellen.

Description

Ein bearbeitbares Textfeld, in dem Sie einen erläuternden Text zur Datenquelle angeben können.

Is public

Ein Kontrollkästchen, das angibt, ob jeder das Projekt sehen kann (aktiviert) oder ob Benutzer und Gruppen explizit der Eignerliste hinzugefügt werden müssen (inaktiviert).

Sharing

Sie können das Eigentumsrecht für eine Datenquelle freigeben, indem Sie Benutzer und Gruppen als Autoren hinzufügen.

- Durch Eingeben eines Suchbegriffs in das Textfeld wird nach Benutzern und Gruppen gefiltert, deren Name den Suchbegriff enthält. Klicken Sie auf die Schaltfläche **Add participant**, um sie der Liste der Autoren hinzuzufügen.
- Zum Entfernen eines Autors wählen Sie einen Benutzer oder eine Gruppe in der Autorenliste aus und klicken Sie auf die Schaltfläche **Remove participant**.

Anmerkung: Administratoren verfügen über Lese- und Schreibzugriff auf jede Datenquelle, unabhängig davon, ob sie als Autor aufgelistet sind.

File Input

Einstellungen, die für die Definition von Datenquellen mit dem Inhaltstyp **File** spezifisch sind.

File Viewer

Hier werden für den Einschluss in die Datenquelle verfügbare Dateien angezeigt. Wählen Sie den Modus **Projects** aus, um Dateien in der Analytic Server-Projektstruktur anzuzeigen, oder **HDFS**, um das restliche verteilte Hadoop-Dateisystem anzuzeigen. Sie können beide Ordnerstrukturen durchsuchen, HDFS ist jedoch überhaupt nicht bearbeitbar und die Analytic Server-Ordnerstruktur ist nur in definierten Projekten bearbeitbar. Sie können also auf der Stammebene des Modus **Projects** keine Dateien hinzufügen, keine Ordner erstellen und keine Elemente löschen. Verwenden Sie zum Erstellen, Bearbeiten oder Löschen eines Projekts das Projektakkordeon.

- Durch Klicken auf **Upload file to HDFS** wird eine Datei in das aktuelle Projekt/den aktuellen Unterordner heruntergeladen.
- Durch Klicken auf **Create a new folder** wird unter dem aktuellen Ordner ein neuer Ordner mit dem Namen erstellt, den Sie im Dialogfeld **New Folder Name** angeben.
- Durch Klicken auf die Schaltfläche **Download file to the local filesystem** werden die ausgewählten Dateien in das lokale Dateisystem heruntergeladen.
- Durch Klicken auf **Delete the selected file(s)** werden die ausgewählten Dateien/Ordner entfernt.

Files included in data source definition

Verwenden Sie die Pfeilschaltfläche, um ausgewählte Dateien der Datenquelle hinzuzufügen.

gen oder aus ihr zu entfernen. Klicken Sie für jede ausgewählte Datei in der Datenquelle auf Settings, um die Spezifikationen für das Lesen der Datei zu definieren.

File Output

An Datenquellen mit dem Inhaltstyp **File** kann die Ausgabe von Streams, die in Analytic Server ausgeführt werden, angehängt werden. Wählen Sie **Make writeable** aus, um das Anhängen zu aktivieren, und wählen Sie einen Ausgabeordner aus, in den die neuen Dateien geschrieben werden.

Database Selections

Geben Sie die Verbindungsparameter für die Datenbank an, die den Datensatzinhalt enthält.

Database

Wählen Sie den Datenbanktyp aus, zu dem Sie eine Verbindung herstellen wollen. Folgendes steht zur Auswahl: DB2, Oracle, SQL Server, TeraData oder Netezza.

Server address

Geben Sie die URL des Servers an, auf dem sich die Datenbank befindet.

Server port

Die Nummer des Ports, an dem die Datenbank empfangsbereit ist.

Database name

Der Name der Datenbank, zu der Sie eine Verbindung herstellen wollen.

Username

Wenn die Datenbank kennwortgeschützt ist, geben Sie Ihren Benutzernamen ein.

Password

Wenn die Datenbank kennwortgeschützt ist, geben Sie Ihr Kennwort ein.

Table name

Geben Sie den Namen einer Tabelle aus der Datenbank an, die Sie verwenden wollen.

Maximum concurrent reads

Database Output

Datenquellen mit dem Inhaltstyp **Database** können über die Ausgabe von Streams angehängt werden, die auf Analytic Server ausgeführt werden. Wählen Sie **Make writeable** aus, um das Anhängen zu aktivieren, und wählen Sie eine Ausgabedatenbanktabelle aus, in die die Ausgabedaten geschrieben werden.

HCatalog Selections

Geben Sie die Parameter für den Zugriff auf Daten an, die unter Apache HCatalog verwaltet werden.

Database

Der Name der HCatalog-Datenbank.

Table name

Geben Sie den Namen einer Tabelle aus der Datenbank an, die Sie verwenden wollen.

Filter Der Partitionsfilter für die Tabelle, wenn die Tabelle als partitionierte Tabelle erstellt wurde.

HCatalog Schema

Zeigt die Struktur der angegebenen Tabelle an. HCatalog kann ein hochgradig strukturiertes Dataset unterstützen. Wenn eine Analytic Server-Datenquelle für solche Daten definiert werden soll, muss die Struktur in einfache Zeilen und Spalten abgeflacht werden. Wählen Sie ein Element im Schema aus und klicken Sie auf die Pfeilschaltfläche, um es für die Analyse einem Feld zuzuordnen. Nicht alle Baumknoten können zugeordnet werden.

Beispielsweise wird ein Array oder eine Zuordnung komplexer Typen als "übergeordnetes Element" angesehen und kann nicht zugeordnet werden. Diese Knoten sind durch eine Beschriftung im Baum gekennzeichnet, die auf "...:array:struct" oder auf "...:map:struct" endet.

HCatalog Field Mappings

Zeigt die Zuordnung eines Elements in HCatalog zu einem Feld in der Datenquelle an. Klicken Sie auf **Preview Raw Data**, um die Datensätze so anzuzeigen, wie sie in HCatalog gespeichert sind. So können Sie leichter festlegen, wie das HCatalog-Schema Feldern zugeordnet werden soll.

HCatalog Element

Doppelklicken Sie auf eine Zelle, um sie zu bearbeiten. Sie müssen die Zelle bearbeiten, wenn das HCatalog-Element ein Array oder eine Zuordnung ist. Geben Sie mit einem Array die ganze Zahl an, die dem Mitglied des Arrays entspricht, das Sie einem Feld zuordnen wollen. Geben Sie mit einer Zuordnung eine in Anführungszeichen eingeschlossene Zeichenfolge an, die dem Schlüssel entspricht, den Sie einem Feld zuordnen wollen. In Abb. 2 finden Sie ein Beispiel, wie anhand der Rohdatenübersicht die Zeichenfolge festgestellt werden kann, die dem Zuordnungsindex entspricht.

Mapping Field

Das Feld, wie es in der Analytic Server-Datenquelle angezeigt wird. Doppelklicken Sie auf eine Zelle, um sie zu bearbeiten. Doppelte Werte in der Spalte **Mapping Field** sind nicht zulässig und führen zu einem Fehler.

Storage

Der Speicherort des Felds. Der Speicherort wird von HCatalog abgeleitet und kann nicht bearbeitet werden.

Anmerkung: Wenn Sie auf Preview and Metadata klicken, um eine HCatalog-Datenquelle fertigzustellen, stehen keine Bearbeitungsoptionen zur Verfügung.

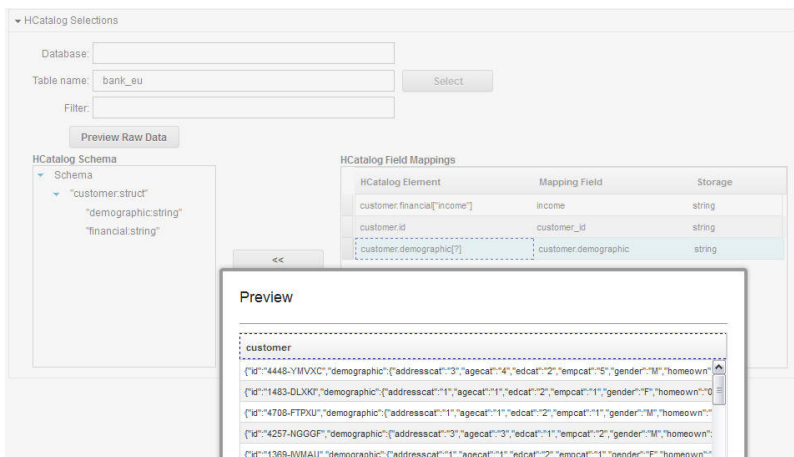


Abbildung 2. Datenquellenakkordeon: Definieren einer HCatalog-Datenquelle

Einstellungen (Dateidatenquellen)

Character set encoding

Die Zeichencodierung der Datei. Wählen Sie einen Java-Zeichensatznamen wie 'UTF-8', 'ISO-8859-2', 'GB18030' usw. aus oder geben Sie diesen an. Der Standardwert ist 'UTF-8'.

Locale Definiert eine Ländereinstellung. Optional. Standardmäßig die Ländereinstellung des Servers. Die Ländereinstellungszeichenfolge sollte wie folgt angegeben werden: <Sprache>[_Land[_Variante]], wobei Folgendes gilt:

Sprache

Ein gültiger, aus zwei Buchstaben bestehender Code in Kleinbuchstaben gemäß ISO-639-Definition. Erforderlich.

Land Ein gültiger, aus zwei Buchstaben bestehender Code in Großbuchstaben gemäß ISO-639-Definition. Optional.

Variante

Ein für den Anbieter oder Browser spezifischer Code. Optional.

Trim white space

Entfernt Leerzeichen am Anfang und/oder Ende der Zeichenfolgefelder. Standardmäßig **None**. Die folgenden Werte werden unterstützt:

None Entfernt Leerzeichen nicht.

Left Entfernt Leerzeichen am Anfang der Zeichenfolge.

Right Entfernt Leerzeichen am Ende der Zeichenfolge.

Both Entfernt Leerzeichen am Anfang und Ende der Zeichenfolge.

Grouping symbols

Legt fest, ob das für die Ländereinstellung spezifische Zeichen, das als Tausendertrennzeichen verwendet wird, verwendet werden sollte.

Field delimiters

Mindestens ein Zeichen, das Feldgrenzen markiert. Jedes Zeichen wird als unabhängiges Trennzeichen gesehen. ",\"\\t\" bedeutet z. B., dass entweder ein Komma oder ein Tabulator Feldgrenzen markiert. Wenn Steuerzeichen Felder begrenzen, werden die hier angegebenen Zeichen neben Steuerzeichen als Trennzeichen betrachtet. Der Standardwert ist \",\", wenn Steuerzeichen keine Felder begrenzen, und andernfalls die leere Zeichenfolge.

Control characters delimit fields

Legt fest, ob ASCII-Steuerzeichen, außer LF und CR, als Feldtrennzeichen betrachtet werden. Standardmäßig **No**.

First row contains field names

Legt fest, ob die erste Zeile für die Festlegung der Feldnamen verwendet werden soll. Standardmäßig **No**.

Number of initial characters to skip

Die Anzahl der Zeichen am Anfang der Datei, die übersprungen werden sollen. Eine nicht negative Ganzzahl. Der Standardwert ist '0' (null).

Merge white space

Legt fest, ob mehrere benachbarte Vorkommen eines Leerzeichens und/oder Tabulators als ein einziges Feldtrennzeichen betrachtet werden. Hat keine Auswirkung, wenn weder das Leerzeichen noch der Tabulator ein Feldtrennzeichen ist. Der Standardwert ist 'true'.

End-of-line comment characters

Mindestens ein Zeichen, das Zeilenendekommentare markiert. Das Zeichen und alles, was im Datensatz darauf folgt, wird, ignoriert. Jedes Zeichen wird als unabhängige Kommentarmarkierung gesehen. "/"* bedeutet z. B., dass ein Kommentar entweder mit einem Schrägstrich oder einem Stern beginnt. Es ist nicht möglich, Kommentarmarkierungen aus mehreren Zeichen zu definieren, z. B. "//". Die leere Zeichenfolge signalisiert, dass keine Kommentarzeichen definiert sind. Wenn Kommentarzeichen definiert sind, werden diese überprüft, bevor Anführungszeichen verarbeitet oder zu überspringende Zeichen am Anfang übersprungen werden. Der Standardwert ist die leere Zeichenfolge.

Invalid characters

Legt fest, wie ungültige Zeichen (Bytesequenzen, die nicht Zeichen in der Codierung entsprechen) behandelt werden sollen. Eine leere Zeichenfolge gibt an, dass sie gelöscht werden sollen. Eine nicht leere Zeichenfolge (in der Regel ein einzelnes Zeichen) gibt an, dass sie durch den Inhalt der Zeichenfolge ersetzt werden sollen. Der Standardwert ist die leere Zeichenfolge.

Single quotes

Gibt die Verarbeitung von einfachen Anführungszeichen (Hochkommas) an. Der Standardwert ist **Keep**.

Keep Hochkommas haben keine besondere Bedeutung und werden wie jedes andere Zeichen behandelt.

Drop Hochkommas werden gelöscht, wenn sie nicht in Anführungszeichen stehen

Pair Hochkommas werden als Anführungszeichen betrachtet und Zeichen zwischen zwei Hochkommas verlieren ihre besondere Bedeutung (sie werden als Zeichen in Anführungszeichen betrachtet). Ob Hochkommas selbst innerhalb von Zeichenfolgen in Hochkommas vorkommen können, wird durch die Einstellung **Quotes can be quoted by doubling** festgelegt.

Double quotes

Gibt die Verarbeitung von Anführungszeichen an. Der Standardwert ist **Pair**.

Keep Anführungszeichen haben keine besondere Bedeutung und werden wie jedes andere Zeichen behandelt.

Drop Anführungszeichen werden gelöscht, wenn sie nicht in Anführungszeichen stehen

Pair Doppelte Anführungszeichen werden als Anführungszeichen betrachtet und Zeichen zwischen zwei Hochkommas verlieren ihre besondere Bedeutung (sie werden als Zeichen in Anführungszeichen betrachtet). Ob doppelte Anführungszeichen selbst innerhalb von Zeichenfolgen in Hochkommas vorkommen können, wird durch die Einstellung **Quotes can be quoted by doubling** festgelegt.

Quotes can be quoted by doubling

Gibt an, ob doppelte Anführungszeichen innerhalb von Zeichenfolgen in Anführungszeichen dargestellt werden können und ob Hochkommas innerhalb von Zeichenfolgen in Hochkommas dargestellt werden können, wenn **Pair** festgelegt ist. Bei Angabe von **Yes** werden doppelte Anführungszeichen innerhalb von Zeichenfolgen in Anführungszeichen und Hochkommas innerhalb von Zeichenfolgen in Hochkommas verdoppelt. Bei Angabe von **No** gibt es keine Möglichkeit, ein Anführungszeichen innerhalb einer Zeichenfolge in Anführungszeichen oder ein Hochkomma innerhalb einer Zeichenfolge in Hochkommas zu setzen. Der Standardwert ist **Yes**.

Vorschau und Metadaten (Datenquellen)

Durch Klicken auf **Preview and Metadata** wird das Datenmodell für die Datenquelle auf der Basis einer Stichprobe der Datensätze generiert. Hier können Sie die grundlegenden Metadateninformationen überprüfen und bearbeiten.

Preview

Auf der Registerkarte **Preview** werden eine kleine Stichprobe der Datensätze und ihre Feldwerte angezeigt.

Edit

Auf der Registerkarte **Edit** können Sie die Hauptfeldmetadaten korrigieren.

Field Doppelklicken Sie auf den Feldnamen, um ihn zu bearbeiten.

Measurement

Dies ist das Messniveau, mit dem Merkmale der Daten in einem bestimmten Feld beschrieben werden.

Role Wird verwendet, um Modellierungsknoten mitzuteilen, ob Felder für einen Computerlernprozess Eingabefelder (Input, Vorhersagefelder) oder Zielfelder (Target, vorherzusagende Felder) sind. **Both** und **None** sind ebenfalls verfügbare Rollen, ebenso wie die Rolle **Partition**, die ein Feld angibt, mit dem Datensätze zu Schulungs-, Test- und Prüfzwecken in separate Stichproben aufgeteilt werden. Der Wert **Split** gibt an, dass für jeden möglichen Wert des Feldes separate Modelle erstellt werden.

Storage

Mit der Option **Storage** wird beschrieben, wie Daten in einem Feld gespeichert werden. In einem Feld mit den Werten 1 und 0 werden z. B. ganzzahlige Daten gespeichert. Dies darf nicht mit dem Messniveau verwechselt werden, das die Verwendung der Daten beschreibt und sich nicht auf die Speicherung auswirkt. Sie können z. B. das Messniveau für ein Feld für ganze Zahlen mit den Werten 1 und 0 auf **Flag** setzen. Dies gibt normalerweise Folgendes an: 1 = True und 0 = False.

Projekte

Projekte sind Arbeitsbereiche, in denen Eingaben gespeichert werden und auf Ausgaben von Jobs zugegriffen wird. Sie stellen die Organisationsstruktur der höchsten Ebene bereit, die Dateien und Ordner enthält. Projekte können mit einzelnen Benutzern und Gruppen gemeinsam genutzt werden.

Linke Spalte

In der linken Spalte werden die vorhandenen Projekte unter der Akkordeonüberschrift angezeigt.

- Wählen Sie ein Projekt aus, um seine Details im Inhaltsbereich anzuzeigen und seine Eigenschaften zu bearbeiten. Bei der Eingabe eines Suchbegriffs in den Suchbereich wird die Liste gefiltert, sodass nur Projekte angezeigt werden, deren Name den Suchbegriff enthält.
- Klicken Sie auf **New Project**, um ein neues Projekt mit dem Namen zu erstellen, den Sie im Dialogfeld **Add New Project** angeben. Bei Namen muss die Groß-/Kleinschreibung beachtet werden, führende und abschließende Leerzeichen müssen ignoriert werden und die Namen müssen vor SQL-Injection geschützt werden.
- Klicken Sie auf **Delete Project**, um das Projekt zu entfernen. Bei dieser Aktion bleiben alle Dateien intakt, die der Datenquelle zugeordnet sind.

Inhaltsbereich

Der Inhaltsbereich enthält die Registerkarten **Settings**, **Owners** und **Version**.

Settings

Project description

Ein bearbeitbares Textfeld, in dem Sie einen erläuternden Text zum Projekt angeben können.

Is public

Ein Kontrollkästchen, das angibt, ob jeder das Projekt sehen kann (aktiviert) oder ob Benutzer und Gruppen explizit als Eigner hinzugefügt werden müssen (inaktiviert).

Durch Klicken auf **Save** wird der aktuelle Status der Einstellungen gespeichert.

Project data sources

Ein nicht bearbeitbarer Bereich, in dem alle Datenquellen aufgelistet werden, die dem Projekt zugeordnet sind.

Projektstrukturbereich

Im rechten Bereich wird die Projekt-/Ordnerstruktur für das zurzeit ausgewählte Projekt angezeigt. Sie können die Ordnerstruktur durchsuchen, sie ist jedoch nur über die Schaltflächen bearbeitbar.

- Klicken Sie auf **Download file to the local filesystem**, um eine ausgewählte Datei in das lokale Dateisystem herunterzuladen.
- Klicken Sie auf **Delete the selected file(s)**, um die ausgewählte Datei/den ausgewählten Ordner zu entfernen.

Owners

Eigner sind Vollmitglieder eines Projekts und können das Projekt sowie die darin enthaltenen Ordner und Dateien ändern.

In der Liste der verfügbaren Benutzer und Gruppen werden die Benutzer und Gruppen des aktiven Nutzers angezeigt, die diesem Projekt zurzeit nicht zugeordnet sind.

- Durch Eingeben eines Suchbegriffs in den Suchbereich wird nach Benutzern und Gruppen gefiltert, deren Name den Suchbegriff enthält.
- Durch Auswählen des Benutzersymbols über der Liste werden die verfügbaren Benutzer angezeigt. Durch Abwählen des Symbols werden die Benutzer ausgeblendet. Dieses Symbol ist standardmäßig ausgewählt.
- Durch Auswählen des Gruppensymbols über der Liste werden die verfügbaren Gruppen angezeigt. Durch Abwählen des Symbols werden die Gruppen ausgeblendet. Dieses Symbol ist standardmäßig ausgewählt.

Benutzer und Gruppen können mithilfe der Schaltfläche **Move** in die Liste der Projektbenutzer und -gruppen verschoben werden. Diese Benutzer und Mitglieder dieser Gruppen haben Lesezugriff (Analytic Server-Quellenknoten) und Schreibzugriff (Analytic Server-Exportknoten) auf dieses Projekt, wenn sie zu Analytic Server eine Verbindung über IBM SPSS Modeler herstellen.

Anmerkung: Änderungen, die an der Registerkarte **Owners** vorgenommen werden, werden sofort und automatisch übernommen.

Anmerkung: Administratoren verfügen über Lese- und Schreibzugriff auf jedes Projekt, unabhängig davon, ob sie namentlich als Eigner aufgelistet sind.

Versions

Projekte werden auf der Basis von Änderungen des Datei- und Ordnerinhalts versioniert. Für Änderungen an den Attributen eines Projekts (z. B. die Beschreibung, ob es öffentlich ist und mit wem es gemeinsam genutzt wird) ist keine neue Version erforderlich. Zum Hinzufügen, Ändern oder Löschen von Dateien oder Ordnern ist eine neue Version erforderlich.

Tabelle für die Projektversionssteuerung

In der Tabelle werden die vorhandenen Projektversionen, ihr Erstellungs- und Festschreibungsdatum, die Benutzer, die für die einzelnen Versionen verantwortlich sind, und die übergeordnete Version angezeigt. Die übergeordnete Version ist die Version, auf der die ausgewählte Version basiert.

- Klicken Sie auf **Lock**, um Änderungen am ausgewählten Projektversionsinhalt vorzunehmen.
- Klicken Sie auf **Commit**, um alle an einem Projekt vorgenommenen Änderungen zu speichern und diese Version zum aktuellen sichtbaren Status des Projekts zu machen.
- Klicken Sie auf **Discard**, um alle an einem gesperrten Projekt vorgenommenen Änderungen zu verwerfen und die zuletzt festgeschriebene Version wieder zum sichtbaren Status des Projekts zu machen.
- Klicken Sie auf **Delete Project**, um die ausgewählte Version zu entfernen.

Automatically clean up when number of versions exceeds

Löscht automatisch die älteste festgeschriebene Projektversion, wenn die Anzahl der Versionen die angegebene Anzahl überschreitet.

Bemerkungen

Die vorliegenden Informationen wurden für Produkte und Services entwickelt, die auf dem deutschen Markt angeboten werden.

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

Trotz sorgfältiger Bearbeitung können technische Ungenauigkeiten oder Druckfehler in dieser Veröffentlichung nicht ausgeschlossen werden. 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 Software Group
ATTN: Licensing
200 W. Madison St.
Chicago, IL; 60606
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.

Alle in diesem Dokument enthaltenen Leistungsdaten stammen aus einer kontrollierten Umgebung. Die Ergebnisse, die in anderen Betriebsumgebungen erzielt werden, können daher erheblich von den hier erzielten Ergebnissen abweichen. Einige Daten stammen möglicherweise von Systemen, deren Entwicklung noch nicht abgeschlossen ist. Eine Gewährleistung, dass diese Daten auch in allgemein verfügbaren Systemen erzielt werden, kann nicht gegeben werden. Darüber hinaus wurden einige Daten unter Umständen durch Extrapolation berechnet. Die tatsächlichen Ergebnisse können davon abweichen. Benutzer dieses Dokuments sollten die entsprechenden Daten in ihrer spezifischen Umgebung prüfen.

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.

Alle von IBM angegebenen Preise sind empfohlene Richtpreise und können jederzeit ohne weitere Mitteilung geändert werden. Händlerpreise können u. U. von den hier genannten Preisen abweichen.

Diese Veröffentlichung dient nur zu Planungszwecken. Die in dieser Veröffentlichung enthaltenen Informationen können geändert werden, bevor die beschriebenen Produkte verfügbar sind.

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.

Kopien oder Teile der Beispielprogramme bzw. daraus abgeleiteter Code müssen folgenden Copyrightvermerk beinhalten:

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.

Kopien oder Teile der Beispielprogramme bzw. daraus abgeleiteter Code müssen folgenden Copyrightvermerk beinhalten:

© (Name Ihrer Firma) (Jahr). Teile des vorliegenden Codes wurden aus Beispielprogrammen der IBM Corp. abgeleitet.

© Copyright IBM Corp. _Jahr/Jahre angeben_. Alle Rechte vorbehalten.

Marken

IBM, das IBM Logo und ibm.com sind Marken oder eingetragene Marken der IBM Corp 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.

IT Infrastructure Library ist eine eingetragene Marke der Central Computer and Telecommunications Agency. Die Central Computer and Telecommunications Agency ist nunmehr in das Office of Government Commerce eingegliedert worden.

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 der Microsoft Corporation in den USA und/oder anderen Ländern.

ITIL ist eine eingetragene Marke, eine eingetragene Gemeinschaftsmarke des Cabinet Office und eine eingetragene Marke, die beim US Patent and Trademark Office registriert ist.

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.

Cell Broadband Engine wird unter Lizenz verwendet und ist eine Marke der Sony Computer Entertainment, Inc. in den USA und/oder anderen Ländern.

Linear Tape-Open, LTO, das LTO-Logo, Ultrium und das Ultrium-Logo sind Marken von HP, IBM und Quantum in den USA und anderen Ländern.

Index

A

Analytic Server-Export 4
Analytic Server-Quelle 3

D

Datenquellen 9

E

Exportknoten
Analytic Server-Export 4

P

Projekte 15

Q

Quellenknoten
Analytic Server-Quelle 3

S

Streameigenschaften
Analytic Server 4

