

*IBM SPSS Modeler Server 18.1
Verwaltungs- und Leistungshandbuch*

IBM

Hinweis

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen“ auf Seite 97 gelesen werden.

Produktinformation

Diese Ausgabe bezieht sich auf Version 18, Release 0, Modifikation 0 von IBM SPSS Modeler und alle nachfolgenden Releases und Modifikationen, bis dieser Hinweis in einer Neuauflage geändert wird.

Diese Veröffentlichung ist eine Übersetzung des Handbuchs *IBM SPSS Modeler Server 18.1, Client Server Administration and Performance Guide*, herausgegeben von International Business Machines Corporation, USA

© Copyright International Business Machines Corporation 2017

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
Juni 2017

Inhaltsverzeichnis

Vorwort	v	Konfigurieren von Serverprofilen	29
Kapitel 1. Informationen zu IBM SPSS Modeler	1	Arbeiten mit Serverprofilen	29
IBM SPSS Modeler-Produkte	1	Profilstruktur	31
IBM SPSS Modeler	1	Profilscrippts	33
IBM SPSS Modeler Server	1	Verwaltung	38
IBM SPSS Modeler Administration Console	2	IBM SPSS Modeler Server-Verwaltung	38
IBM SPSS Modeler Batch	2	Starten von Modeler Administration Console	38
IBM SPSS Modeler Solution Publisher	2	Neustart des Web-Service.	38
IBM SPSS Modeler Server-Adapter für IBM SPSS Collaboration and Deployment Services	2	Konfigurieren des Zugriffs mithilfe von Modeler Administration Console	39
IBM SPSS Modeler-Editionen	2	SPSS Modeler Server-Verbindungen	39
IBM SPSS Modeler-Dokumentation	3	SPSS Modeler Server-Konfiguration	40
SPSS Modeler Professional-Dokumentation	3	SPSS Modeler Server-Überwachung	48
SPSS Modeler Premium-Dokumentation	4	Verwenden der Datei options.cfg	48
Anwendungsbeispiele	4	Schließen nicht verwendeter Datenbankverbindungen	49
Ordner "Demos"	4	Verwenden von SSL zur sicheren Datenübertragung	50
Lizenzüberwachung	5	Funktionsweise von SSL	50
Kapitel 2. Architektur- und Hardwareempfehlungen	7	Schützen der Client/Server- und Server/Server-Kommunikation durch SSL	50
IBM SPSS Modeler-Architektur	7	Cognos-SSL-Verbindung	54
Architekturbeschreibung	7	Cognos TM1-SSL-Verbindung	55
Hardwareempfehlungen	9	Konfigurieren von Gruppen	55
Anforderungen an temporären Festplattenspeicher und Arbeitsspeicher	10	Serverprotokoll	57
Datenzugriff	11	Kapitel 5. Leistungsüberblick	59
Referenzieren von Datendateien	12	Leistungs- und Optimierungseinstellungen für Ihren Server	59
Importieren von IBM SPSS Statistics-Datendateien	13	Leistungs- und Optimierungseinstellungen für Ihren Client	59
Installationsanweisungen	13	Datenbankverwendung und -optimierung	61
Kapitel 3. IBM SPSS Modeler-Support	15	SQL-Optimierung	61
Verbinden mit IBM SPSS Modeler Server	15	Kapitel 6. SQL-Optimierung	63
Konfigurieren von Single Sign-on	16	Funktionsweise der SQL-Generierung	64
Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung	22	Beispiel für SQL-Generierung	65
Suchen nach Servern in IBM SPSS Collaboration and Deployment Services.	22	Konfigurieren der SQL-Optimierung	66
Daten- und Dateisysteme	23	Vorschau für generiertes SQL	67
Benutzerauthentifizierung	23	Anzeigen von SQL für Modellnuggets	67
Berechtigungen	24	Tipps für die Maximierung der SQL-Generierung.	67
Dateierstellung	25	Knoten zur Unterstützung der SQL-Generierung	69
Unterschiede bei den Ergebnissen	25	CLEM-Ausdrücke und Operatoren, die die SQL-Generierung unterstützen.	73
Kapitel 4. IBM SPSS Modeler-Verwaltung	27	Verwenden von SQL-Funktionen in CLEM-Ausdrücken	75
Starten und Stoppen von IBM SPSS Modeler Server	27	Schreiben von SQL-Abfragen	75
Starten, Stoppen und Statusprüfung unter Windows	27	Scoring-Adapter für Teradata - doppelte Zeilen	76
Starten, Stoppen und Statusprüfung unter UNIX	27	Anhang A. Konfigurieren von Oracle für UNIX-Plattformen	77
Umgang mit nicht reagierenden Serverprozessen (UNIX-Systeme)	28	Konfigurieren von Oracle für die SQL-Optimierung	77

Anhang B. Konfigurieren von UNIX-Startscripts. 79

Einführung 79
Scripts 79
Automatisches Starten und Stoppen von IBM SPSS Modeler Server 79
Manuelles Starten und Stoppen von IBM SPSS Modeler Server 80
Bearbeiten von Scripts. 80
Steuern der Berechtigungen für die Dateierstellung IBM SPSS Modeler Server und das Data Access Pack 80
 Fehlerbehebung bei der ODBC-Konfiguration . . 83
 Bibliothekspfade. 86

Anhang C. Konfigurieren und Ausführen von SPSS Modeler Server als Nicht-Root-Prozess unter UNIX 87

Einführung 87
Konfigurieren als Nicht-Root-Benutzer ohne private Kennwortdatenbank 87
Konfigurieren als Nicht-Root-Benutzer mit privater Kennwortdatenbank 88

Ausführen von SPSS Modeler Server als Nicht-Root-Benutzer 89
Fehlerbehebung bei fehlgeschlagener Benutzerauthentifizierung 90

Anhang D. Konfigurieren und Ausführen von SPSS Modeler Server mit privater Kennwortdatei unter Windows . . 91

Einführung 91
Konfigurieren einer privaten Kennwortdatenbank 91

Anhang E. Lastverteilung mit Server-Clustern 93

Anhang F. LDAP-Authentifizierung . . . 95

Bemerkungen. 97

Marken. 98
Bedingungen für Produktdokumentation 98

Index 101

Vorwort

IBM® SPSS Modeler ist die auf Unternehmensebene einsetzbare Data-Mining-Workbench von IBM. Mit SPSS Modeler können Unternehmen und Organisationen die Beziehungen zu ihren Kunden bzw. zu den Bürgern durch ein tief greifendes Verständnis der Daten verbessern. Organisationen verwenden die mithilfe von SPSS Modeler gewonnenen Erkenntnisse zur Bindung profitabler Kunden, zur Ermittlung von Cross-Selling-Möglichkeiten, zur Gewinnung neuer Kunden, zur Ermittlung von Betrugsfällen, zur Reduzierung von Risiken und zur Verbesserung der Verfügbarkeit öffentlicher Dienstleistungen.

Die grafische Schnittstelle von SPSS Modeler erleichtert die Anwendung des jeweiligen Fachwissens der Benutzer, was zu leistungsstärkeren Vorhersagemodellen und einer schnelleren Lösungserstellung führt. SPSS Modeler bietet zahlreiche Modellierungsverfahren, beispielsweise Algorithmen für Vorhersage, Klassifizierung, Segmentierung und Assoziationserkennung. Nach der Modellerstellung ermöglicht IBM SPSS Modeler Solution Publisher die unternehmensweite Bereitstellung des Modells für Entscheidungsträger oder in einer Datenbank.

Informationen zu IBM Business Analytics

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

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

Technical Support

Kunden mit Wartungsvertrag können den Technical Support in Anspruch nehmen. Kunden können sich an den Technical Support wenden, wenn sie Hilfe bei der Arbeit mit IBM Produkten oder bei der Installation in einer der unterstützten Hardwareumgebungen benötigen. Den Technical Support erreichen Sie über die Website von IBM unter <http://www.ibm.com/support>. Sie müssen bei der Kontaktaufnahme Ihren Namen, Ihr Unternehmen und Ihre Supportvereinbarung angeben.

Kapitel 1. Informationen zu IBM SPSS Modeler

IBM SPSS Modeler ist ein Set von Data-Mining-Tools, mit dem Sie auf der Grundlage Ihres Fachwissens schnell und einfach Vorhersagemodelle erstellen und zur Erleichterung der Entscheidungsfindung in die Betriebsabläufe einbinden können. Das Produkt IBM SPSS Modeler, das auf der Grundlage des den Industrienormen entsprechenden Modells CRISP-DM entwickelt wurde, unterstützt den gesamten Data-Mining-Prozess, von den Daten bis hin zu besseren Geschäftsergebnissen.

IBM SPSS Modeler bietet eine Vielzahl von Modellbildungsmethoden, die aus dem maschinellen Lernen, der künstlichen Intelligenz und der Statistik stammen. Mit den in der Modellierungspalette verfügbaren Methoden können Sie aus Ihren Daten neue Informationen ableiten und Vorhersagemodelle erstellen. Jede Methode hat ihre speziellen Stärken und eignet sich besonders für bestimmte Problemtypen.

SPSS Modeler kann als Standalone-Produkt oder als Client in Verbindung mit SPSS Modeler Server erworben werden. Außerdem ist eine Reihe von Zusatzoptionen verfügbar, die in den folgenden Abschnitten kurz zusammengefasst werden. Weitere Informationen finden Sie unter <https://www.ibm.com/analytics/us/en/technology/spss/>.

IBM SPSS Modeler-Produkte

Zur IBM SPSS Modeler-Produktfamilie und der zugehörigen Software gehören folgende Elemente.

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten)
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- IBM SPSS Modeler Server-Adapter für IBM SPSS Collaboration and Deployment Services

IBM SPSS Modeler

SPSS Modeler ist eine funktionell in sich abgeschlossene Produktversion, die Sie auf Ihrem PC installieren und ausführen können. Sie können SPSS Modeler im lokalen Modus als Standalone-Produkt oder im verteilten Modus zusammen mit IBM SPSS Modeler Server verwenden, um bei Datasets die Leistung zu verbessern.

Mit SPSS Modeler können Sie schnell und intuitiv genaue Vorhersagemodelle erstellen, und das ohne Programmierung. Mithilfe der speziellen grafischen Schnittstelle können Sie den Data-Mining-Prozess auf einfache Weise visualisieren. Mit der Unterstützung der in das Produkt eingebetteten erweiterten Analyseprozesse können Sie zuvor verborgene Muster und Trends in Ihren Daten aufdecken. Sie können Ergebnisse modellieren und Einblick in die Faktoren gewinnen, die Einfluss auf diese Ergebnisse haben, wodurch Sie in die Lage versetzt werden, Geschäftschancen zu nutzen und Risiken zu mindern.

SPSS Modeler ist in zwei Editionen erhältlich: SPSS Modeler Professional und SPSS Modeler Premium. Weitere Informationen finden Sie in „IBM SPSS Modeler-Editionen“ auf Seite 2.

IBM SPSS Modeler Server

SPSS Modeler verwendet eine Client/Server-Architektur zur Verteilung von Anforderungen für ressourcenintensive Vorgänge an leistungsstarke Serversoftware, wodurch bei größeren Datasets eine höhere Leistung erzielt werden kann.

SPSS Modeler Server ist ein separat lizenziertes Produkt, das durchgehend im Modus für verteilte Analysen auf einem Server-Host in Verbindung mit einer oder mehreren IBM SPSS Modeler-Installationen ausgeführt wird. Auf diese Weise bietet SPSS Modeler Server eine herausragende Leistung bei großen Datensätzen, da speicherintensive Vorgänge auf dem Server ausgeführt werden können, ohne Daten auf den Client-Computer herunterladen zu müssen. IBM SPSS Modeler Server bietet außerdem Unterstützung für SQL-Optimierung sowie Möglichkeiten zur Modellierung innerhalb der Datenbank, was weitere Vorteile hinsichtlich Leistung und Automatisierung mit sich bringt.

IBM SPSS Modeler Administration Console

Modeler Administration Console ist eine grafische Benutzerschnittstelle zur Verwaltung einer Vielzahl von SPSS Modeler Server-Konfigurationsoptionen, die auch mithilfe einer Optionsdatei konfiguriert werden können. Die Konsole gehört zum Lieferumfang von IBM SPSS Deployment Manager, kann zum Überwachen und Konfigurieren Ihrer SPSS Modeler Server-Installationen verwendet werden und stehen aktuellen SPSS Modeler Server-Kunden kostenlos zur Verfügung. Die Anwendung kann nur unter Windows installiert werden. Der von ihr verwaltete Server kann jedoch auf einer beliebigen unterstützten Plattform installiert sein.

IBM SPSS Modeler Batch

Das Data-Mining ist zwar in der Regel ein interaktiver Vorgang, es ist jedoch auch möglich, SPSS Modeler über eine Befehlszeile auszuführen, ohne dass die grafische Benutzerschnittstelle verwendet werden muss. Beispielsweise kann es sinnvoll sein, langwierige oder sich wiederholende Aufgaben ohne Eingreifen des Benutzers durchzuführen. SPSS Modeler Batch ist eine spezielle Version des Produkts, die die vollständigen Analysefunktionen von SPSS Modeler ohne Zugriff auf die reguläre Benutzerschnittstelle bietet. SPSS Modeler Server ist für die Verwendung von SPSS Modeler Batch erforderlich.

IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher ist ein Tool, mit dem Sie eine gepackte Version eines SPSS Modeler-Streams erstellen können, der durch eine externe Runtime-Engine ausgeführt oder in eine externe Anwendung eingebettet werden kann. Auf diese Weise können Sie vollständige SPSS Modeler-Streams für die Verwendung in Umgebungen veröffentlichen und bereitstellen, in denen SPSS Modeler nicht installiert ist. SPSS Modeler Solution Publisher wird als Teil des Diensts für IBM SPSS Collaboration and Deployment Services - Scoring verteilt, für den eine separate Lizenz erforderlich ist. Mit dieser Lizenz erhalten Sie SPSS Modeler Solution Publisher Runtime, womit Sie die veröffentlichten Streams ausführen können.

Weitere Informationen zu SPSS Modeler Solution Publisher finden Sie in der Dokumentation zu IBM SPSS Collaboration and Deployment Services. Das IBM SPSS Collaboration and Deployment Services Knowledge Center enthält die Abschnitte "IBM SPSS Modeler Solution Publisher" und "IBM SPSS Analytics Toolkit".

IBM SPSS Modeler Server-Adapter für IBM SPSS Collaboration and Deployment Services

Für IBM SPSS Collaboration and Deployment Services ist eine Reihe von Adaptern verfügbar, mit denen SPSS Modeler und SPSS Modeler Server mit einem Repository von IBM SPSS Collaboration and Deployment Services interagieren können. Auf diese Weise kann ein im Repository bereitgestellter SPSS Modeler-Stream von mehreren Benutzern gemeinsam verwendet werden. Auch der Zugriff über die Thin-Client-Anwendung IBM SPSS Modeler Advantage ist möglich. Sie installieren den Adapter auf dem System, das als Host für das Repository fungiert.

IBM SPSS Modeler-Editionen

SPSS Modeler ist in den folgenden Editionen erhältlich.

SPSS Modeler Professional

SPSS Modeler Professional bietet sämtliche Tools, die Sie für die Arbeit mit den meisten Typen von strukturierten Daten benötigen, beispielsweise in CRM-Systemen erfasste Verhaltensweisen und Interaktionen, demografische Daten, Kaufverhalten und Umsatzdaten.

SPSS Modeler Premium

SPSS Modeler Premium ist ein separat lizenziertes Produkt, das SPSS Modeler Professional für die Arbeit mit spezialisierten Daten sowie für die Arbeit mit unstrukturierten Textdaten erweitert. SPSS Modeler Premium schließt IBM SPSS Modeler Text Analytics ein:

IBM SPSS Modeler Text Analytics verwendet hoch entwickelte linguistische Technologien und die Verarbeitung natürlicher Sprache (Natural Language Processing, NLP), um eine schnelle Verarbeitung einer großen Vielfalt an unstrukturierten Textdaten zu ermöglichen, um die Schlüsselkonzepte zu extrahieren und zu ordnen und um diese Konzepte in Kategorien zusammenzufassen. Extrahierte Konzepte und Kategorien können mit vorhandenen strukturierten Daten, beispielsweise demografischen Informationen, kombiniert und mithilfe der vollständigen Suite der Data-Mining-Tools von IBM SPSS Modeler auf die Modellierung angewendet werden, um bessere und fokussiertere Entscheidungen zu ermöglichen.

IBM SPSS Modeler-Dokumentation

Dokumentation ist über das Hilfemenü in SPSS Modeler verfügbar. Dadurch wird das SPSS Modeler Knowledge Center geöffnet, das außerhalb des Produkts öffentlich verfügbar ist.

Die vollständige Dokumentation für die einzelnen Produkte (einschließlich Installationsanweisungen) ist über den Produktdownload in einem separaten komprimierten Ordner auch im PDF-Format verfügbar. Die PDF-Dokumente können auch über das Web unter <http://www.ibm.com/support/docview.wss?uid=swg27046871> heruntergeladen werden.

SPSS Modeler Professional-Dokumentation

Die SPSS Modeler Professional-Dokumentationssuite (ohne Installationsanweisungen) umfasst folgende Dokumente:

- **IBM SPSS Modeler Benutzerhandbuch.** Allgemeine Einführung in die Verwendung von SPSS Modeler, in der u. a. die Erstellung von Datenstreams, der Umgang mit fehlenden Werten, die Erstellung von CLEM-Ausdrücken, die Arbeit mit Projekten und Berichten sowie das Packen von Streams für die Bereitstellung in IBM SPSS Collaboration and Deployment Services oder IBM SPSS Modeler Advantage beschrieben werden.
- **IBM SPSS Modeler Quellen-, Prozess- und Ausgabeknoten.** Beschreibung aller Knoten, die zum Lesen, zum Verarbeiten und zur Ausgabe von Daten in verschiedenen Formaten verwendet werden. Im Grunde sind sie alle Knoten, mit Ausnahme der Modellierungsknoten.
- **IBM SPSS Modeler Modellierungsknoten.** Beschreibungen sämtlicher für die Erstellung von Data-Mining-Modellen verwendeter Knoten. IBM SPSS Modeler bietet eine Vielzahl von Modellbildungsmethoden, die aus dem maschinellen Lernen, der künstlichen Intelligenz und der Statistik stammen.
- **IBM SPSS Modeler Anwendungshandbuch.** Die Beispiele in diesem Handbuch bieten eine kurze, gezielte Einführung in bestimmte Modellierungsmethoden und -verfahren. Eine Online-Version dieses Handbuchs kann auch über das Hilfemenü aufgerufen werden. Weitere Informationen finden Sie in „Anwendungsbeispiele“ auf Seite 4.
- **IBM SPSS Modeler Python Handbuch für Scripterstellung und Automatisierung.** Informationen zur Automatisierung des Systems über Python-Scripterstellung, einschließlich der Eigenschaften, die zur Bearbeitung von Knoten und Streams verwendet werden können.
- **IBM SPSS Modeler Bereitstellungshandbuch.** Informationen zum Ausführen von IBM SPSS Modeler-Streams als Schritte bei der Verarbeitung von Jobs im IBM SPSS Deployment Manager.

- **IBM SPSS Modeler CLEF-Entwicklerhandbuch.** CLEF bietet die Möglichkeit, Drittanbieterprogramme, wie Datenverarbeitungsroutinen oder Modellierungsalgorithmen, als Knoten in IBM SPSS Modeler zu integrieren.
- **IBM SPSS Modeler Datenbankinternes Mining.** Informationen darüber, wie Sie Ihre Datenbank dazu einsetzen, die Leistung zu verbessern, und wie Sie die Palette der Analysefunktionen über Drittanbieteralgorithmen erweitern.
- **IBM SPSS Modeler Server Verwaltungs- und Leistungshandbuch.** Informationen zur Konfiguration und Verwaltung von IBM SPSS Modeler Server.
- **IBM SPSS Deployment Manager Benutzerhandbuch.** Informationen zur Verwendung der zum Lieferumfang von Deployment Manager gehörenden Benutzerschnittstelle der Administrationskonsole zum Überwachen und Konfigurieren von IBM SPSS Modeler Server.
- **IBM SPSS Modeler CRISP-DM Handbuch.** Schritt-für-Schritt-Anleitung für das Data Mining mit SPSS Modeler unter Verwendung der CRISP-DM-Methode.
- **IBM SPSS Modeler Batch Benutzerhandbuch.** Vollständiges Handbuch für die Verwendung von IBM SPSS Modeler im Stapelmodus, einschließlich Details zur Ausführung des Stapelmodus und zu Befehlszeilenargumenten. Dieses Handbuch steht nur im PDF-Format zur Verfügung.

SPSS Modeler Premium-Dokumentation

Die SPSS Modeler Premium-Dokumentationssuite (ohne Installationsanweisungen) umfasst folgende Dokumente:

- **SPSS Modeler Text Analytics Benutzerhandbuch.** Informationen zur Verwendung von Textanalysen mit SPSS Modeler, unter Behandlung der Text Mining-Knoten, der interaktiven Workbench sowie von Vorlagen und anderen Ressourcen.

Anwendungsbeispiele

Mit den Data-Mining-Tools in SPSS Modeler kann eine große Bandbreite an geschäfts- und unternehmensbezogenen Problemen gelöst werden; die Anwendungsbeispiele dagegen bieten jeweils eine kurze, gezielte Einführung in spezielle Modellierungsmethoden und -verfahren. Die hier verwendeten Datasets sind viel kleiner als die großen Datenbestände, die von einigen Data-Mining-Experten verwaltet werden müssen, die zugrunde liegenden Konzepte und Methoden können jedoch auch auf reale Anwendungen übertragen werden.

Klicken Sie im Menü "Hilfe" in SPSS Modeler auf die Option **Anwendungsbeispiele**, um auf die Beispiele zuzugreifen.

Die Datendateien und Beispielstreams wurden im Ordner *Demos*, einem Unterordner des Produktinstallationsverzeichnis, installiert. Weitere Informationen finden Sie in „Ordner "Demos"“.

Beispiele für die Datenbankmodellierung. Die Beispiele finden Sie im Handbuch *IBM SPSS Modeler Datenbankinternes Mining*.

Scriptbeispiele. Die Beispiele finden Sie im *IBM SPSS Modeler Handbuch für Scripterstellung und Automatisierung*.

Ordner "Demos"

Die in den Anwendungsbeispielen verwendeten Datendateien und Beispielstreams wurden im Ordner *Demos*, einem Unterordner des Produktinstallationsverzeichnis (beispielsweise `C:\Programme\IBM\SPSS\Modeler\<Version>\Demos`), installiert. Auf diesen Ordner können Sie auch über die Programmgruppe **SPSS Modeler** im Windows-Startmenü oder durch Klicken auf *Demos* in der Liste der zuletzt angezeigten Verzeichnisse im Dialogfeld **Datei > Stream öffnen** zugreifen.

Lizenzüberwachung

Bei der Verwendung von SPSS Modeler wird die Lizenznutzung überwacht und in regelmäßigen Intervallen protokolliert. Es werden die Lizenzmetriken *AUTHORIZED_USER* und *CONCURRENT_USER* protokolliert und der Typ der protokollierten Metrik ist von Ihrem Lizenztyp für SPSS Modeler abhängig.

Die erstellten Protokolldateien können vom Produkt IBM License Metric Tool verarbeitet werden, über das Sie Lizenznutzungsberichte generieren können.

Die Lizenzprotokolldateien werden in demselben Verzeichnis erstellt, in dem SPSS Modeler Client-Protokolldateien aufgezeichnet werden (standardmäßig %ALLUSERSPROFILE%/IBM/SPSS/Modeler/<Version>/log).

Kapitel 2. Architektur- und Hardwareempfehlungen

IBM SPSS Modeler-Architektur

In diesem Kapitel wird die Architektur von IBM SPSS Modeler Server beschrieben, inklusive der Server-Software, der Client-Software und der Datenbank. Es umfasst Informationen zum Design von IBM SPSS Modeler Server, das auf optimale Leistung ausgerichtet ist, sowie Empfehlungen zur erforderlichen Hardware, um diese Leistung voll ausschöpfen zu können. Das Kapitel wird mit einem Abschnitt über den Datenzugriff abgeschlossen. Hier erfahren Sie, wo die erforderlichen ODBC-Treiber eingerichtet werden müssen.

Architekturbeschreibung

IBM SPSS Modeler Server verwendet eine verteilte Architektur mit drei Schichten. Softwareoperationen werden zwischen den Client- und Server-Computern aufgeteilt. Die Vorteile von IBM SPSS Modeler Server gegenüber der Standalone-Version von IBM SPSS Modeler, vor allem beim Verarbeiten großer Datensets, sind zahlreich:

- IBM SPSS Modeler Server kann nicht nur unter Windows, sondern auch unter UNIX ausgeführt werden, sodass Sie eine größere Flexibilität hinsichtlich des Installationsortes haben. Sie können auf jeder Plattform einen schnelleren, größeren Server-Computer für Data-Mining-Prozesse verwenden.
- IBM SPSS Modeler Server wurde vorwiegend für Schnelligkeit optimiert. Wenn Vorgänge nicht an die Datenbank übertragen werden können, speichert IBM SPSS Modeler Server die Zwischenergebnisse als temporäre Dateien auf der Festplatte und nicht im Arbeitsspeicher. Da Server normalerweise über ausreichend Festplattenspeicher verfügen, kann IBM SPSS Modeler Server Sortier-, Zusammenführungs- und Aggregationsvorgänge mit sehr großen Datensets durchführen.
- Mit der Client/Server-Architektur können Sie die Data-Mining-Prozesse in Ihrem Unternehmen zentralisieren. Auf diese Weise können Sie die Rolle des Data-Mining in Ihren Geschäftsprozessen formalisieren.
- Mithilfe von Verwaltungstools wie IBM SPSS Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten) sowie IBM SPSS Collaboration and Deployment Services (separat erhältlich) können Sie Data-Mining-Prozesse überwachen und sicherstellen, dass ausreichend Computerressourcen verfügbar sind. Mit IBM SPSS Collaboration and Deployment Services können Sie bestimmte Data-Mining-Aufgaben automatisieren, den Zugriff auf Datenmodelle verwalten und Ergebnisse unternehmensweit verfügbar machen.

Die Komponenten der verteilten Architektur von IBM SPSS Modeler sind in der Grafik "IBM SPSS Modeler Server-Architektur" dargestellt.

- **IBM SPSS Modeler.** Die Client-Software wird auf dem Computer des Endbenutzers installiert. Sie stellt die Benutzerschnittstelle zur Verfügung und zeigt Data-Mining-Ergebnisse an. Der Client ist eine vollständige Installation der IBM SPSS Modeler-Software. Wird der Client jedoch für die verteilte Analyse mit IBM SPSS Modeler Server verbunden, wird seine Ausführungseingabe inaktiviert. IBM SPSS Modeler kann nur unter Windows ausgeführt werden.
- **IBM SPSS Modeler Server.** Die Server-Software wird auf einem Server-Computer installiert, der Netzkonnektivität sowohl zu IBM SPSS Modeler-Instanzen als auch zur Datenbank aufweist. IBM SPSS Modeler Server wird als Service bzw. Dienst (unter Windows) oder als Dämonprozess (unter UNIX) ausgeführt und wartet auf die Verbindung von Clients. Er verarbeitet die Ausführung von Streams und Scripts, die mit IBM SPSS Modeler erstellt wurden.
- **Datenbankserver.** Der Datenbankserver könnte ein Livedaten-Warehouse (z. B. Oracle auf einem großen UNIX-Server) oder, um die Auswirkungen auf andere Betriebssysteme zu verringern, ein Datamart auf einem lokalen Server bzw. einem Abteilungsserver (z. B. SQL Server unter Windows) sein.

IBM SPSS Modeler Server-Architektur

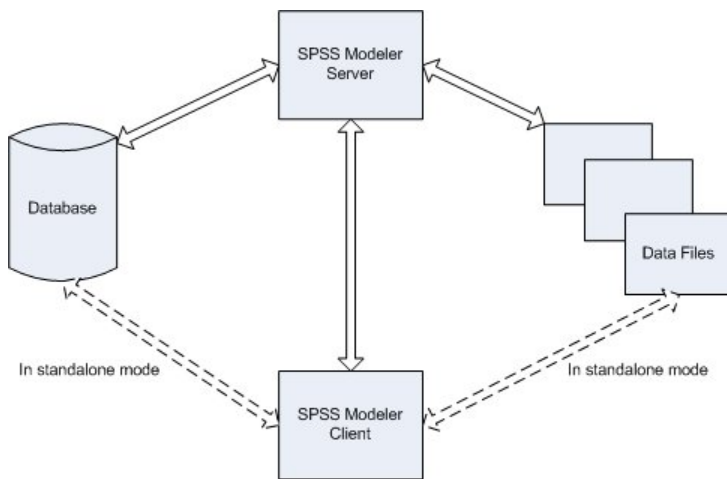


Abbildung 1. IBM SPSS Modeler-Architektur

Bei der verteilten Architektur findet ein Großteil der Verarbeitung auf dem Server-Computer statt. Wenn der Endbenutzer einen Stream ausführt, sendet IBM SPSS Modeler eine Beschreibung des Streams an den Server. Der Server ermittelt, welche Vorgänge in SQL ausgeführt werden können, und erstellt die passenden Anfragen. Diese Anfragen werden in der Datenbank ausgeführt und die Ergebnisse werden an den Server übergeben, der die Verarbeitung der Daten übernimmt, die nicht mithilfe von SQL ausgedrückt werden können. Nachdem die Verarbeitung abgeschlossen ist, werden nur relevante Ergebnisse an den Client zurückgegeben.

Falls erforderlich, kann IBM SPSS Modeler Server alle IBM SPSS Modeler-Vorgänge außerhalb der Datenbank durchführen. Die Verwendung von Arbeitsspeicher und Festplattenspeicher für die Speicherung von Daten für die Verarbeitung wird automatisch im Gleichgewicht gehalten. Durch diesen Prozess ist IBM SPSS Modeler Server vollständig kompatibel mit Flatfiles.

Außerdem kann die Last auch durch die Verwendung von Server-Clustern für die Verarbeitung verteilt werden. Clustering ist in IBM SPSS Collaboration and Deployment Services ab Version 3.5 über das Plugin Coordinator of Processes verfügbar. Weitere Informationen finden Sie in Anhang E, „Lastverteilung mit Server-Clustern“, auf Seite 93. Die Verbindung mit einem im Coordinator of Processes verwalteten Server oder Cluster ist direkt über das IBM SPSS Modeler-Dialogfeld "Anmelden beim Server" möglich. Weitere Informationen finden Sie in „Verbinden mit IBM SPSS Modeler Server“ auf Seite 15.

Standalone-Client

IBM SPSS Modeler kann auch für die Ausführung als eigenständige Desktopanwendung konfiguriert werden, wie in der unten stehenden Grafik gezeigt. Weitere Informationen finden Sie in Kapitel 3, „IBM SPSS Modeler-Support“, auf Seite 15.



Abbildung 2. IBM SPSS Modeler (Standalone)

Hardwareempfehlungen

Bei der Planung Ihrer IBM SPSS Modeler Server-Installation sollten Sie auch die Hardware berücksichtigen, die Sie verwenden werden. Auch wenn das Design von IBM SPSS Modeler Server auf Geschwindigkeit ausgerichtet ist, können Sie die Effizienz der Anwendung nur mithilfe der richtigen Hardware für Ihre Data-Mining-Aufgaben vollständig ausschöpfen. Ein Aufrüsten der Hardware ist häufig die einfachste und wirtschaftlichste Lösung zum Verbessern der Systemleistung.

Dedizierter Server. Installieren Sie IBM SPSS Modeler Server auf einem dedizierten Server-Computer, dessen Ressourcen allein der Server-Software zur Verfügung stehen. Dies betrifft auch alle Datenbanken, auf die IBM SPSS Modeler Server zugreift. Vor allem Modellerstellungsvorgänge sind besonders ressourcenintensiv und können wesentlich schneller durchgeführt werden, wenn keine Ressourcen für andere Anwendungen abgezweigt werden.

Anmerkung: Obwohl die Installation von IBM SPSS Modeler Server und der Datenbank auf demselben Computer die Übertragungszeit zwischen diesen beiden Komponenten aufgrund der verringerten Netzlast reduzieren kann, ist es in den meisten Fällen sinnvoll, Server und Datenbank auf getrennten Computern zu konfigurieren, um Ressourcenkonflikte zu vermeiden. Mit einer schnellen Verbindung zwischen den beiden Komponenten können Sie die Kosten der Datenübertragung minimieren.

Prozessoren. Die Anzahl an Prozessoren im Computer sollte nicht unter der Anzahl an gleichzeitigen Aufgaben (gleichzeitigen Streams) liegen, die regelmäßig ausgeführt werden müssen. Im Allgemeinen gilt: Je mehr Prozessoren, desto besser.

- Eine einzelne Instanz von IBM SPSS Modeler Server akzeptiert Verbindungen von mehreren Clients (Benutzern). Jede Clientverbindung kann mehrere Streamausführungen starten. Auf einem Server können daher stets mehrere Ausführungsaufgaben gleichzeitig laufen.
- Als Faustregel gehen Sie von einem Prozessor für einen bis zwei Benutzer, zwei Prozessoren für bis zu vier Benutzer und acht Prozessoren für bis zu acht Benutzer aus. Je nach Arbeitsanforderung fügen Sie darüber hinaus je einen weiteren Prozessor für alle zwei bis vier Benutzer hinzu.
- Der Anteil der Verarbeitung, der durch die SQL-Optimierung an die Datenbank übergeben werden kann, bestimmt, ob eine CPU von zwei oder mehr Benutzern gemeinsam verwendet werden kann, bei einem nur geringen Leistungsverlust.
- Multithreading-Fähigkeiten ermöglichen, dass eine Aufgabe mehrere Prozessoren verwenden kann, sodass ein Aufstocken der Prozessoren die Leistung auch in Fällen verbessert, in denen jeweils nur eine Aufgabe gleichzeitig durchgeführt wird. Im Allgemeinen wird das Multithreading für die C5.0-Modellerstellung und für bestimmte Datenvorbereitungsvorgänge (Sortieren, Aggregieren und Zusammenführen) verwendet. Multithreading wird auch für alle Knoten unterstützt, die in IBM SPSS Analytic Server ausgeführt werden (z. B. GLE, Linear-AS, Random Forest, LSVM, Tree-AS, Zeitreihen, TCM, Assoziationsregeln und STP).

64-Bit-Plattformen. Wenn Sie Modelle mit großen Datenmengen erstellen oder verarbeiten möchten, verwenden Sie einen 64-Bit-Computer als Plattform für IBM SPSS Modeler Server und verwenden Sie die maximal mögliche Menge an RAM für diesen Computer. Bei großen Datasets schöpft der Server die verfügbaren Speichermengen (die den Einschränkungen der 32-Bit-Plattform unterliegen) für Prozesse schnell aus, sodass Daten auf die Festplatte geschrieben werden müssen, was zu beträchtlichen Zeitverlusten bei der Ausführung führt. 64-Bit-Serverimplementierungen können zusätzlichen Arbeitsspeicher nutzen. Es werden jedoch mindestens 8 GB empfohlen.

Langfristige Anforderungen. Sofern möglich, stellen Sie sicher, dass die Serverhardware hinsichtlich Arbeitsspeicher und CPUs erweiterbar ist, um später einer höheren Auslastung (z. B. einer größeren Anzahl an gleichzeitigen Benutzern oder höheren Anforderungen an die Verarbeitung) sowie erweiterten Multithreading-Fähigkeiten von nachfolgenden IBM SPSS Modeler Server-Versionen gerecht werden zu können.

Anforderungen an temporären Festplattenspeicher und Arbeitsspeicher

IBM SPSS Modeler Server verwendet für die Verarbeitung von großen Datenmengen temporären Festplattenspeicher. Die Menge an erforderlichem Festplattenspeicher hängt von der verarbeiteten Datenart und -menge sowie von den jeweils durchgeführten Vorgängen ab. Die Datenmenge verhält sich proportional zur Anzahl an Zeilen *und* Spalten. Je mehr Zeilen und Spalten verarbeitet werden, desto mehr Festplattenspeicher ist erforderlich.

In diesem Abschnitt werden die Umstände beschrieben, unter denen temporärer Festplattenspeicher und zusätzlicher Arbeitsspeicher erforderlich sind und wie die erforderliche Menge ermittelt wird. Beachten Sie, dass in diesem Abschnitt nicht die Anforderungen an temporären Festplattenspeicher beschrieben werden, die in einer Datenbank auftreten, denn diese Anforderungen hängen von der jeweiligen Datenbank ab.

Bedingungen, die temporären Festplattenspeicher erfordern

Die leistungsfähige SQL-Optimierungsfunktion von IBM SPSS Modeler Server ermöglicht in vielen Fällen die Verarbeitung innerhalb der Datenbank (statt auf dem Server). Unter den folgenden Umständen kann die SQL-Optimierung jedoch nicht verwendet werden:

- Die zu verarbeitenden Daten befinden sich in einer Flatfile und nicht in einer Datenbank.
- Die SQL-Optimierung ist inaktiviert.
- Die Verarbeitung kann mit SQL nicht optimiert werden.

Wenn die SQL-Optimierung nicht verwendet werden kann, erstellen die folgenden Datenbearbeitungsknoten und CLEM-Funktionen auf der Festplatte temporäre Kopien von einigen oder allen Daten. Wenn die auf Ihrer Site verwendeten Streams diese Verarbeitungsbefehle oder -funktionen enthalten, müssen Sie möglicherweise zusätzlichen Festplattenspeicher auf Ihrem Server bereitstellen.

- Aggregatknotten
- Duplikatknotten
- Klassierknotten
- Zusammenführungsknotten bei Verwendung der Option "Zusammenführen nach Schlüssel"
- Jeder Modellierungsknotten
- Sortierknotten
- Tabellenausgabeknotten
- @OFFSET-Funktionen, bei denen die Suchbedingung die Funktion @THIS verwendet.
- Jede @-Funktion wie @MIN, @MAX und @AVE, in denen der Offset-Parameter berechnet wird.

Berechnen der Menge an erforderlichem Festplattenspeicher

Im Allgemeinen muss IBM SPSS Modeler Server in der Lage sein, eine temporäre Datei zu erstellen, die mindestens *dreimal so groß* ist, wie das ursprüngliche Dataset. Wenn die Datendatei beispielsweise 2 GB groß ist und die SQL-Generierung nicht verwendet wird, benötigt IBM SPSS Modeler Server 6 GB Festplattenspeicher, um die Daten zu verarbeiten. Da jedes gleichzeitig angemeldete Benutzerkonto eigene temporäre Dateien erstellt, müssen Sie den Festplattenspeicher für jeden dieser Benutzer entsprechend anpassen.

Wenn Sie feststellen, dass Ihre Site häufig große temporäre Dateien verwendet, können Sie für die temporären Dateien von IBM SPSS Modeler ein separates Dateisystem auf einer separaten Festplatte einrichten. Die besten Ergebnisse, um die Festplattenvorgänge zu beschleunigen, erzielen Sie mit einem RAID 0 oder Striped Data Set, das mehrere physische Festplatten umfasst, idealerweise mit jeder Festplatte im Striped-Dateisystem auf einem separaten Festplattencontroller.

Arbeitsspeicheranforderungen

Für die meisten Vorgänge, die nicht in der Datenbank verarbeitet werden können, speichert IBM SPSS Modeler Server die Zwischenergebnisse als temporäre Dateien auf der Festplatte und nicht im Arbeitsspeicher (RAM). Im Falle von Modellierungsknoten wird jedoch der Arbeitsspeicher verwendet, wenn möglich. Die Knoten "Neuronales Netz", "Kohonen" und "K-Means" benötigen eine große Menge an Arbeitsspeicher. Wenn diese Knoten häufig auf Ihrer Site verwendet werden, rüsten Sie gegebenenfalls den Arbeitsspeicher Ihres Servers auf.

Im Allgemeinen kann die Menge an erforderlichem Arbeitsspeicher wie folgt geschätzt werden:
(Anzahl_an_Datensätzen * Anzahl_an_Zellen_pro_Datensatz) * Anzahl_an_Byte_pro_Zelle.

Hierbei kann Anzahl_an_Zellen_pro_Datensatz sehr groß werden, wenn nominale Felder verwendet werden.

Die aktuell gültigen RAM-Anforderungen finden Sie im Serverinstallationshandbuch im Abschnitt zu den Systemanforderungen. Ab vier gleichzeitigen Benutzern wird noch mehr RAM empfohlen. Der Arbeitsspeicher muss unter gleichzeitig durchgeführten Aufgaben aufgeteilt werden, daher sollten Sie entsprechend aufrüsten. Grundsätzlich ist das Aufrüsten von Arbeitsspeicher einer der kostengünstigsten Wege, die Systemleistung zu verbessern.

Datenzugriff

Wenn Sie aus einer Datenbank lesen oder in eine Datenbank schreiben wollen, müssen Sie für die relevante Datenbank eine ODBC-Datenquelle mit Lese- bzw. Schreibberechtigung installiert und konfiguriert haben. Das IBM SPSS Data Access Pack umfasst eine Reihe von ODBC-Treibern, die zu diesem Zweck verwendet werden können. Diese Treiber stehen auf der Download-Site zur Verfügung. Wenn Sie Fragen zur Erstellung oder Einstellung von Berechtigungen für ODBC-Datenquellen haben, wenden Sie sich an Ihren Datenbankadministrator.

Unterstützte ODBC-Treiber

Die neuesten Informationen zu Datenbanken und ODBC-Treibern, die für die Verwendung mit IBM SPSS Modeler getestet wurden und unterstützt werden, finden Sie in den Produktkompatibilitätsmatrizes auf der unternehmensweiten Support-Site (<http://www.ibm.com/support>).

Installationsort der Treiber

Anmerkung: ODBC-Treiber müssen auf jedem Computer installiert und konfiguriert werden, auf dem eine Verarbeitung möglich ist.

- Wenn Sie IBM SPSS Modeler im lokalen Modus (Standalone-Modus) ausführen, müssen die Treiber auf dem lokalen Computer installiert sein.
- Wenn Sie IBM SPSS Modeler im verteilten Modus mit einer fernen IBM SPSS Modeler Server-Instanz ausführen, müssen die ODBC-Treiber auf dem Computer installiert sein, auf dem IBM SPSS Modeler Server installiert ist. Beachten Sie bei IBM SPSS Modeler Server auf UNIX-Systemen auch "Konfiguration von ODBC-Treibern auf UNIX-Systemen" weiter hinten in diesem Abschnitt.
- Wenn Sie von IBM SPSS Modeler und IBM SPSS Modeler Server auf dieselben Datenquellen zugreifen müssen, müssen die ODBC-Treiber auf beiden Computern installiert sein.
- Wenn Sie IBM SPSS Modeler über Terminaldienste ausführen, müssen die ODBC-Treiber auf dem Terminaldiensteserver installiert sein, auf dem Sie IBM SPSS Modeler installiert haben.

Konfiguration von ODBC-Treibern auf UNIX-Systemen

Standardmäßig ist der DataDirect-Treibermanager nicht für IBM SPSS Modeler Server auf UNIX-Systemen konfiguriert. Geben Sie folgende Befehle ein, um UNIX so zu konfigurieren, dass der DataDirect-Treibermanager geladen wird:

```
cd <Modeler Server-Installationsverzeichnis>/bin
rm -f libspssodbc.so
ln -s libspssodbc_datadirect.so libspssodbc.so
```

Dadurch wird die Standardverknüpfung entfernt und eine Verknüpfung zum DataDirect-Treibermanager wird erstellt.

Anmerkung: Der UTF16-Treiberwrapper ist erforderlich, um SAP HANA- oder IBM DB2-CLI-Treiber für einige Datenbanken verwenden zu können. DashDB erfordert den IBM DB2 CLI-Treiber. Geben Sie stattdessen die folgenden Befehle ein, um einen Link für den UTF16-Treiberwrapper zu erstellen:

```
rm -f libspssodbc.so
ln -s libspssodbc_datadirect_utf16.so libspssodbc.so
```

So konfigurieren Sie SPSS Modeler Server:

1. Konfigurieren Sie das Startscript `modelersrv.sh` von SPSS Modeler Server zur Verwendung der Umgebungsdatei `odbc.sh` von IBM SPSS Data Access Pack als Quelle, indem Sie `modelersrv.sh` die folgende Zeile hinzufügen:

```
. /<PfadzurSDAP-Installation>/odbc.sh
```

Dabei ist `<PfadzurSDAP-Installation>` der vollständige Pfad zu Ihrer IBM SPSS Data Access Pack-Installation.

2. Starten Sie SPSS Modeler Server erneut.

Fügen Sie dem Datenquellennamen in Ihrer Datei `odbc.ini` zusätzlich und nur für SAP HANA und IBM DB2 die folgende Parameterdefinition hinzu, um Pufferüberläufe während der Verbindung zu vermeiden:
`DriverUnicodeType=1`

Anmerkung: Der Wrapper `libspssodbc_datadirect_utf16.so` ist auch mit den anderen von SPSS Modeler Server unterstützten ODBC-Treibern kompatibel.

Anmerkung: Die oben genannten Regeln gelten speziell für den Zugriff auf Daten in einer Datenbank. Andere Dateizugriffe, wie das Öffnen und Speichern von Stream-, Projekt-, Modell-, Knoten-, PMMI-, Ausgabe- und Scriptdateien, finden stets auf dem Client statt und werden stets unter Berücksichtigung des Dateisystems auf dem Client-Computer angegeben. Ferner legt der Befehl "Verzeichnis festlegen" in SPSS Modeler das Arbeitsverzeichnis von *lokalen* Clientobjekten (z. B. Streams) fest, er hat jedoch keinerlei Auswirkung auf das Arbeitsverzeichnis des Servers.

UNIX und SPSS Statistics

Informationen zum Konfigurieren von SPSS Modeler Server unter UNIX, um mit der IBM SPSS Statistics-Datenzugriffstechnologie zu arbeiten, finden Sie in Anhang B, „Konfigurieren von UNIX-Startscripts“, auf Seite 79.

Referenzieren von Datendateien

Windows. Wenn Sie Daten auf demselben Computer wie IBM SPSS Modeler Server speichern, wird empfohlen, den Pfad zu den Daten aus der Perspektive des Server-Computers anzugeben (z. B. `C:\ServerData\Sales 1998.csv`). Die Leistung ist besser, wenn die Datei nicht über das Netz gesucht werden muss.

Sind die Daten auf einem anderen Host gespeichert, wird die Verwendung von UNC-Dateiverweisen empfohlen (z. B. `\\Mein_Datenserver\ServerData\Sales 1998.csv`). Beachten Sie, dass UNC-Namen nur funktionieren, wenn der Pfad den Namen der gemeinsamen Netzressource enthält. Der referenzierende Computer muss die Berechtigung zum Lesen der angegebenen Datei haben. Wenn Sie häufig vom Modus für verteilte Analysen in den Modus für lokale Analysen wechseln, verwenden Sie UNC-Dateiverweise, da diese unabhängig vom Modus immer funktionieren.

UNIX. Zum Referenzieren von Daten auf einem UNIX-Server verwenden Sie die vollständige Dateispezifizierung und normale Schrägstriche (z. B. `/public/data/ServerData/Umsatz1998.csv`). Vermeiden Sie im UNIX-Verzeichnis und in Dateinamen für Daten, die mit IBM SPSS Modeler Server verwendet werden, den umgekehrten Schrägstrich. Es spielt keine Rolle, ob eine Textdatei im UNIX- oder DOS-Format vorhanden ist, da beide automatisch verarbeitet werden.

Importieren von IBM SPSS Statistics-Datendateien

Wenn Sie auf Ihrer Site IBM SPSS Statistics Server ausführen, kann es vorkommen, dass Benutzer IBM SPSS Statistics-Daten im verteilten Modus importieren oder exportieren möchten. Beachten Sie, dass IBM SPSS Modeler im verteilten Modus das Dateisystem des Servers verwendet. Der IBM SPSS Statistics Client arbeitet auf dieselbe Weise. Zum Importieren und Exportieren zwischen den beiden Anwendungen müssen die beiden Clients in demselben Modus arbeiten. Ist dies nicht der Fall, weichen die Ansichten der Dateisysteme voneinander ab und die gemeinsame Nutzung der Dateien ist nicht möglich. Die IBM SPSS Statistics-Knoten in IBM SPSS Modeler können den IBM SPSS Statistics Client automatisch starten, die Benutzer müssen jedoch zunächst sicherstellen, dass der IBM SPSS Statistics Client und IBM SPSS Modeler in demselben Modus arbeiten.

Installationsanweisungen

Informationen zur Installation von IBM SPSS Modeler Server finden Sie in den Installationsanweisungen, die als PDF-Dateien über den Produktdownload zur Verfügung stehen. Weitere Dokumente sind für Windows und UNIX verfügbar.

Vollständige Informationen zum Installieren und Verwenden des IBM SPSS Modeler-Clients finden Sie in den PDF-Dateien, die über den Produktdownload zur Verfügung stehen. Separate Installationsdokumente sind abhängig von Ihrer jeweiligen Lizenz verfügbar.

Kapitel 3. IBM SPSS Modeler-Support

Dieser Abschnitt wurde für Administratoren und Help-Desk-Mitarbeiter geschrieben, die Support für die Benutzer von IBM SPSS Modeler anbieten. Die folgenden Themen werden beschrieben:

- Anmeldung bei IBM SPSS Modeler Server (oder Standalone-Ausführung durch Trennen der Verbindung zu einem Server)
- Vom Benutzer benötigte Daten- und Dateisysteme
- Benutzerkonten und Dateiberechtigungen für IBM SPSS Modeler Server
- Unterschiede bei den Ergebnissen, die Benutzer beim Wechsel zwischen IBM SPSS Modeler Server und IBM SPSS Modeler sehen

Verbinden mit IBM SPSS Modeler Server

IBM SPSS Modeler kann als eigenständige Anwendung oder als Client ausgeführt werden, der direkt mit IBM SPSS Modeler Server oder über das Plug-in Coordinator of Processes von IBM SPSS Collaboration and Deployment Services mit einer IBM SPSS Modeler Server-Instanz oder einem Server-Cluster verbunden ist. Der aktuelle Verbindungsstatus wird unten links im IBM SPSS Modeler-Fenster angezeigt.

Wenn Sie eine Verbindung zu einem Server herstellen möchten, können Sie den Namen des Servers, mit dem eine Verbindung hergestellt werden soll, manuell eingeben oder einen zuvor definierten Namen auswählen. Wenn Sie IBM SPSS Collaboration and Deployment Services verwenden, können Sie im Dialogfeld für die Anmeldung beim Server eine Liste von Servern bzw. Server-Clustern durchsuchen. Die Möglichkeit, die auf einem Netz ausgeführten Statistics-Dienste zu durchsuchen, wird über den Coordinator of Processes bereitgestellt.

So stellen Sie eine Verbindung mit einem Server her:

1. Klicken Sie im Menü "Tools" auf die Option **Anmelden beim Server**. Das Dialogfeld "Anmelden beim Server" wird geöffnet. Alternativ können Sie auf den Bereich des Verbindungsstatus im IBM SPSS Modeler-Fenster doppelklicken.
2. Legen Sie in diesem Dialogfeld die Optionen zum Verbinden mit dem lokalen Server-Computer fest oder wählen Sie eine Verbindung in der Tabelle aus.
 - Klicken Sie auf **Hinzufügen** bzw. **Bearbeiten**, um eine Verbindung hinzuzufügen bzw. zu bearbeiten. Weitere Informationen finden Sie in „Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung“ auf Seite 22.
 - Klicken Sie auf **Suche**, um auf einen Server bzw. Server-Cluster in Coordinator of Processes zuzugreifen. Weitere Informationen finden Sie in „Suchen nach Servern in IBM SPSS Collaboration and Deployment Services“ auf Seite 22.

Servertabelle. Diese Tabelle enthält die Menge der definierten Serververbindungen. In der Tabelle werden die Standardverbindung, der Servername sowie die Beschreibung und Portnummer angegeben. Sie können manuell eine neue Verbindung hinzufügen sowie eine bestehende Verbindung auswählen bzw. danach suchen. Um einen bestimmten Server als Standardverbindung einzurichten, aktivieren Sie in der Tabelle für die Verbindung das Kontrollkästchen in der Spalte "Standard".

Standarddatenpfad. Geben Sie einen Pfad an, der für Daten auf dem Server-Computer verwendet wird. Mit der Auslassungsschaltfläche (...) wechseln Sie zum gewünschten Verzeichnis.

Berechnungsnachweise festlegen. Lassen Sie dieses Kontrollkästchen unausgewählt, um die Funktion für **Single Sign-on** zu aktivieren. Diese versucht, Sie mithilfe Ihres lokalen Benutzernamens und Kennworts beim Server anzumelden. Falls ein Single Sign-on nicht möglich ist oder Sie das Kontrollkästchen zur Inaktivierung des Single Sign-on aktivieren (z. B. zur Anmeldung an einem Administratorkonto), wird ein weiteres Fenster angezeigt, in dem Sie aufgefordert werden, Ihre Berechnungsnachweise einzugeben.

Benutzer-ID. Geben Sie den Benutzernamen ein, mit dem die Anmeldung beim Server erfolgen soll.

Kennwort. Geben Sie das Kennwort ein, das zum angegebenen Benutzernamen gehört.

Domäne. Geben Sie die Domäne an, mit der die Anmeldung beim Server erfolgen soll. Ein Domänenname ist nur dann erforderlich, wenn sich der Server-Computer in einer anderen Windows-Domäne befindet als der Client-Computer.

3. Klicken Sie auf **OK**, um die Verbindung herzustellen.

So trennen Sie eine Verbindung mit einem Server:

1. Klicken Sie im Menü "Tools" auf die Option **Anmelden beim Server**. Das Dialogfeld "Anmelden beim Server" wird geöffnet. Alternativ können Sie auf den Bereich des Verbindungsstatus im IBM SPSS Modeler-Fenster doppelklicken.
2. Wählen Sie im Dialogfeld den lokalen Server aus und klicken Sie auf **OK**.

Konfigurieren von Single Sign-on

Sie können mit Single Sign-on eine Verbindung zu einer Instanz von IBM SPSS Modeler Server herstellen, die auf einer beliebigen unterstützten Plattform ausgeführt wird. Wenn Sie eine Verbindung mit Single Sign-on herstellen wollen, müssen Sie zunächst Ihren IBM SPSS Modeler-Server und die Client-Computer konfigurieren.

Wenn Sie Single Sign-On verwenden, um die Verbindung sowohl zu IBM SPSS Modeler Server als auch zu IBM SPSS Collaboration and Deployment Services herzustellen, müssen Sie zuerst die Verbindung zu IBM SPSS Collaboration and Deployment Services herstellen, bevor Sie die Verbindung zu IBM SPSS Modeler herstellen.

IBM SPSS Modeler Server verwendet Kerberos für Single Sign-On.

Kerberos ist eine Kernkomponente von Windows Active Directory und bei den folgenden Informationen wird davon ausgegangen, dass eine Active Directory-Infrastruktur vorliegt. Insbesondere gilt Folgendes:

- Der Client-Computer ist ein Windows-Computer, der mit einer Active Directory-Domäne verbunden ist
- Der Clientbenutzer hat sich mit einem Domänenkonto am Computer angemeldet. Das zum Anmelden verwendete Verfahren ist dabei nicht relevant, d. h., die Anmeldung kann über eine Smartcard, einen elektronischen Fingerabdruck usw. erfolgen
- IBM SPSS Modeler Server kann die Clientbenutzer-Berechtigungsnachweise durch Bezugnahme auf den Active Directory-Domänencontroller validieren

In dieser Dokumentation wird beschrieben, wie Windows- und UNIX-Server für diese Art der Authentifizieren konfiguriert werden können. Es sind auch andere Konfigurationen möglich, diese wurden jedoch nicht getestet.

Zur Interaktion mit den meisten modernen, sicheren Active Directory-Installationen müssen Sie das Softwarepaket für starke Verschlüsselung für Java installieren, da die erforderlichen Verschlüsselungsalgorithmen nicht standardmäßig unterstützt werden. Sie müssen das Softwarepaket für den Client und den Server installieren. Eine Fehlermeldung wie `Illegal key size` (Unzulässige Schlüsselgröße) wird auf dem Client angezeigt, wenn eine Serververbindung aufgrund eines nicht installierten Softwarepakets fehlschlägt. Siehe „Installieren der Verschlüsselung mit unbegrenzter Stärke“ auf Seite 54.

Name des Service-Principals (SPN)

Jede Serverinstanz muss zur eigenen Identifizierung einen eindeutigen *Namen des Service-Principals* (SPN - Service Principal Name) registrieren und der Client muss denselben SPN angeben, wenn er eine Verbindung zum Server herstellt.

Ein SPN für eine Instanz von SPSS Modeler Server hat das folgende Format:

```
modelerserver/<Host>:<Port>
```

Beispiel:

```
modelerserver/jdoemachine.spss.com:28054
```

Beachten Sie, dass der Hostname mit der zugehörigen DNS-Domäne (in diesem Beispiel: spss.com) qualifiziert und die Domäne dem Kerberos-Realm zugeordnet sein muss.

Die Kombination aus Hostname und Portnummer stellt die Eindeutigkeit des SPN sicher (da jede Instanz auf einem Host auf einem anderen Port empfangsbereit sein muss). Da sowohl Client als auch Server bereits Hostnamen und Portnummer aufweisen, sind sie in der Lage, den jeweiligen SPN für die Instanz zu bilden. Als zusätzlicher Konfigurationsschritt muss der SPN in der Kerberos-Datenbank registriert werden.

Registrieren des SPN unter Windows

Wenn Sie Active Directory als Kerberos-Implementierung nutzen, verwenden Sie den Befehl `setspn`, um den SPN zu registrieren. Zum Ausführen des Befehls müssen die folgenden Bedingungen erfüllt sein:

- Sie müssen an einem Domänencontroller angemeldet sein
- Sie müssen die Eingabeaufforderung mit erweiterten Berechtigungen (als Administrator) ausführen
- Sie müssen ein Mitglied der Gruppe der Domänenadministratoren sein (oder über die entsprechende, Ihnen von einem Domänenadministrator delegierte Berechtigung verfügen)

Weitere Informationen finden Sie in den folgenden Artikeln:

- Referenz für `Setspn`-Befehlszeile
- Delegating Authority to Modify SPNs

Für die Standardinstanz, die auf dem Standardport empfangsbereit ist (z. B. 28054) und unter dem Konto "Lokales System" ausgeführt wird, müssen Sie den SPN für den Namen des Server-Computers registrieren. Beispiel:

```
setspn -s modelerserver/jdoemachine.spss.com:28054 jdoemachine
```

Für jede nachfolgende (Profil-)Instanz, die auf einem benutzerdefinierten Port (z. B. 29000) empfangsbereit ist und unter einem beliebigen Benutzerkonto (z. B. jdoe) mit der auf Y gesetzten Option `start_process_as_login_user` ausgeführt wird, müssen Sie den SPN für den Namen des Dienstbenutzerkontos registrieren:

```
setspn -s modelerserver/jdoemachine.spss.com:29000 jdoe
```

Beachten Sie, dass in diesem Fall (wenn als Dienstkonto nicht "Lokales System" verwendet wird), es nicht ausreicht, den SPN zu registrieren, damit ein Client eine Verbindung herstellen kann. Weitere Konfigurationsschritte werden im nächsten Abschnitt beschrieben.

So zeigen Sie die für das Konto jdoe registrierten SPNs an:

```
setspn -l jdoe
```

Registrieren des SPN unter UNIX

Wenn Sie Active Directory als Kerberos-Implementierung nutzen, können Sie den Befehl `setspn` verwenden, wie im vorherigen Abschnitt zu Windows beschrieben. Dies setzt voraus, dass Sie bereits ein Computer- oder Benutzerkonto im Verzeichnis erstellt haben. Alternativ können Sie `ktpass` verwenden, wie in „Konfigurieren von IBM SPSS Modeler Server unter UNIX und Linux“ auf Seite 19 dargestellt.

Wenn Sie eine andere Kerberos-Implementierung verwenden, können Sie mit Ihrem Kerberos-Verwaltungstool der Kerberos-Datenbank den Service-Principal hinzufügen. Um den SPN in einen Kerberos-Principal umzuwandeln, müssen Sie den Namen des Kerberos-Realms anhängen. Beispiel:

modelerserver/jdoemachine.spss.com:28054@MODELERSSO.COM

Fügen Sie diesen Principal und das Kennwort dem Chiffrierschlüssel des Servers hinzu. Der Chiffrierschlüssel muss für jede auf dem Host ausgeführte Instanz einen Eintrag enthalten.

Konfigurieren von IBM SPSS Modeler Server unter Windows

Im Standardszenario, bei dem der SPSS Modeler Server-Dienst unter dem Konto "Lokales System" ausgeführt wird, verwendet der Server native Windows-APIs zum Authentifizieren der Benutzerberechtigungs-nachweise und es ist keine zusätzliche Konfiguration erforderlich.

Im alternativen Szenario, bei dem der SPSS Modeler Server-Dienst unter einem dedizierten Benutzerkonto ausgeführt wird und `start_process_as_login_user` auf `Y` gesetzt ist, verwendet der Server Java-APIs zum Authentifizieren der Benutzerberechtigungs-nachweise und es ist eine zusätzliche Konfiguration auf dem Server erforderlich.

Stellen Sie zuerst sicher, dass das Standardszenario funktioniert. Der Client sollte in der Lage sein, über SSO die Verbindung zur Standardinstanz herzustellen, die unter dem Konto "Lokales System" ausgeführt wird. Damit wird die clientseitige (unveränderte) Konfiguration validiert. Sie müssen den SPN wie zuvor beschrieben für die Standardinstanz registrieren.

Führen Sie dann die folgenden Schritte aus:

1. Erstellen Sie das Verzeichnis `<MODELERSERVER>\config\sso`.
2. Erstellen Sie eine Datei mit dem Namen `krb5.conf` im Ordner `sso`, den Sie in Schritt 1 erstellt haben. Anweisungen zum Erstellen dieser Datei finden Sie unter Schritt 3 in „Konfigurieren des IBM SPSS Modeler-Clients“ auf Seite 19. Die Datei muss auf dem Server und dem Client identisch sein.
3. Erstellen Sie mithilfe des folgenden Befehls die Datei `krb5.keytab` im SSO-Verzeichnis des Servers:

```
<MODELERSERVER>\jre\bin\ktab -a <spn>@<Realm> -k krb5.keytab
```

Beispiel:

```
"..\jre\bin\ktab.exe" -a modelerserver/jdoemachine.spss.com:29000@SPSS.COM  
-k krb5.keytab
```

Daraufhin werden Sie zur Angabe eines Kennworts aufgefordert. Als Kennwort müssen Sie das Kennwort des Dienstkontos eingeben. Beim Dienstkonto `jdoe` zum Beispiel, müssen Sie das Kennwort für den Benutzer `jdoe` eingeben.

Das Dienstkonto selbst wird im Chiffrierschlüssel nicht genannt, allerdings haben Sie zuvor den SPN mithilfe von `setspn` für dieses Konto registriert. Dies bedeutet, der Dienstprincipal und das Dienstkonto haben ein und dasselbe Kennwort.

Bei jeder neuen Instanz (Profil), die Sie erstellen, müssen Sie den SPN für diese Instanz mithilfe von `setspn` (siehe „Konfigurieren von Serverprofilen“ auf Seite 29 und „Name des Service-Principals (SPN)“ auf Seite 16) registrieren und im Chiffrierschlüssel mithilfe von `jre\bin\ktab` einen Eintrag hinzufügen. Es gibt nur eine Chiffrierschlüsseldatei. Diese muss einen Eintrag für jede Instanz enthalten, die nicht als lokales System ausgeführt wird. Die Standardinstanz oder eine andere als lokales System ausgeführte Instanz muss im Chiffrierschlüssel nicht angegeben werden, da sie für Authentifizierungen die Windows-APIs verwendet. Windows-APIs verwenden den Chiffrierschlüssel nicht.

So prüfen Sie, ob im Chiffrierschlüssel eine Instanz angegeben ist:

```
ktab.exe -l -e -k krb5.keytab
```

Möglicherweise werden für jeden Principal mehrere Einträge mit verschiedenen Verschlüsselungstypen angezeigt, dies ist jedoch normal.

Konfigurieren von IBM SPSS Modeler Server unter UNIX und Linux Voraussetzungen

IBM SPSS Modeler Server verwendet Windows Active Directory (AD) zum Aktivieren von Single Sign-on. Hierfür müssen die folgenden Voraussetzungen erfüllt sein:

- Der SPSS Modeler Client-Computer (Windows) ist Mitglied einer Active Directory-Domäne.
- Der Clientbenutzer meldet sich mit einem AD-Domänenkonto am Computer an.
- Der SPSS Modeler Server-Computer (UNIX) wird über einen vollständig qualifizierten Domänennamen angegeben, dessen Stammelement sich in der AD-DNS-Domäne befindet. Wenn die DNS-Domäne `modelersso.com` ist, könnte der Hostname des Servers z. B. `myserver.modelersso.com` sein.
- Die AD-DNS-Domäne unterstützt Vorwärtssuchen und umgekehrte Suchen für den SPSS Modeler Server-Hostnamen.

Wenn der SPSS Modeler Server-Computer kein Mitglied der AD-Domäne ist, müssen Sie ein Domänenbenutzerkonto zur Darstellung des Service im Verzeichnis erstellen. Sie könnten z. B. ein Domänenkonto `ModelerServer` erstellen.

So konfigurieren Sie SPSS Modeler Server unter UNIX oder Linux:

1. Erstellen Sie im SPSS Modeler Server-Ordner `config` den Unterordner `sso`.
2. Erstellen Sie im Ordner `sso` eine `keytab`-Datei. Die Generierung der `keytab`-Datei kann auf der AD-Seite erfolgen. Allerdings sind die Anforderungen unterschiedlich, je nachdem, ob der SPSS Modeler Server-Computer ein Mitglied der AD-Domäne ist:

- Wenn der SPSS Modeler Server-Computer ein Mitglied der AD-Domäne ist, verwenden Sie den Computerkontonamen als Servicebenutzernamen:

```
ktpass -princ <SPN>@<realm> -mapUser <Domäne>\<Computerkonto> -pass <Kennwort> -out <Ausgabedatei> -ptype KRB5_NT_PRINCIPAL
```

Beispiel:

```
ktpass -princ modelerserver/myserver.modelersso.com:28054@MODELERSSO.COM -mapUser MODELERSSO\myserver$ -pass Pass1234 -out c:\myserver.keytab -ptype KRB5_NT_PRINCIPAL
```

- Wenn der SPSS Modeler Server-Computer **kein** Mitglied der AD-Domäne ist, geben Sie das Domänenbenutzerkonto, das Sie als Voraussetzung erstellt haben, als Servicebenutzer an:

```
ktpass -princ <SPN>@<realm> -mapUser <Domäne>\<Benutzerkonto> -mapOp set -pass <Kennwort> -out <Ausgabedatei> -ptype KRB5_NT_PRINCIPAL
```

Beispiel:

```
ktpass -princ modelerserver/myserver.modelersso.com:28054@MODELERSSO.COM -mapUser MODELERSSO\ModelerServer -mapOp set -pass Pass1234 -out c:\myserver.keytab -ptype KRB5_NT_PRINCIPAL
```

Weitere Informationen finden Sie in Referenz für `Ktpass`-Befehlszeile.

3. Benennen Sie die `keytab`-Datei im Ordner `sso` in `krb5.keytab` um.

Anmerkung: Wenn Sie den Servercomputer wieder mit der Domäne verbinden, müssen Sie eine neue `keytab`-Datei generieren.

4. Erstellen Sie eine Datei mit dem Namen `krb5.conf` im Ordner `sso`, den Sie in Schritt 1 erstellt haben. Anweisungen zum Erstellen dieser Datei finden Sie unter Schritt 3 in „Konfigurieren des IBM SPSS Modeler-Clients“. Die Datei muss auf dem Server und dem Client identisch sein.

Konfigurieren des IBM SPSS Modeler-Clients

1. Aktivieren Sie Java für den Zugriff auf den TGT-Sitzungsschlüssel:
 - a. Klicken Sie im Menü **Start** auf **Ausführen**.
 - b. Geben Sie `regedit` ein und klicken Sie auf **OK**, um den **Registrierungseditor** zu öffnen.

- c. Navigieren Sie zu der Registry-Position für das Betriebssystem der lokalen Maschine:
 - Unter Windows XP: Computer\HKEY_LOCAL_MACHINE\System\CurrentControlSet\Control\Lsa\Kerberos
 - Unter Windows Vista oder Windows 7: Computer\HKEY_LOCAL_MACHINE\System\CurrentControlSet\Control\Lsa\Kerberos\Parameters
 - d. Klicken Sie mit der rechten Maustaste auf den Ordner und wählen Sie **Neu > DWORD** aus. Der Name des neuen Wertes sollte allowtgtsessionkey lauten.
 - e. Setzen Sie den Wert für allowtgtsessionkey auf den Hexadezimalwert 1, also 0x0000001.
 - f. Schließen Sie den **Registrierungseditor**.
 - g. **Hinweis:** Es gibt ein bekanntes Problem, wenn das Benutzerkonto zur Gruppe der lokalen Administratoren gehört und die Benutzerkontensteuerung (UAC - User Account Control) aktiviert ist. In diesem Fall ist der Sitzungsschlüssel im abgerufenen Service-Ticket leer, sodass die SSO-Authentifizierung fehlschlägt. Führen Sie zur Vermeidung dieses Problems eine der folgenden Aktionen aus:
 - Führen Sie die Anwendung als Administrator aus
 - Inaktivieren Sie die Benutzerkontensteuerung
 - Verwenden Sie ein Konto, das kein Administratorkonto ist
2. Erstellen Sie im Ordner config des IBM SPSS Modeler-Installationsverzeichnis den Ordner sso.
 3. Erstellen Sie im Ordner sso die Datei krb5.conf. Anweisungen zum Erstellen einer Datei krb5.conf finden Sie unter http://web.mit.edu/kerberos/krb5-current/doc/admin/conf_files/krb5_conf.html. Im Folgenden finden Sie ein Beispiel für eine Datei krb5.conf:

```
[libdefaults]
    default_realm = MODELERSO.COM
    dns_lookup_kdc = true
    dns_lookup_realm = true

[realms]
    MODELERSO.COM = {
        kdc = ad.modelersso.com:88
        admin_server = ad.modelersso.com:749
        default_domain = modelersso.com
    }

[domain_realm]
    .modelersso.com = MODELERSO.COM
    modelersso.com = MODELERSO.COM
```

4. Starten Sie die lokale Maschine und die Servermaschine erneut.

Abrufen der Gruppenzugehörigkeit des SSO-Benutzers

Wenn sich ein Benutzer über SSO bei SPSS Modeler Server anmeldet und der Server als Nicht-Root-Server ausgeführt wird, ist der Name des authentifizierten Benutzers keinem Betriebssystembenutzerkonto zugeordnet. Der Server ist nicht in der Lage, die Betriebssystemgruppenzugehörigkeit des Benutzers abzurufen. Wie wird in diesem Fall die Gruppenkonfiguration durchgeführt?

Angenommen, der Benutzer ist in einem LDAP-Verzeichnis (z. B. einem Active Directory-Verzeichnis) registriert und kann die Gruppenzugehörigkeit vom LDAP-Server anfordern. SPSS Modeler Server kann den LDAP-Provider in IBM SPSS Collaboration and Deployment Services nach der Gruppenzugehörigkeit abfragen.

Zwei Eigenschaften in der Datei options.cfg in der SPSS Modeler Server-Instanz steuern den Zugriff des Servers auf das Repository von IBM SPSS Collaboration and Deployment Services:

```
repository_enabled, N
repository_url, ""
```

Zur Aktivierung der Gruppensuche müssen beide Eigenschaften gesetzt werden. Beispiel:

```
repository_enabled, Y
repository_url, "http://jdoemachine.spss.ibm.com:9083"
```

Die Repository-Verbindung wird nur für SSO-Gruppensuchen verwendet, daher müssen Sie diese Eigenschafteneinstellungen nicht ändern, es sei denn, Sie möchten diese Funktion nutzen.

Damit die Gruppensuche ordnungsgemäß funktioniert, müssen Sie zuerst in der Konfiguration für Ihr Repository einen LDAP- oder Active Directory-Provider hinzufügen und anschließend SSO über diesen Provider wie folgt aktivieren:

1. Starten Sie den IBM SPSS Deployment Manager-Client und wählen Sie **Datei > Neu > Verbindung zu verwaltetem Server...** aus, um eine Verbindung zu einem verwalteten Server für Ihr Repository zu erstellen (sofern noch keine vorhanden ist).
2. Melden Sie sich bei der Verbindung zum verwalteten Server an und erweitern Sie den Ordner **Konfiguration**.
3. Klicken Sie mit der rechten Maustaste auf **Sicherheitsprovider**, wählen Sie **Neu > Sicherheitsproviderdefinition...** aus und geben Sie die entsprechenden Werte ein. Klicken Sie im Dialogfeld auf **Hilfe**, um weitere Informationen zu erhalten.
4. Erweitern Sie den Ordner für Single Sign-On Provider, klicken Sie mit der rechten Maustaste auf **Kerberos-SSO-Provider** und wählen Sie **Öffnen** aus.
5. Klicken Sie auf **Aktivieren**, wählen Sie Ihren Sicherheitsprovider aus und klicken Sie dann auf **Speichern**. Sie müssen hier keine weiteren Angaben machen, es sei denn, Sie möchten den SSO nutzen (die Aktivierung des Providers allein reicht aus, um Gruppensuchen zu ermöglichen).

Wichtig: Damit die Gruppensuche ordnungsgemäß funktioniert, müssen Sie hier den gleichen Kerberos-Provider konfigurieren, der auch für SPSS Modeler Server konfiguriert wurde. Insbesondere ist zu beachten, dass die Provider innerhalb desselben Kerberos-Realms eingesetzt werden müssen. Wenn sich also ein Benutzer bei der SPSS Modeler Server-Instanz über SSO anmeldet und diese ihn als `jdoe@SPSS.COM` identifiziert (wobei `SPSS.COM` für das Realm steht), wird davon ausgegangen, dass der Sicherheitsprovider in IBM SPSS Collaboration and Deployment Services diesen Namen des Benutzerprincipals erkennt und die entsprechende Gruppenzugehörigkeit aus dem LDAP-Verzeichnis zurückgibt.

Single Sign-on für Datenquellen

In IBM SPSS Modeler haben Sie die Möglichkeit, über Single Sign-on eine Verbindung zu Datenbanken herzustellen. Wenn Sie eine Datenbankverbindung über Single Sign-on erstellen wollen, müssen Sie zuerst mithilfe Ihrer ODBC-Verwaltungssoftware eine Datenquelle und ein Single-Sign-on-Token ordnungsgemäß konfigurieren. Wenn dann in IBM SPSS Modeler eine Verbindung zu einer Datenbank hergestellt wird, verwendet IBM SPSS Modeler dieses Single-Sign-on-Token und der Benutzer wird nicht aufgefordert, sich an der Datenquelle anzumelden.

Wurde die Datenquelle jedoch nicht ordnungsgemäß für Single Sign-on konfiguriert, fordert IBM SPSS Modeler den Benutzer auf, sich an der Datenquelle anzumelden. Nach Angabe der gültigen Berechtigungsnachweise kann der Benutzer weiter auf die Datenquelle zugreifen.

Detaillierte Informationen zur Konfiguration von ODBC-Datenquellen auf Ihrem System bei aktiviertem Single Sign-on finden Sie in der Dokumentation Ihres Datenbankanbieters. Nachfolgend sind beispielhaft allgemeine Schritte aufgeführt, die möglicherweise durchzuführen sind:

1. Konfigurieren Sie Ihre Datenbank so, dass sie Kerberos-Single Sign-on unterstützt.
2. Erstellen Sie auf dem IBM SPSS Modeler Server-System eine ODBC-Datenquelle und testen Sie sie. Für die DSN-Verbindung sollte keine Benutzer-ID und kein Kennwort benötigt werden.
3. Stellen Sie über Single Sign-on eine Verbindung zu IBM SPSS Modeler Server her und verwenden Sie die in Schritt 2 erstellte und validierte ODBC-Datenquelle.

Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung

Serververbindungen können manuell im Dialogfeld "Anmelden beim Server" bearbeitet bzw. hinzugefügt werden. Durch Klicken auf "Hinzufügen" können Sie auf ein leeres Dialogfeld vom Typ "Server hinzufügen/bearbeiten" zugreifen, in dem Sie Details zur Serververbindung eingeben können. Durch Auswahl einer bestehenden Verbindung und Klicken auf "Bearbeiten" im Dialogfeld "Anmelden beim Server" wird das Dialogfeld "Server hinzufügen/bearbeiten" mit den Details für die betreffende Verbindung geöffnet, sodass Sie etwaige Änderungen vornehmen können.

Anmerkung: Sie können eine Serververbindung, die über IBM SPSS Collaboration and Deployment Services hinzugefügt wurde, nicht bearbeiten, da der Name, der Port und andere Details in IBM SPSS Collaboration and Deployment Services definiert sind. Es ist ein bewährtes Verfahren, für die Kommunikation mit IBM SPSS Collaboration and Deployment Services und dem SPSS Modeler-Client dieselben Ports zu verwenden. Diese können als `max_server_port` und `min_server_port` in der Datei `options.cfg` festgelegt werden.

So fügen Sie Serververbindungen hinzu:

1. Klicken Sie im Menü "Tools" auf die Option **Anmelden beim Server**. Das Dialogfeld "Anmelden beim Server" wird geöffnet.
2. Klicken Sie in diesem Dialogfeld auf **Hinzufügen**. Das Dialogfeld "Anmelden beim Server: Server hinzufügen/bearbeiten" wird angezeigt.
3. Geben Sie die Details für die Serververbindung ein und klicken Sie auf **OK**, um die Verbindung zu speichern und zum Dialogfeld "Anmelden beim Server" zurückzukehren.
 - **Server.** Geben Sie einen verfügbaren Server an oder wählen Sie einen aus der Liste aus. Der Server-Computer lässt sich anhand eines alphanumerischen Namens (z. B. *meinserver*) oder der dem Server-Computer zugewiesenen IP-Adresse (z. B. 202.123.456.78) identifizieren.
 - **Port.** Geben Sie die Portnummer an, die der Server überwacht. Wenn der Standardwert nicht funktioniert, fragen Sie Ihren Systemadministrator nach der richtigen Portnummer.
 - **Beschreibung.** Geben Sie eine optionale Beschreibung für diese Serververbindung ein.
 - **Verbindung verschlüsseln (mit SSL).** Legt fest, ob eine SSL-Verbindung (**Secure Sockets Layer**) verwendet werden soll. SSL ist ein weit verbreitetes Protokoll zum Schutz der über ein Netz versendeten Daten. Um diese Funktion verwenden zu können, muss SSL auf dem Server, auf dem sich IBM SPSS Modeler Server befindet, aktiviert sein. Wenden Sie sich gegebenenfalls an den lokalen Administrator, wenn Sie weitere Informationen benötigen.

So bearbeiten Sie Serververbindungen:

1. Klicken Sie im Menü "Tools" auf die Option **Anmelden beim Server**. Das Dialogfeld "Anmelden beim Server" wird geöffnet.
2. Wählen Sie in diesem Dialogfeld die zu bearbeitende Verbindung aus und klicken Sie dann auf **Bearbeiten**. Das Dialogfeld "Anmelden beim Server: Server hinzufügen/bearbeiten" wird angezeigt.
3. Ändern Sie die Details für die Serververbindung und klicken Sie auf **OK**, um die Änderungen zu speichern und zum Dialogfeld "Anmelden beim Server" zurückzukehren.

Suchen nach Servern in IBM SPSS Collaboration and Deployment Services

Anstatt eine Serververbindung manuell einzugeben, können Sie einen im Netz verfügbaren Server oder Server-Cluster über Coordinator of Processes auswählen. Diese Funktion ist in IBM SPSS Collaboration and Deployment Services verfügbar. Ein Server-Cluster ist eine Gruppe von Servern, aus denen Coordinator of Processes den Server ermittelt, der am besten für die Beantwortung einer Verarbeitungsanforderung geeignet ist.

Sie können zwar auch manuell Server im Dialogfeld "Anmelden beim Server" hinzufügen, durch die Suche nach verfügbaren Servern können Sie jedoch eine Verbindung zu Servern herstellen, ohne den richti-

gen Servernamen und die Portnummer kennen zu müssen. Diese Informationen werden automatisch bereitgestellt. Allerdings benötigen Sie auch bei dieser Variante die richtigen Anmeldeinformationen, wie Benutzername, Domäne und Kennwort.

Hinweis: Wenn Sie keinen Zugriff auf die Funktion Coordinator of Processes haben, können Sie dennoch den Namen des Servers, mit dem eine Verbindung hergestellt werden soll, manuell eingeben oder einen zuvor definierten Namen auswählen. Weitere Informationen finden Sie in „Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung“ auf Seite 22.

So suchen Sie nach Servern und Clustern:

1. Klicken Sie im Menü "Tools" auf die Option **Anmelden beim Server**. Das Dialogfeld "Anmelden beim Server" wird geöffnet.
2. Klicken Sie in diesem Dialogfeld auf **Suche**, um das Dialogfeld "Nach Servern suchen" zu öffnen. Wenn Sie versuchen, Coordinator of Processes zu durchsuchen, ohne bei IBM SPSS Collaboration and Deployment Services angemeldet zu sein, werden Sie zur Anmeldung aufgefordert.
3. Wählen Sie den Server bzw. Server-Cluster in der Liste aus.
4. Klicken Sie auf **OK**, um das Dialogfeld zu schließen und diese Verbindung der Tabelle im Dialogfeld "Anmelden beim Server" hinzuzufügen.

Daten- und Dateisysteme

Benutzer, die mit IBM SPSS Modeler Server arbeiten, müssen wahrscheinlich auf Datendateien und andere Datenquellen im Netz zugreifen sowie Dateien auf dem Netz speichern. Je nachdem benötigen sie die folgenden Informationen:

- **ODBC-Datenquelle.** Wenn Benutzer auf ODBC-Datenquellen zugreifen müssen, die auf dem Server-Computer definiert sind, benötigen Sie die Namen, Beschreibungen und Anmeldeinformationen (inklusive Datenanmelde-IDs und Kennwörter) für die Datenquellen.
- **Datendateizugriff.** Wenn Benutzer auf dem Server-Computer oder an einem anderen Standort im Netz auf Datendateien zugreifen wollen, müssen Sie die Namen und Speicherorte der Datendateien kennen.
- **Speicherort zum Speichern von Dateien.** Wenn Benutzer während einer Verbindung mit IBM SPSS Modeler Server Daten speichern, kann es vorkommen, dass sie die Dateien auf dem Server-Computer speichern möchten. Der Server ist jedoch häufig schreibgeschützt. In diesem Fall müssen Sie den Benutzern mitteilen, an welchem Speicherort sie die Datendateien speichern sollen. (Normalerweise handelt es sich bei diesem Speicherort um das Basisverzeichnis des Benutzers.)

Benutzerauthentifizierung

IBM SPSS Modeler Server verwendet das Betriebssystem des Server-Computers zum Authentifizieren der Benutzer, die eine Verbindung zum Server herstellen. Wenn ein Benutzer eine Verbindung mit SPSS Modeler Server herstellt, werden alle für den Benutzer ausgeführten Vorgänge im Sicherheitskontext des Benutzers ausgeführt. Der Zugriff auf Datenbanktabellen unterliegt den Benutzer- und/oder Kennwortberechtigungen für die Datenbank.

Windows. Unter Windows kann sich jeder Benutzer anmelden, der im Hostnetz über ein gültiges Konto verfügt. Bei der Standardauthentifizierung benötigen Benutzer Zugriff einschließlich Änderungsrecht auf das Verzeichnis `<Modeler-Server-Installation>\Tmp`. Ohne dieses Recht können sich die Benutzer unter Windows vom Client aus nicht mit der Standardauthentifizierung bei SPSS Modeler Server anmelden.

UNIX. Standardmäßig wird SPSS Modeler Server unter UNIX als Root ausgeführt. Auf diese Weise kann sich jeder Benutzer anmelden, der im Hostnetz über ein gültiges Konto verfügt. Außerdem wird der Zugriff auf die eigenen Dateien und Verzeichnisse eingeschränkt. Sie können jedoch SPSS Modeler Server so konfigurieren, dass er ohne die Rootberechtigung ausgeführt werden kann. In diesem Fall müssen Sie eine private Kennwortdatenbank erstellen, die für die Authentifizierung verwendet wird. Alle SPSS Mo-

deler-Benutzer müssen ein gemeinsames UNIX-Benutzerkonto verwenden (und somit gemeinsamen Zugriff auf Datendateien haben). Weitere Informationen finden Sie in „Konfigurieren als Nicht-Root-Benutzer mit privater Kennwortdatenbank“ auf Seite 88.

Konfigurieren von PAM

Unter Linux- und AIX-Plattformen verwendet SPSS Modeler Server PAM (Pluggable Authentication Module) zur Authentifizierung.

Um die PAM-Authentifizierung verwenden zu können, müssen die entsprechenden PAM-Module auf dem Hostsystem ordnungsgemäß konfiguriert sein. Soll zum Beispiel PAM mit LDAP zusammen eingesetzt werden, muss ein PAM-LDAP-Modul auf dem Betriebssystem des Hosts vorhanden und ordnungsgemäß konfiguriert sein. Weitere Informationen finden Sie in der Dokumentation zum Betriebssystem. Dies ist eine notwendige Voraussetzung, damit SPSS Modeler Server PAM verwenden kann.

Wenn SPSS Modeler Server für die Verwendung von PAM konfiguriert werden soll, bearbeiten Sie die SPSS Modeler Server-Datei "options.cfg" und fügen Sie die Zeile **authentication_methods, pam** hinzu bzw. bearbeiten Sie diese.

Sie können den Servicenamen `modelerserver` verwenden, um für SPSS Modeler Server eine bestimmte PAM-Konfiguration bereitzustellen, falls erforderlich. Die folgenden Schritte sind zum Beispiel für eine Konfiguration für Red Hat Linux erforderlich:

1. Wechseln Sie in das PAM-Konfigurationsverzeichnis. Beispiel: `/etc/pam.d`.
2. Erstellen Sie mit einem Texteditor eine neue Datei mit dem Namen "modelerserver".
3. Fügen Sie die PAM-Konfigurationsinformationen hinzu, die verwendet werden sollen. Beispiel:

```
auth      include    system-auth
account   include    system-auth
password  required  pam_deny.so
session   required  pam_deny.so
```

Anmerkung: Diese Zeilen können je nach Ihrer speziellen Konfiguration variieren. Weitere Informationen finden Sie in der Linux-Dokumentation.

4. Speichern Sie die Datei und starten Sie den Modeler-Service erneut.

Berechtigungen

Windows. Ein Benutzer, der eine Verbindung zu einer auf einem *NTFS*-Laufwerk installierten Server-Software herstellt, muss sich mit einem Konto anmelden, das über die folgenden Berechtigungen verfügt:

- Lese- und Ausführungsberechtigungen für das Installationsverzeichnis des Servers und dessen Unterverzeichnisse
- Lese-, Ausführungs- und Schreibberechtigungen für die Verzeichnisposition für temporäre Dateien

Unter Windows Server 2008 und höher kann nicht davon ausgegangen werden, dass Benutzer über diese Berechtigung verfügen. Erforderliche Berechtigungen müssen daher gegebenenfalls explizit festgelegt werden.

Wenn die Server-Software auf einem *FAT*-Laufwerk installiert ist, müssen Sie keine Berechtigungen festlegen, da die Benutzer uneingeschränkten Zugriff auf alle Dateien haben.

UNIX. Wenn Sie nicht die interne Authentifizierung verwenden, muss sich ein Benutzer, der eine Verbindung zur Server-Software herstellt, mit einem Konto anmelden, das über die folgenden Berechtigungen verfügt:

- Lese- und Ausführungsberechtigungen für das Installationsverzeichnis des Servers und dessen Unterverzeichnisse

- Lese-, Ausführungs- und Schreibberechtigungen für die Verzeichnisposition für temporäre Dateien

Dateierstellung

Wenn IBM SPSS Modeler Server auf Daten zugreift und diese verarbeitet, muss das Programm häufig eine temporäre Kopie dieser Daten auf der Festplatte speichern. Die Menge an erforderlichem Festplattenspeicher für die temporären Dateien hängt von der Größe der vom Endbenutzer analysierten Datendatei ab sowie von der jeweils durchgeführten Analyse. Weitere Informationen finden Sie in „Anforderungen an temporären Festplattenspeicher und Arbeitsspeicher“ auf Seite 10.

UNIX. Die UNIX-Versionen von IBM SPSS Modeler Server verwenden den UNIX-Befehl `umask`, um für die temporären Dateien Dateiberechtigungen festzulegen. Sie können die Standardberechtigungen des Servers überschreiben. Weitere Informationen finden Sie in „Steuern der Berechtigungen für die Dateierstellung“ auf Seite 80.

Unterschiede bei den Ergebnissen

Benutzer, die in beiden Modi Analysen durchführen, erhalten möglicherweise leichte Unterschiede bei den Ergebnissen von IBM SPSS Modeler und IBM SPSS Modeler Server. Diese Diskrepanz tritt normalerweise auf, wenn Datensätze anders sortiert oder gerundet werden.

Sortieren von Datensätzen. Wenn ein Stream Datensätze nicht ausdrücklich sortiert, kann die Reihenfolge der dargestellten Datensätze zwischen lokal und auf dem Server ausgeführten Streams voneinander leicht abweichen. Außerdem kann es zwischen Vorgängen, die in einer Datenbank ausgeführt werden, und Vorgängen, die in IBM SPSS Modeler Server ausgeführt werden, zu Unterschieden in der Reihenfolge kommen. Die Ursache für diese Unterschiede liegt in den verschiedenen Algorithmen, die jedes System zum Implementieren von Funktionen zum erneuten Sortieren der Datensätze verwendet, beispielsweise die Aggregation. Außerdem legt SQL die Reihenfolge, in der Datensätze aus einer Datenbank zurückgegeben werden, ohne ausdrücklichen Sortiervorgang nicht fest.

Unterschiede beim Runden. IBM SPSS Modeler verwendet im lokalen Modus ein anderes internes Format zum Speichern von Gleitkommawerten als IBM SPSS Modeler Server. Aufgrund der Unterschiede beim Runden können die Ergebnisse leicht voneinander abweichen.

Kapitel 4. IBM SPSS Modeler-Verwaltung

Dieses Kapitel enthält Informationen zum Starten und Stoppen von IBM SPSS Modeler Server, zum Konfigurieren verschiedener Serveroptionen, zum Konfigurieren von Gruppen und zur Interpretation der Protokolldatei. Es beschreibt die Verwendung von IBM SPSS Modeler Administration Console, einer Anwendung zur Konfiguration und Überwachung des Servers. Installationsanweisungen zu dieser Komponente finden Sie in den Installationsanweisungen zu IBM SPSS Modeler Server, die im Lieferumfang des Produkts enthalten sind.

Starten und Stoppen von IBM SPSS Modeler Server

IBM SPSS Modeler Server wird unter Windows als Dienst und unter UNIX als Dämonprozess ausgeführt.

Hinweis zur Planung: Beim Stoppen von IBM SPSS Modeler Server wird die Verbindung der einzelnen Endbenutzer unterbrochen und ihre Sitzung wird beendet. Versuchen Sie also, Serverneustarts in Zeiten geringer Nutzung zu legen. Wenn dies nicht möglich ist, sollten Sie die Benutzer unbedingt vor dem Anhalten des Servers benachrichtigen.

Starten, Stoppen und Statusprüfung unter Windows

Unter Windows steuern Sie IBM SPSS Modeler Server mit dem Dialogfeld "Dienste" in der Windows-Systemsteuerung.

1. **Windows XP.** Öffnen Sie das Windows-Startmenü. Wählen Sie **Einstellungen** und dann **Systemsteuerung** aus. Doppelklicken Sie auf **Verwaltung** und dann auf **Dienste**.
Windows 2003 oder 2008. Öffnen Sie das Windows-Startmenü. Wählen Sie **Systemsteuerung**, dann **Verwaltung** und anschließend **Dienste** aus.
2. Wählen Sie den Dienst **IBM SPSS Modeler Server <nn.n>** aus. Sie können nun seinen Status überprüfen, ihn starten bzw. stoppen und nach Bedarf die Startparameter bearbeiten.

Standardmäßig ist der Dienst für automatischen Start konfiguriert, d. h., wenn Sie ihn stoppen, wird er automatisch bei einem erneuten Booten des Computers neu gestartet. Wenn der Dienst auf diese Weise gestartet wird, wird er unbeaufsichtigt ausgeführt und der Computer kann abgemeldet werden, ohne dass dies Einfluss auf den Dienst hätte.

Starten, Stoppen und Statusprüfung unter UNIX

Unter UNIX können Sie IBM SPSS Modeler Server starten bzw. stoppen, indem Sie das Script *modelersrv.sh* im Installationsverzeichnis von IBM SPSS Modeler Server ausführen.

1. Wechseln Sie zum IBM SPSS Modeler Server-Installationsverzeichnis. Geben Sie beispielsweise an einer UNIX-Eingabeaufforderung Folgendes ein:

```
cd /usr/modelersrv
```

Dabei steht *modelersrv* für das Installationsverzeichnis von IBM SPSS Modeler Server.
2. Um den Server zu starten, geben Sie an der Eingabeaufforderung Folgendes ein:

```
./modelersrv.sh start
```
3. Um den Server zu stoppen, geben Sie an der Eingabeaufforderung Folgendes ein:

```
./modelersrv.sh stop
```
4. Um den Status von IBM SPSS Modeler Server zu überprüfen, geben Sie an einer UNIX-Eingabeaufforderung Folgendes ein:

```
./modelersrv.sh list
```

Untersuchen Sie dann die Ausgabe, die den Ausgaben des UNIX-Befehls `ps` ähnelt. Der erste Prozess in der Liste ist der IBM SPSS Modeler Server-Dämonprozess, die restlichen Prozesse sind IBM SPSS Modeler-Sitzungen.

Das IBM SPSS Modeler Server-Installationsprogramm beinhaltet ein Script (*auto.sh*), das das System so konfiguriert, dass der Serverdämon automatisch beim Booten gestartet wird. Wenn Sie dieses Script ausgeführt haben und dann den Server stoppen, wird der Dämon automatisch neu gestartet, wenn der Computer erneut gebootet wird. Weitere Informationen finden Sie in „Automatisches Starten und Stoppen von IBM SPSS Modeler Server“ auf Seite 79.

Grenzwerte für UNIX-Kernel

Sie müssen sicherstellen, dass Kernelgrenzwerte im System für den Betrieb von IBM SPSS Modeler Server ausreichen. Die "ulimits" für Daten, Arbeitsspeicher und Dateien sind besonders wichtig und sollten in der IBM SPSS Modeler Server-Umgebung auf unbegrenzt (unlimited) eingestellt werden. Gehen Sie dazu wie folgt vor:

1. Fügen Sie *modelersrv.sh* die folgenden Befehle hinzu:

```
ulimit -d unlimited
ulimit -m unlimited
ulimit -f unlimited
```

Setzen Sie außerdem den Stackgrenzwert auf das für Ihr System zulässige Maximum (`ulimit -s XXXX`).
Beispiel:

```
ulimit -s 64000
```

2. Starten Sie IBM SPSS Modeler Server neu.

Umgang mit nicht reagierenden Serverprozessen (UNIX-Systeme)

IBM SPSS Modeler Server-Prozesse können aus unterschiedlichen Gründen nicht mehr reagieren, z. B. wenn sie einen System- oder ODBC-Treiberaufruf starten, der blockiert wird (Aufruf wird nicht zurückgegeben oder der Vorgang dauert sehr lange). Tritt dies bei UNIX-Prozessen auf, können sie mithilfe des `kill`-Befehls in UNIX bereinigt werden (durch den IBM SPSS Modeler-Client ausgelöste Interrupts oder das Schließen des IBM SPSS Modeler-Clients sind wirkungslos). Ein `kill`-Befehl ist als Alternative zum normalen `stop`-Befehl vorgesehen und ermöglicht dem Administrator die Verwendung von *modelersrv.sh*, um den geeigneten `kill`-Befehl auf einfache Weise auszugeben.

Bei Systemen, auf denen sich nicht verwendbare Serverprozesse ("Geisterprozesse") häufen könnten, wird empfohlen, IBM SPSS Modeler Server in regelmäßigen Intervallen mit der folgenden Befehlsfolge zu stoppen und neu zu starten:

```
cd Modeler_Server-Installationsverzeichnis
./modelersrv.sh stop
./modelersrv.sh kill
```

Die IBM SPSS Modeler-Prozesse, die mithilfe des Befehls `modelersrv.sh kill` beendet werden, hinterlassen temporäre Dateien (aus dem temporären Verzeichnis), die manuell entfernt werden müssen. Temporäre Dateien bleiben unter Umständen auch dann zurück, wenn die Anwendung wegen Ressourcenauslastung abstürzt, sowie bei Benutzerinterrupts, Systemabstürzen oder aus anderen Gründen. Es wird daher empfohlen, IBM SPSS Modeler Server nicht nur regelmäßig neu zu starten, sondern alle verbleibenden Dateien aus dem temporären Verzeichnis von IBM SPSS Modeler zu entfernen.

Sobald alle Serverprozesse geschlossen und die temporären Dateien entfernt wurden, kann IBM SPSS Modeler Server sicher neu gestartet werden.

Konfigurieren von Serverprofilen

Mithilfe von Serverprofilen können Sie über eine einzelne Installation mehrere unabhängige Instanzen von SPSS Modeler Server ausführen. Einem Client gegenüber stellen sie sich als separate Server dar, die sich auf demselben Host befinden, aber auf verschiedenen Portnummern empfangsbereit sind. Die gemeinsame Nutzung einer Installation durch mehrere Instanzen bietet Administratoren den Vorteil der einfacheren Wartung. Nach der ersten Instanz lassen sich nachfolgende Instanzen mit geringerem Aufwand erstellen und löschen als mit einer vollständigen Installation und Deinstallation und auch Fixpacks müssen nur einmal angewendet werden.

Das Ausführen mehrerer Serverinstanzen auf demselben Host dient dem Zweck, jede Instanz separat konfigurieren zu können. Wenn alle Instanzen identisch sind, ergibt sich dadurch kein zusätzlicher Nutzen. Insbesondere wenn die Instanzen als *Nicht-Root*-Instanzen ausgeführt werden (damit alle Sitzungen dasselbe Benutzerkonto gemeinsam nutzen), kann für jede Instanz ein anderes Benutzerkonto verwendet werden, um zwischen den Benutzergruppen eine Datenisolation zu gewährleisten. Beispiel: Einem Benutzer, der sich bei einer Instanz A anmeldet, wird eine Sitzung zugeordnet, deren Eigner ein bestimmter Benutzer-A ist und er hat nur Zugriff auf die Dateien und Ordner dieses Benutzers. Ein Benutzer, der sich bei Instanz B anmeldet, sieht dagegen ein anderes Set von Dateien und Ordnern, die dem Benutzer-B zugänglich sind. Bei Verwendung in Verbindung mit der Gruppenkonfiguration können Anmeldungen bei einer bestimmten Instanz auf bestimmte Gruppen beschränkt werden, sodass sich Endbenutzer nur bei der Instanz (oder den Instanzen) anmelden können, die ihrer Rolle entsprechen. Siehe „Konfigurieren von Gruppen“ auf Seite 55.

In einer SPSS Modeler Server-Standardinstallation sind die Ordner `config`, `data` und `tmp` spezifisch für eine Serverinstanz. Der Ordner `config` hat den Zweck, eine private Konfiguration der Instanz zu ermöglichen, während die Ordner `data` und `tmp` der Datenisolation dienen. Jede Instanz verfügt über eine private Kopie dieser Ordner, alles andere wird gemeinsam genutzt.

Große Teile der Serverkonfiguration können weiter gemeinsam genutzt werden (z. B. Datenbankeinstellungen), sodass eine Profilkonfiguration die gemeinsame Konfiguration überschreibt. Der Server liest zuerst die Profilkonfiguration und greift danach auf die Standardkonfiguration zurück. Zu den Dateien, die am wahrscheinlichsten für ein Profil geändert werden, zählen die Dateien `options`, `groups` und `passwords`.

Weitere Informationen finden Sie in „Profilstruktur“ auf Seite 31.

Informationen zur Konfiguration eines Profils für die Verwendung von SSO finden Sie in „Konfigurieren von Single Sign-on“ auf Seite 16. Hierfür müssen Sie den Namen eines Service-Principals (SPN - Service Principal Name) registrieren, einige Konfigurationsarbeiten ausführen, wenn das Konto für den Windows-Dienst kein lokales Konto ist, und in einigen Fällen die Gruppensuche ermöglichen.

Arbeiten mit Serverprofilen

Nachfolgend werden einige gebräuchliche Anwendungsfälle für Serverprofile beschrieben. Einige dieser Verwendungen werden durch Nutzung von Scripts unterstützt (siehe „Profilscripsts“ auf Seite 33) und erfordern möglicherweise Administrator-/Rootberechtigungen.

Erstellen eines Serverprofils

Eine SPSS Modeler Server-Administratorin mit den Namen Jane erstellt mithilfe eines Scripts ein neues Profil:

- Jane muss einen eindeutigen Namen für das Profil angeben (es darf kein bereits vorhandener Profilname sein). Wenn das Verzeichnis `profiles` noch nicht vorhanden ist, wird es für Jane erstellt. Anschließend wird im Verzeichnis `profiles` ein neues Unterverzeichnis mit dem von Jane angegebenen Namen erstellt, das die Verzeichnisse `config`, `data`, `log` und `tmp` enthält.
- Bei Bedarf kann Jane auch den Namen eines vorhandenen Profils angeben, das als Vorlage verwendet werden soll, was bewirkt, dass der Inhalt des Ordners `config` im vorhandenen Profil in das neue Profil

kopiert wird. Wenn sie keine Vorlage angibt oder wenn das vorhandene Profil keine Optionsdatei enthält, obwohl es eine solche Datei enthalten sollte, wird in dem neuen Profil eine leere Optionsdatei erstellt.

- Es steht Jane außerdem offen, für das Profil eine Portnummer anzugeben, die dann als Wert der Eigenschaft `port_number` in die Optionsdatei des Profils geschrieben wird. Wenn sie keine Portnummer angibt, wird ein Wert ausgewählt und in die Optionsdatei geschrieben.
- Darüber hinaus kann Jane den Namen einer Betriebssystemgruppe angeben, die exklusiven Zugriff auf das Profil hat, wodurch in der Optionsdatei eine Gruppenkonfiguration aktiviert wird. In diesem Fall wird eine Gruppendatei erstellt, die allen außer der angegebenen Gruppe die Anmeldung verweigert.

Konfigurieren eines Serverprofils

Die Serveradministratorin Jane konfiguriert ein Profil, indem sie entweder die Profilkonfigurationsdateien manuell bearbeitet oder mithilfe von IBM SPSS Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten) eine Verbindung zum Profildienst herstellt.

Erstellen eines Windows-Diensts für ein Serverprofil

Unter Windows verwendet der Administrator ein Script, um einen Dienst für ein angegebenes Profil zu erstellen:

- Jane muss den Namen eines vorhandenen Profils angeben, woraufhin eine Dienstinstanz für dieses Profil erstellt wird. Die Befehlszeile für den Dienst enthält das Argument `profile`. Der Name des Diensts folgt einem Standardmuster und enthält den Profilnamen.
- Möglicherweise muss Jane, wenn Sie zu einem späteren Zeitpunkt den Benutzernamen und das Kennwort für den Dienst ändern will (als Nicht-Root), hierfür die Dienstadministrationskonsole verwenden.

Auch unter UNIX bestehen verschiedene Möglichkeiten zum Erstellen von "Services", die automatisch gestartet werden, wenn das System bootet. Administratoren, die Programme zur Profilerstellung anhand dieser Mechanismen erstellen wollen, sollten beachten, dass sie von IBM SPSS Modeler nicht offiziell unterstützt werden.

Verwalten von Windows-Diensten für Serverprofile

Administratoren können mithilfe eines Scripts die folgenden Aufgaben ausführen:

- Ermitteln, welche Serverprofildienste ausgeführt werden
- Einen bestimmten Dienst starten
- Alle Dienste starten
- Einen bestimmten Dienst stoppen
- Alle Dienste stoppen

Beim Starten oder Stoppen aller Dienste werden die Unterverzeichnisse des Verzeichnisses `profiles` durchsucht, um die Liste der Profile abzurufen.

Löschen des Windows-Diensts eines Serverprofils

Unter Windows können Administratoren ein Script verwenden, um einen Dienst für ein angegebenes Profil zu löschen (wenn für das Profil ein Dienst vorhanden ist). Der Name des Profils muss angegeben werden.

Entfernen eines Serverprofils

Nachdem der Profildienst gestoppt wurde, kann das Profil vom Administrator entfernt werden, indem er den zugehörigen Ordner im Verzeichnis `profiles` löscht.

Aktualisieren von SPSS Modeler Server

Wenn ein Fixpack auf SPSS Modeler Server angewendet wird, wird es auf alle Serverprofile angewendet. Unter Windows werden alle Profildienste automatisch gestoppt und erneut gestartet. Unter UNIX müssen Sie sie manuell stoppen und erneut starten.

Deinstallieren von SPSS Modeler Server

Wenn SPSS Modeler Server deinstalliert wird, werden alle Serverprofile deinstalliert. Beachten Sie, dass das Verzeichnis `profiles` und darin befindliche Profile nicht automatisch entfernt werden. Sie müssen manuell gelöscht werden. Unter Windows werden alle Profildienste automatisch deinstalliert. Unter UNIX müssen Sie sie manuell entfernen.

Installieren einer neuen Version von SPSS Modeler Server

Bei der Installation einer neuen Version von SPSS Modeler Server werden vorhandene Serverprofile nicht automatisch migriert. Die Profile müssen von einem Administrator manuell von einer Installation in die nächste kopiert (und die Konfigurationen bei Bedarf bearbeitet) werden, um die Services neu zu erstellen.

Profilstruktur

Das Profilverzeichnis

Serverprofile werden in einer Verzeichnisposition gespeichert, die vom Serveradministrator ausgewählt wird. Die Standardposition ist das Verzeichnis `profiles` im SPSS Modeler Server-Verzeichnis `[Serverinstallationspfad]\config\`, aber es wird aus folgenden Gründen ein anderes Verzeichnis zum Speichern der Profile empfohlen:

- Profile können von Knoten in einem Cluster gemeinsam genutzt werden.
- Profile können bei Upgrades beibehalten werden.
- Administratoren und anderen Benutzern, die Profile konfigurieren, muss keine Schreibberechtigung für das SPSS Modeler Server-Installationsverzeichnis erteilt werden.

Nach einer Neuinstallation von SPSS Modeler Server ist kein Verzeichnis `profiles` vorhanden. Es wird bei der Erstellung des ersten Profils erstellt.

Das Verzeichnis `profiles` enthält für jedes Profil ein Unterverzeichnis, wobei der Name des Unterverzeichnisses dem Namen des Profils entspricht. Da Verzeichnis- und Profilnamen identisch sind, kann der Profilename keine Zeichen enthalten, die in Dateinamen nicht gültig sind. Des Weiteren sollten Profilnamen keine Leerzeichen enthalten, da diese in Scripts häufig zu Problemen führen. Beachten Sie außerdem, dass Profilnamen innerhalb einer Installation eindeutig sein müssen.

Welche Profile in einer Installation vorhanden sind, lässt sich nur an den Unterverzeichnissen des Verzeichnisses `profiles` erkennen. Eine separate Liste mit Profilen gibt es darüber hinaus nicht. Die Anzahl der Profile, die für eine Installation erstellt werden kann, ist (innerhalb der durch das Hostsystem vorgegebenen Grenzen) unbegrenzt.

Im Verzeichnis `profiles` muss das Unterverzeichnis für ein Profil mindestens ein Verzeichnis `config` enthalten und in diesem Verzeichnis muss sich mindestens eine Datei namens `options.cfg` befinden, die die Profilkonfiguration definiert. Diese Datei enthält ein Subset mit allen Einstellungen aus der SPSS Modeler Server-Standarddatei `options.cfg` (im `[Serverinstallationspfad]/config`), die für das Profil erforderlich sind. Einstellungen, die in der Profilkonfiguration nicht vorhanden sind, müssen über die Datei mit den gemeinsamen Optionen im Installationsverzeichnis `config` festgelegt werden. Die Profilkonfiguration muss mindestens eine Einstellung für `port_number` enthalten, da jeder Profilservice auf einer anderen Portnummer empfangsbereit sein muss.

Die Profilkonfiguration kann andere *.cfg-Dateien enthalten, die sich normalerweise im Installationsverzeichnis config befinden. In diesem Fall werden diese anstelle der Standarddateien gelesen (nur die Optionsdatei ist kumulativ). Weitere Dateien, die höchstwahrscheinlich in eine Profilkonfiguration aufgenommen werden, sind groups und passwords. Zu den Dateien, die in einer Profilkonfiguration ignoriert werden, zählen die JVM- und SSO-Konfigurationsdateien, die profilübergreifend gemeinsam genutzt werden.

Ein Profilverzeichnis kann auch die Verzeichnisse data und tmp enthalten, die die Positionen gemeinsamer data- und tmp-Dateien überschreiben, es sei denn, in der Profilkonfiguration werden alternative Positionen angegeben.

Wenn Sie zum Erreichen einer Datenisolation Profile verwenden, achten Sie darauf, dass die Berechtigungen für die relevanten Verzeichnisse entsprechend festgelegt sind.

Profilkonfigurationsdatei

Die Position des Verzeichnisses profiles wird in einer neuen Konfigurationsdatei mit dem Namen [Serverinstallationspfad]\config\profiles.cfg angegeben. Diese hat dasselbe Format wie andere Konfigurationsdateien in demselben Verzeichnis. Der Schlüssel für das Setzen des Verzeichnisses profiles lautet profiles_directory. Beispiel:

```
profiles_directory, "C:\\SPSS\\Modeler\\profiles"
```

Zur Profilkonfiguration wird aus zwei Gründen eine separate Datei verwendet (anstatt Einstellungen zur Standardoptionsdatei hinzuzufügen):

- Die Profilkonfiguration bestimmt, wie die Optionsdateien gelesen werden, wodurch sich Schwierigkeiten beim Definieren der Profilkonfiguration in den Optionsdateien und umgekehrt ergeben.
- Die Profilkonfigurationsdatei ist so konzipiert, dass sie mithilfe von Scripts automatisch verwaltet werden kann, sodass sich Benutzer normalerweise nicht selbst damit befassen müssen. Zur Unterstützung komplexerer Szenarios kann die Datei aber problemlos auch manuell bearbeitet werden.

Neben der Position des Verzeichnisses profiles enthält die Datei profiles.cfg nur eine Portnummer als weiteren Eintrag. Beispiel:

```
profile_port, 28501
```

Dies ist die Standardportnummer für das nächste zu erstellende Profil. Bei jeder Erstellung eines Profils wird sie mithilfe eines Scripts automatisch erhöht. Die Datei profiles.cfg wird nur im Bedarfsfall erstellt und ist daher in einer Neuinstallation nicht vorhanden.

Starten eines Profils

Die ausführbare Servicedatei (modelerserver.exe) akzeptiert ein weiteres Argument (profile), das das Profil für den Service angibt:

```
modelerserver -server profile=<Profilname>
```

Es ist möglich, mehrere Services für dieselbe Installation auszuführen, wenn jeder Service ein anderes Profil verwendet. Wenn das Argument "profile" nicht angegeben wird, verwendet der Service die allgemeinen Installationsstandardwerte ohne Profilüberschreibungen.

Wenn der Service mit dem Argument profile aufgerufen wird, werden folgende Aktionen ausgeführt:

- Die Datei [Serverinstallationspfad]\config\profiles.cfg wird gelesen, um die Position des Verzeichnisses "profiles" abzurufen
- Die Datei [Verzeichnis "profiles"]\[Profilname]\config\options.cfg wird gelesen, um die Profilkonfiguration (insbesondere die Portnummer) abzurufen

Wenn einer dieser Schritte fehlschlägt, gibt der Service im Protokoll eine Fehlermeldung aus und wird gestoppt. Wenn der Service mit einem Profil aufgerufen wird und er das Profil nicht laden kann, wird er nicht ausgeführt.

Umgebungsvariablen

Der Service definiert eine Reihe weiterer Umgebungsvariablen, sodass Pfadnamen usw. angegeben werden können, ohne das aktuelle Profil zu kennen:

Tabelle 1. Umgebungsvariablen

Variable	Wert
PROFILE_NAME	Der Name des aktuellen Profils oder die leere Zeichenfolge, wenn kein Profil angegeben wurde.
MODELERPROFILE	Der vollständige Pfad zum Verzeichnis für das aktuelle Profil (z. B. <code>\$MODELERSERVER\profiles\\${PROFILE_NAME}</code>). Wenn kein Profil angegeben wurde, entspricht der Wert der Angabe <code>\$MODELERSERVER</code> .
MODELERDATA	Der vollständige Pfad zum Verzeichnis data für das aktuelle Profil (z. B. <code>\$MODELERPROFILE\data</code>). Wenn kein Profil angegeben wurde, verweist der Wert auf das Standarddatenverzeichnis <code>\$MODELERSERVER\data</code> .

Diese Umgebungsvariablen werden vom Serviceprozess gesetzt, daher sind sie nur in diesem Prozess und in von diesem erstellten untergeordneten Prozessen sichtbar. Wenn Sie diese Variablen außerhalb des Serviceprozesses setzen, werden sie ignoriert und im Prozess wie beschrieben neu definiert.

Protokollierung

Für jeden Profilservice wird erwartet, dass er über einen separaten, privaten Ordner zur Ablage der Protokolldateien verfügt. Für jedes Profil ist eine Kopie von `server_logging.log` usw. vorhanden.

In der Standardkonfiguration `log4cxx.properties` im Installationsverzeichnis `config` gibt die Umgebungsvariable `PROFILE_NAME` das Protokollverzeichnis für den Service an:

```
log4j.appender.LoggingAppender.File=${ALLUSERSPROFILE}/IBM/SPSS/Modeler
Server/17/log/${PROFILE_NAME}/server_logging.log
```

Sie können die Position der Protokolle für alle Profile ändern, indem Sie die oben aufgeführte Zeile ändern und eine oder beide der profilspezifischen Umgebungsvariablen (`PROFILE_NAME` oder `MODELERPROFILE`) aufnehmen. So kann zum Beispiel das Protokollverzeichnis innerhalb des Profilverzeichnisses verlagert werden:

```
log4j.appender.LoggingAppender.File=${MODELERPROFILE}/log
/server_logging.log
```

Alternativ können Sie die Protokollposition eines bestimmten Profils ändern, indem Sie eine Kopie der Datei `log4cxx.properties` in der Profilkonfiguration erstellen und bearbeiten.

Profilscrippts

Die in diesem Abschnitt beschriebenen Scripts sollen die Erstellung und Verwaltung von SPSS Modeler Server-Profilen erleichtern. Alle Scripts befinden sich im Verzeichnis `scripts/profiles` des Installationsverzeichnisses von SPSS Modeler Server (z. B. `C:\Programme\IBM\SPSS\ModelerServer\18\scripts\profiles`).

Allgemeines Script (für alle Plattformen)

Das folgende Script bietet Unterstützung bei der Erstellung und Verwaltung von Profilen. Dieses Script steht in mehreren Varianten mit jeweils unterschiedlichen Erweiterungen für die verschiedenen Plattformen (`.bat` für Windows und `.sh` für UNIX) zur Verfügung. Die Operation ist für alle Fällen identisch.

Erstellen eines Profils

create_profile [Optionen] <Profilname>

Erstellt ein neues Profil mit dem angegebenen Namen. Der Profilname muss für die Verwendung als Verzeichnisname auf dem Server-Host geeignet sein (da das Script ein Verzeichnis mit diesem Namen erstellt) und sollte keine Leerzeichen enthalten. Der Name muss sich von den vorhandenen Profilnamen unterscheiden.

Optionen:

-d, --profiles-directory<Profilverzeichnis>

Gibt das Verzeichnis profiles an, in dem dieses und alle nachfolgenden Profile erstellt werden sollen. Obwohl Sie diese Angabe nur für das erste Profil festlegen müssen, empfiehlt es sich, sie jedes Mal erneut festzulegen. Wenn Sie beim ersten Mal keine Angabe machen, wird eine Standardposition ausgewählt. Wenn Sie das Verzeichnis profiles bei einem nachfolgenden Aufruf ändern, wird das neue Profil an der neuen Position erstellt, vorhandene Profile jedoch werden ignoriert, wenn sie nicht explizit an die neue Position verschoben werden.

-t, --template <Profilname>

Gibt den Namen eines vorhandenen Profils an, das als Vorlage verwendet werden soll. Die Profilkonfiguration wird aus dem vorhandenen Profil in das neue Profil kopiert, wobei nur die Portnummer geändert wird.

-p, --port-number <Portnummer>

Gibt die Portnummer für den Profildienst an. Die Portnummer muss für dieses Profil eindeutig sein. Wenn Sie keine Portnummer angeben, wird ein Standardwert ausgewählt.

-g, --group-name <Gruppenname>

Gibt den Namen einer Betriebssystemgruppe an, die exklusiven Zugriff auf dieses Profil haben soll. Das Profil ist so konfiguriert, dass nur Mitglieder dieser Gruppe Anmeldezugriff erhalten.

Die Dateisystemberechtigungen werden nicht geändert, sodass Sie diese Aktion separat durchführen müssen.

Beispiele:

```
scripts\profiles\create_profile.bat -d C:\Modeler\Profiles comet
```

Erstellt im Verzeichnis C:\Modeler\Profiles ein neues Profil mit dem Namen comet. Das Profil ist auf einer Standardportnummer empfangsbereit. Wenn Sie die Portnummer ermitteln wollen, öffnen Sie die Datei options.cfg, die für das Profil generiert wurde (in diesem Beispiel: C:\Modeler\Profiles\comet\config\options.cfg).

```
scripts\profiles\create_profile.bat --template comet --group-name "Meteor Users" --port-number 28510 meteor
```

Erstellt ein neues Profil mit dem Namen meteor (aus dem vorherigen Befehl) im Verzeichnis C:\Modeler\Profiles. Das Profil ist auf dem Port 28510 empfangsbereit und Anmeldezugriff wird nur Mitgliedern der Gruppe Meteor Users erteilt. Alle anderen Konfigurationsoptionen werden aus dem vorhandenen Profil comet kopiert.

Windows-Scripts

Diese Scripts bieten Unterstützung bei der Erstellung und Verwaltung von Windows-Diensten für SPSS Modeler Server-Profile. Diese führen die angeforderten Operationen mithilfe des Windows-Programms zur Steuerung von Diensten (SC.EXE) aus und die Scriptausgabe wird, sofern nicht anders angegeben, von SC.EXE bereitgestellt. Zum Ausführen der meisten dieser Aufgaben müssen Sie auf dem lokalen System über Administratorberechtigungen verfügen.

Weitere Informationen zu SC.EXE finden Sie in der TechNet-Dokumentation von Microsoft.

Erstellen eines Windows-Diensts für ein Profil

create_windows_service [Optionen] <Profilname>

Erstellt den Windows-Dienst für ein angegebenes Profil. Zum Erstellen eines Diensts benötigen Sie Administratorberechtigungen. Verwenden Sie die Dienstverwaltungskonsole, um nach der Erstellung eines Diensts weitere Einstellungen für diesen festzulegen (z. B. zum Festlegen der Kontodetails für die Dienstanmeldung).

Optionen:

-u, --service-user <Kontoname>

Gibt das für die Dienstanmeldung verwendete Konto an (passim). Dies kann ein lokales Benutzerkonto, ein Domänenbenutzerkonto oder der Name des lokalen Computers sein (der für lokale Systemkonto steht). Der Standardwert ist das lokale Systemkonto. Wenn Sie ein anderes Konto als das lokale Systemkonto angeben, müssen Sie über die Dienstverwaltungskonsole das Kennwort für das Konto festlegen, bevor der Dienst gestartet wird.

-s, --register-spn

Registriert einen SPN (Service Principal Name - Name des Service-Principals) für den Dienst, damit Clients über Kerberos-SSO eine Verbindung herstellen können. In diesem Fall müssen Sie das Anmeldekonto für den Dienst (-u) angeben, damit der SPN für dieses Konto registriert werden kann. Um diese Option verwenden zu können, benötigen Sie Domänenadministratorberechtigungen (oder es muss Ihnen die Berechtigung zum Registrieren eines SPN erteilt worden sein).

-H, --service-host <Hostname>

Gibt den bei der Bildung des SPN zu verwendenden Hostnamen an. Dies muss der Hostname sein, über den Clients eine Verbindung herstellen. Dieser Name muss mit einem Domännennamen qualifiziert sein, der dem Kerberos-Realm zugeordnet ist (in einer einfachen Active Directory-Konfiguration sind Domännennamen und Kerberos-Realm ein und dasselbe).

Beispiele:

scripts\profiles\create_windows_service.bat comet

Erstellt einen Windows-Dienst für das Profil comet. Eigner des Diensts ist das lokale Systemkonto und es wird erwartet, dass sich Clients mit einem Benutzernamen und einem Kennwort anmelden.

scripts\profiles\create_windows_service.bat -s -H modelerserver.mycompany.com -u MYCOMPANY\ProjectMeteor meteor

Erstellt einen Windows-Dienst für das Profil meteor. Eigner des Diensts ist das Domänenkonto ProjectMeteor. Clients können sich über SSO anmelden. Der Dienst startet erst, wenn Sie über die Dienstverwaltungskonsole das Kennwort für das Konto ProjectMeteor festgelegt haben. Dem Konto wird automatisch das Recht erteilt, sich als Dienst anzumelden.

Löschen eines Windows-Diensts für ein Profil

delete_windows_service [Optionen] <Profilnamen...>

Löscht die Windows-Dienste für die angegebenen Profile. Zum Löschen eines Diensts benötigen Sie Administratorberechtigungen.

Optionen:

-s, --summary

Listet die Namen der Dienste auf, die gelöscht wurden. Dienste, die nicht vorhanden sind oder nicht gelöscht werden können, werden nicht aufgelistet. Wird diese Option nicht angegeben, wird der Löschstaus aller angegebenen Dienste aufgelistet.

-a, --all

Löscht die Dienste für alle Profile.

Beispiele:

scripts\profiles\delete_windows_service.bat comet

Löscht die Windows-Dienste für das Profil comet.

scripts\profiles\delete_windows_service.bat --all

Löscht die Windows-Dienste für alle Profile.

Starten eines Windows-Diensts für ein Profil

start_windows_service [Optionen] <Profilnamen...>

Startet die Windows-Dienste für die angegebenen Profile. Zum Starten eines Diensts benötigen Sie Administratorberechtigungen.

Optionen:

-s, --summary

Listet die Namen der Dienste auf, die gestartet wurden. Dienste, die bereits ausgeführt werden oder nicht gestartet werden können, werden nicht aufgelistet. Wird diese Option nicht angegeben, wird der Status aller aufgeführten Dienste aufgelistet.

-a, --all

Startet die Dienste für alle Profile.

Beispiele:

scripts\profiles\start_windows_service.bat -s comet meteor

Versucht, die Windows-Dienste für die Profile comet und meteor zu starten, und listet die Namen derjenigen auf, die erfolgreich gestartet wurden.

Stoppen eines Windows-Diensts für ein Profil

stop_windows_service [Optionen] <Profilnamen...>

Stoppt die Windows-Dienste für die angegebenen Profile. Zum Stoppen eines Diensts benötigen Sie Administratorberechtigungen.

Optionen:

-s, --summary

Listet die Namen der Dienste auf, die gestoppt wurden. Dienste, die bereits gestoppt sind oder nicht gestoppt werden können, werden nicht aufgelistet. Wird diese Option nicht angegeben, wird der Status aller aufgeführten Dienste aufgelistet.

-a, --all

Stoppt die Dienste für alle Profile.

Beispiele:

scripts\profiles\stop_windows_service.bat -a -s

Versucht, die Windows-Dienste für alle Profile zu stoppen, und gibt die Namen derjenigen aus, die erfolgreich gestoppt wurden. Das Set aller Profile wird über das Verzeichnis profiles abgerufen.

Abfragen des Status eines Windows-Diensts für ein Profil

query_windows_service [Optionen] <Profilnamen...>

Zeigt den Status der Windows-Dienste für die angegebenen Profile an. Zum Abfragen eines Diensts benötigen Sie keine Administratorberechtigungen.

Optionen:

-s, --summary

Listet nur die Namen der Dienste und deren aktuellen Status auf (RUNNING, STOPPED usw.). Wenn ein Dienst nicht abgefragt werden kann (z. B. wenn er nicht vorhanden ist), wird für ihn UNKNOWN als Status gemeldet. Wenn diese Option nicht angegeben wird, wird der vollständige Status von allen aufgeführten Diensten aufgelistet.

-a, --all

Fragt den Dienststatus für alle Profile ab.

Beispiele:

```
scripts\profiles\query_windows_service.bat -a
```

Meldet den vollständigen Dienststatus für alle Profile.

UNIX-Script

Das vorhandene UNIX-Script für die Verwaltung des SPSS Modeler Server-Service akzeptiert jetzt ein weiteres Argument `profile`, sodass SPSS Modeler Server-Profilservices unabhängig voneinander verwaltet werden können.

modelersrv.sh [Optionen] {start|stop|kill|list}

Verwaltet den SPSS Modeler Server-Hauptservice. Weitere Informationen finden Sie in Kapitel 4, „IBM SPSS Modeler-Verwaltung“, auf Seite 27.

Optionen:

-p, --profile <Profilname>

Verwaltet die Serviceinstanz für das angegebene Profil. Bei Verwendung dieses Arguments wird der angegebene Befehl nur auf die Instanz für das angegebene Profil angewendet. Wird dieses Argument nicht angegeben, startet der Befehl `start` nur die Standardinstanz (ein Service ohne Profil), die Befehle `stop`, `kill` und `list` dagegen werden auf alle aktiven Instanzen angewendet.

Beispiele:

```
./modelersrv.sh --profile comet start
```

Startet den Service für das Profil `comet`.

```
./modelersrv.sh --profile meteor start
```

Startet den Service für das Profil `meteor`.

```
./modelersrv.sh list
```

Listet die Prozesse für alle aktiven Services auf.

```
./modelersrv.sh --profile comet stop
```

Stoppt den Service für das Profil `comet`.

```
./modelersrv.sh stop
```

Stoppt alle aktiven Services.

Gegenwärtig wird unter UNIX keine Methode für ein automatisches Starten der SPSS Modeler Server-Profilservices unterstützt. Das Standardscript `auto.sh` kann verwendet werden, um das System so zu konfigurieren, dass es den SPSS Modeler Server-Hauptservice mit dem Betriebssystem startet und stoppt, dies gilt jedoch nur für den Standardservice - nicht für Profilservices.

Verwaltung

IBM SPSS Modeler Server weist eine Reihe von konfigurierbaren Optionen auf, die sein Verhalten steuern. Sie können diese Optionen auf zwei verschiedene Weisen festlegen:

- Verwenden Sie die IBM SPSS Modeler Administration Console, die aktuellen IBM SPSS Modeler-Kunden kostenlos zur Verfügung steht. Weitere Informationen finden Sie in „IBM SPSS Modeler Server-Verwaltung“.
- Verwenden Sie die Textdatei `options.cfg`, die sich im Verzeichnis `[Serverinstallationspfad]/config` befindet. Weitere Informationen finden Sie in „Verwenden der Datei `options.cfg`“ auf Seite 48.

Es wird empfohlen, IBM SPSS Deployment Manager zu installieren und die dort enthaltene IBM SPSS Modeler Administration Console als Verwaltungstool zu installieren und zu verwenden und nicht die Datei `options.cfg` zu bearbeiten. Für die Bearbeitung der Datei ist Zugriff auf das Dateisystem von IBM SPSS Modeler Server erforderlich, mit IBM SPSS Modeler Administration Console dagegen können Sie jede Person mit einem Benutzerkonto zur Anpassung dieser Optionen berechtigen. Außerdem bietet IBM SPSS Modeler Administration Console weitere Informationen zu den Serverprozessen, sodass Sie Nutzung und Leistung überwachen können. Anders als bei der Bearbeitung der Konfigurationsdatei können die meisten Konfigurationsoptionen ohne Neustart von IBM SPSS Modeler Server geändert werden.

Weitere Informationen zur Verwendung von IBM SPSS Modeler Administration Console und der Datei `options.cfg` finden Sie in folgenden Abschnitten.

IBM SPSS Modeler Server-Verwaltung

Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten) bietet eine Konsolenbenutzerschnittstelle zur Überwachung und Konfiguration der SPSS Modeler Server-Installationen und steht aktuellen SPSS Modeler Server-Kunden kostenlos zur Verfügung. Die Anwendung kann nur unter Windows installiert werden. Der von ihr verwaltete Server kann jedoch auf einer beliebigen unterstützten Plattform installiert sein.

Viele der über Modeler Administration Console verfügbaren Optionen können auch in der Datei `options.cfg` angegeben werden, die sich im Installationsverzeichnis von SPSS Modeler Server unter `/config` befindet. Modeler Administration Console bietet jedoch eine für die gemeinsame Verwendung freigegebene grafische Benutzerschnittstelle, mit der Sie eine Verbindung mit mehreren Servern herstellen und diese konfigurieren und überwachen können.

Starten von Modeler Administration Console

Wählen Sie im Startmenü von Windows [Alle] **Programme**, anschließend **IBM SPSS Collaboration and Deployment Services** und dann **Deployment Manager** aus.

Wenn Sie die Anwendung zum ersten Mal ausführen, sehen Sie leere Serveradministrations- und Eigenschaftsbereiche (außer Sie haben Deployment Manager bereits mit einer schon eingerichteten IBM SPSS Collaboration and Deployment Services-Serververbindung installiert). Nach der Konfiguration von Modeler Administration Console wird im Serveradministratorbereich auf der linken Seite ein Knoten für jeden SPSS Modeler Server angezeigt, den Sie verwalten können. Im rechten Bereich werden die Konfigurationsoptionen für den ausgewählten Server angezeigt. Sie müssen zuerst für jeden Server, den Sie verwalten möchten, eine Verbindung einrichten.

Neustart des Web-Service

Sobald Sie in der Administrationskonsole Änderungen an IBM SPSS Modeler Server oder IBM SPSS Modeler Text Analytics Server vornehmen, müssen Sie den Web-Service neu starten.

So starten Sie den Web-Service unter Microsoft Windows neu:

1. Wählen Sie auf dem Computer, auf dem Sie IBM SPSS Modeler installiert haben, in der Systemsteuerung unter "Verwaltung" die Option **Dienste** aus.
2. Suchen Sie in der Liste nach **IBM SPSS Modeler Server** und starten Sie es neu.
3. Klicken Sie auf **OK**, um das Dialogfeld zu schließen.

So starten Sie den Web-Service unter UNIX neu:

Unter UNIX müssen Sie IBM SPSS Modeler Server neu starten, indem Sie das Script **modelersrv.sh** im Installationsverzeichnis von IBM SPSS Modeler Server ausführen.

1. Wechseln Sie zum IBM SPSS Modeler Server-Installationsverzeichnis. Geben Sie beispielsweise an einer UNIX-Eingabeaufforderung Folgendes ein:
`cd /usr/<modelersrv>`. Dabei steht `modelersrv` für das Installationsverzeichnis von IBM SPSS Modeler Server.
2. Um den Server zu stoppen, geben Sie an der Eingabeaufforderung Folgendes ein:
`./modelersrv.sh stop`
3. Um den Server neu zu starten, geben Sie an der Eingabeaufforderung Folgendes ein:
`./modelersrv.sh start`

Konfigurieren des Zugriffs mithilfe von Modeler Administration Console

Der Administratorzugriff auf SPSS Modeler Server über Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten) wird über die Zeile `administrators` in der Datei `options.cfg` gesteuert. Diese Datei befindet sich im Installationsverzeichnis von SPSS Modeler Server unter `/config`. Diese Zeile ist standardmäßig auf Kommentar gesetzt. Sie müssen diese Zeile also bearbeiten, um bestimmten Personen Zugriff zu ermöglichen, oder Sie verwenden `*`, um allen Benutzern Zugriff zu ermöglichen (siehe folgende Beispiele):

```
administrators, "*"
administrators, "jsmith,mjones,achavez"
```

- Die Zeile muss mit `administrators`, beginnen und die Einträge müssen in Anführungszeichen gesetzt werden. Bei den Einträgen wird zwischen Groß- und Kleinschreibung unterschieden.
- Trennen Sie mehrere Benutzer-IDs mit Kommas.
- Verwenden Sie im Falle von Windows-Konten keine Domänennamen.
- Verwenden Sie den Stern (*) umsichtig. Mit dieser Einstellung kann sich jede Person mit einem gültigen Benutzerkonto für IBM SPSS Modeler Server (meistens jede beliebige Person im Netz) anmelden und die Konfigurationsoptionen ändern.

Konfigurieren des Zugriffs mithilfe der Benutzerzugriffssteuerung

Wenn Sie mit Modeler Administration Console Aktualisierungen an einer SPSS Modeler Server-Konfiguration vornehmen wollen, die auf einem Windows-Computer installiert ist, auf dem die Benutzerzugriffssteuerung (UAC) aktiviert ist, müssen Sie Lese-, Schreib- und Ausführungsberechtigungen für das Verzeichnis `config` und die Datei `options.cfg` haben. Diese NTFS-Berechtigungen müssen wegen der Art, wie UAC- und NTFS-Berechtigungen interagieren, auf der jeweiligen Benutzerebene und nicht auf der Gruppenebene definiert sein.

Modeler Administration Console ist im Lieferumfang von IBM SPSS Deployment Manager enthalten.

SPSS Modeler Server-Verbindungen

Sie müssen für jede SPSS Modeler Server-Instanz in Ihrem Netz, die Sie verwalten möchten, eine Verbindung einrichten. Anschließend müssen Sie sich bei jedem Server anmelden. Die Serververbindung wird für Modeler Administration Console-Sitzungen in IBM SPSS Deployment Manager gespeichert, nicht jedoch die Anmeldeberechtigungs-nachweise. Sie müssen sich bei jedem Start von IBM SPSS Deployment Manager erneut anmelden.

So richten Sie eine Serververbindung ein:

1. Stellen Sie sicher, dass der IBM SPSS Modeler Server-Service gestartet wurde.
2. Wählen Sie im Menü "Datei" die Option **Neu** und dann **Verwaltete Serververbindung** aus.
3. Geben Sie auf der ersten Seite des Assistenten einen Namen für die Serververbindung ein. Der Name ist zu Ihrer eigenen Verwendung und sollte beschreibend sein, zum Beispiel *Produktionsserver*. Stellen Sie sicher, dass der Typ auf **Verwaltet IBM SPSS Modeler Server** gesetzt ist und klicken Sie dann auf **Weiter**.
4. Auf der zweiten Seite geben Sie den Hostnamen oder die IP-Adresse des Servers ein. Wenn Sie die Standardeinstellung für den Port geändert haben, geben Sie die Portnummer ein. Klicken Sie auf **Fertigstellen**. Die neue Serververbindung wird im Serveradministratorbereich angezeigt.

Um Verwaltungsaufgaben durchzuführen, müssen Sie sich anmelden.

So melden Sie sich beim Server an:

1. Doppelklicken Sie im Serveradministratorbereich auf den Server, bei dem Sie sich anmelden möchten.
2. Geben Sie im Dialogfeld "Anmelden" Ihre Berechtigungsnachweise ein. (Verwenden Sie Ihr Benutzerkonto für den Server-Host.) Klicken Sie auf **OK**.

Wenn die Anmeldung fehlschlägt und die Nachricht **Administratorrechte auf dem Server nicht verfügbar** angezeigt wird, ist die wahrscheinlichste Ursache, dass der Administratorzugriff nicht korrekt konfiguriert wurde. Weitere Informationen finden Sie in „Konfigurieren des Zugriffs mithilfe von Modeler Administration Console“ auf Seite 39.

Wenn die Anmeldung fehlschlägt und die Fehlermeldung **Verbindung zum Server '<Server>' fehlgeschlagen** angezeigt wird, überprüfen Sie, ob die Benutzer-ID und das Kennwort korrekt sind und ob der IBM SPSS Modeler Server-Service ausgeführt wird. Rufen Sie z. B. unter Windows "Systemsteuerung > Verwaltung > Dienste" auf und überprüfen Sie den Eintrag für IBM SPSS Modeler Server. Wenn die Spalte nicht **Gestartet** anzeigt, wählen Sie diese Zeile auf dem Bildschirm aus und klicken Sie auf **Starten**. Versuchen Sie anschließend erneut, sich anzumelden.

Nach der Anmeldung bei IBM SPSS Modeler Server werden unter dem Servernamen die beiden Optionen Konfiguration und Überwachung angezeigt. Doppelklicken Sie auf eine dieser Optionen.

SPSS Modeler Server-Konfiguration

Im Konfigurationsbereich werden Konfigurationsoptionen für SPSS Modeler Server angezeigt. In diesem Bereich können Sie die Optionen wunschgemäß ändern. Klicken Sie in der Symbolleiste auf **Speichern**, um die Änderungen zu speichern. Beachten Sie, dass beim Ändern von mit einem Stern (*) markierten Optionen ein Serverneustart erforderlich ist, damit die Änderungen wirksam werden.

Die Optionen werden in den nachfolgenden Abschnitten beschrieben und die entsprechende Zeile in der Datei `options.cfg` wird für jede Option in Klammern angegeben. Optionen, die nur in `options.cfg` sichtbar sind, werden am Ende dieses Abschnitts beschrieben.

Anmerkung: Wenn ein Benutzer ohne Rootberechtigung diese Optionen ändern will, ist Schreibberechtigung für das Verzeichnis `config` von SPSS Modeler Server erforderlich.

Verbindungen/Sitzungen

Maximale Anzahl an Verbindungen. (`max_sessions`) Maximale Anzahl gleichzeitiger Serversitzungen. Der Wert -1 bedeutet, dass keinerlei Beschränkung vorliegt.

Portnummer. (`port_number`) Die Portnummer, die SPSS Modeler Server überwacht. Ändern Sie die Portnummer, wenn der Standardport bereits von einer anderen Anwendung verwendet wird. Endbenutzer müssen die Portnummer kennen, um mit SPSS Modeler Server arbeiten zu können.

Analytic Server-Verbindung

SSL für Analytic Server aktivieren (as_ssl_enabled) Geben Sie Y an, wenn die Kommunikation zwischen Analytic Server und SPSS Modeler verschlüsselt durchgeführt werden soll. Geben Sie andernfalls N an.

Host (as_host) Die IP-Adresse von Analytic Server.

Portnummer (as_port) Die Portnummer von Analytic Server.

Kontextstammverzeichnis (as_context_root) Das Kontextstammverzeichnis von Analytic Server.

Nutzer (as_tenant) Der Nutzer, dem die SPSS Modeler Server-Installation als Mitglied angehört.

Eingabeaufforderung für Kennwort (as_prompt_for_password) Geben Sie N an, wenn SPSS Modeler Server mit demselben Authentifizierungssystem für Benutzer und Kennwörter konfiguriert wird wie das System, das auch für Analytic Server verwendet wird, beispielsweise wenn Sie die Kerberos-Authentifizierung verwenden. Geben Sie andernfalls Y an.

Anmerkung: Falls Sie Kerberos-SSO verwenden möchten, müssen Sie Zusatzoptionen in der Datei options.cfg festlegen. Weitere Informationen finden Sie in "In options.cfg sichtbare Optionen" weiter unten in diesem Kapitel.

Zugriff auf Datendateien

Zugriff ausschließlich auf Datendateipfad beschränken (data_files_restricted) Ist dieser Wert auf **Ja** eingestellt, wird der Zugriff auf Datendateien im standardmäßigen Datenverzeichnis und den im **Datendateipfad** aufgelisteten Dateien eingeschränkt.

Datendateipfad (data_file_path) Eine Liste zusätzlicher Verzeichnisse, für die Clients Lese- und Schreibrechte besitzen. Diese Option wird ignoriert, außer die Option **Zugriff auf Datendateipfad beschränken** ist aktiviert. Beachten Sie, dass bei allen Pfadangaben normale Schrägstriche (/) zu verwenden sind. Trennen Sie in Windows mehrere Verzeichnisse durch Semikolons (z. B. [Serverinstallationspfad]/data;c:/data;c:/temp). Unter Linux und UNIX müssen Sie Doppelpunkte (:) statt Semikolons verwenden. Der Datendateipfad muss alle Pfade enthalten, die durch den nachfolgend beschriebenen Parameter temp_directory angegeben sind.

Zugriff ausschließlich auf Programmdateipfad beschränken (program_files_restricted) Ist dieser Wert auf **Ja** eingestellt, wird der Programmdateizugriff auf das Standardverzeichnis bin und die im **Programmdateipfad** aufgelisteten Verzeichnisse eingeschränkt. Ab Release 17 ist die einzige Programmdatei, auf die der Zugriff beschränkt ist, die ausführbare Datei von Python (siehe **Pfad zur ausführbaren Datei von Python** unten).

Programmdateipfad (program_file_path) Eine Liste zusätzlicher Verzeichnisse, in denen Clients Programme ausführen dürfen. Diese Option wird ignoriert, außer die Option **Zugriff ausschließlich auf Programmdateipfad beschränken** ist aktiviert. Beachten Sie, dass bei allen Pfadangaben normale Schrägstriche (/) zu verwenden sind. Trennen Sie mehrere Verzeichnisse durch Semikolons.

Maximale Dateigröße (max_file_size) Maximale Größe (in Byte) von temporären und exportierten Datendateien, die während der Streamausführung erstellt wurden (gilt nicht für SAS- und SPSS Statistics-Datendateien). Der Wert -1 bedeutet, dass keinerlei Beschränkung vorliegt.

Temporäres Verzeichnis (temp_directory) Das zum Speichern temporärer Datendateien (Cachedateien) verwendete Verzeichnis. Idealerweise sollte sich dieses Verzeichnis auf einem separaten Hochgeschwindigkeitslaufwerk oder Controller befinden, da die Zugriffszeit auf dieses Verzeichnis erhebliche Auswirkungen auf die Leistung haben kann. Es können mehrere temporäre Verzeichnisse (durch Komma getrennt) angegeben werden. Diese sollten sich auf unterschiedlichen Datenträgern befinden; das erste Verzeichnis wird am häufigsten verwendet, die zusätzlichen Verzeichnisse dienen zur Speicherung temporärer Arbeitsdateien, wenn bestimmte Datenaufbereitungsoperationen (beispielsweise die Sortierung)

bei der Ausführung Parallelität verwenden. Dadurch, dass die einzelnen Ausführungsthreads gesonderte Datenträger für die temporäre Speicherung verwenden können, ist eine Leistungssteigerung möglich. Verwenden Sie bei allen Pfadangaben normale Schrägstriche (/).

Anmerkung:

- In diesem Verzeichnis werden während des Starts von SPSS Modeler Server temporäre Dateien generiert. Stellen Sie sicher, dass Sie über die notwendigen Zugriffsrechte für dieses Verzeichnis verfügen (beispielsweise wenn es sich beim temporären Verzeichnis um einen freigegebenen Netzordner handelt), andernfalls kann SPSS Modeler Server nicht gestartet werden.
- Die Einstellung `temp_directory` gilt nicht, wenn Evaluierungsstreams über Jobs von IBM SPSS Collaboration and Deployment Services ausgeführt werden. Wenn Sie einen solchen Job ausführen, wird eine temporäre Datei erstellt. Die Datei wird standardmäßig im Installationsverzeichnis des IBM SPSS Modeler-Servers gespeichert. Sie können den Standarddateiordner, in dem die temporären Dateien gespeichert werden, beim Erstellen der IBM SPSS Modeler-Serververbindung in IBM SPSS Modeler ändern.

Pfad zur ausführbaren Python-Datei (`python_exe_path`) Vollständiger Pfad zur ausführbaren Python-Datei einschließlich des Namens der exe-Datei. Wenn der Zugriff auf Programmdateien beschränkt ist, müssen Sie das Verzeichnis, das die ausführbare Datei von Python enthält, dem Pfad für die Programmdateien hinzufügen (siehe **Zugriff ausschließlich auf Programmdateipfad beschränken** oben).

Leistung und Optimierung

Streamneuerstellung. (`stream_rewriting_enabled`) Ermöglicht dem Server die Optimierung von Streams durch Neuschreiben. Der Server kann beispielsweise die Datenreduktionsvorgänge näher an den Quellenknoten bringen, um die Größe des Datensets so früh wie möglich zu verringern. Die Inaktivierung dieser Option wird normalerweise nur empfohlen, wenn durch die Optimierung Fehler oder andere unerwartete Ergebnisse verursacht werden. Diese Einstellung überschreibt die entsprechende Optimierungseinstellung am Client. Wenn diese Einstellung auf dem Server inaktiviert ist, kann der Client sie nicht aktivieren. Ist sie auf dem Server aktiviert, kann der Client sie jedoch inaktivieren.

Parallelismus. (`max_parallelism`) Gibt die Anzahl an parallelen Arbeitsthreads an, die SPSS Modeler beim Ausführen eines Streams verwenden darf. Wenn Sie diesen Wert auf 0 oder eine negative Zahl setzen, wird IBM SPSS Modeler angewiesen, die Anzahl an Threads der Anzahl an verfügbaren Prozessoren auf dem Computer anzupassen. Der Standardwert für diese Option ist -1. Um die parallele Verarbeitung (für Computer mit mehreren Prozessoren) zu inaktivieren, setzen Sie diese Option auf den Wert 1. Um eine eingeschränkte parallele Verarbeitung zu gestatten, setzen Sie den Wert auf eine Zahl, die kleiner als die auf Ihrem Computer verfügbare Anzahl an Prozessoren ist. Beachten Sie, dass ein Hyperthreaded- oder Dual-Core-Prozessor als zwei Prozessoren betrachtet wird.

Puffergröße (Byte). (`io_buffer_size`) Die vom Server zum Client übertragenen Datendateien werden durch einen Puffer mit dieser Größe geleitet.

Cachekomprimierung. (`cache_compression`) Ein ganzzahliger Wert von 0 bis 9, der die Komprimierung des Cache und von anderen Dateien im temporären Verzeichnis des Servers steuert. Durch die Komprimierung reduziert sich die verwendete Speichermenge, was bei begrenzter Speicherkapazität von Bedeutung ist. Durch die Komprimierung erhöht sich die Prozessorzeit. Dies wird jedoch fast immer durch die verringerte Datenträgerzugriffszeit ausgeglichen. Beachten Sie, dass eine Komprimierung nur bei bestimmten Caches möglich ist, nämlich bei solchen, auf die der Zugriff sequenziell erfolgt. Diese Option steht nicht bei Caches mit Zufallszugriff zur Verfügung, beispielsweise denjenigen, die von den Netztrainingsalgorithmen verwendet werden. Mit dem Wert 0 wird die Komprimierung vollständig inaktiviert. Die Werte von 1 aufwärts bieten einen steigenden Komprimierungsgrad, jedoch jeweils mit entsprechend erhöhter Zugriffszeit. Der Standardwert ist 1; höhere Werte können erforderlich sein, wenn der Datenträgerspeicher besonders knapp ist.

Multiplikator für Arbeitsspeicherverwendung. (`memory_usage`) Legt fest, welcher Anteil an physischem Speicher für die Sortierung und andere speicherinterne Cachevorgänge zugeordnet wird. Die Standardeinstellung ist 100. Dies entspricht etwa 10 % des physischen Speichers. Wenn freier Arbeitsspeicher ver-

fürbar ist, können Sie diesen Wert erhöhen, um die Sortierleistung zu verbessern. Sie sollten jedoch bei der Erhöhung vorsichtig vorgehen, um übermäßig große Auslagerungsdateien zu vermeiden.

Obergrenze für Modellierungsarbeitsspeicher (Prozent). (`modelling_memory_limit_percentage`) Steuert, welcher Anteil an physischem Speicher für das Training von Kohonen- und K-Means-Modellen zugeordnet wird. Der Standardwert ist 25 %. Wenn freier Arbeitsspeicher verfügbar ist, können Sie diesen Wert erhöhen, um die Trainingsleistung zu verbessern. Sie sollten jedoch bei der Erhöhung vorsichtig vorgehen, damit Sie den Wert nicht so hoch setzen, dass die Auslagerungsdatei übermäßig groß wird, wenn Daten auf den Datenträger geschrieben werden müssen.

Umgehung von Modellierungsarbeitsspeicher erlauben. (`allow_modelling_memory_override`) Aktiviert bzw. inaktiviert die Option **Auf Geschwindigkeit optimiert** in bestimmten Modellierungsknoten. Diese Option ist standardmäßig aktiviert. Mit dieser Option kann der Modellierungsalgorithmus den gesamten Arbeitsspeicher beanspruchen (die Option für die Prozentgrenze wird übergangen). Diese Option sollte inaktiviert werden, wenn Arbeitsspeicherressourcen auf dem Server-Computer für die gemeinsame Verwendung freigegeben sein müssen.

Maximale und minimale Nummer für Server-Port. (`max_server_port` und `min_server_port`) Gibt den Bereich der Portnummern an, die für die zusätzlichen Socketverbindungen zwischen Client und Server verwendet werden können, die für interaktive Modelle und die Streamausführung erforderlich sind. Dafür muss der Server einen anderen Port überwachen; wenn der Bereich nicht eingeschränkt wird, kann dies zu Problemen für Benutzer führen, die Systeme mit Firewalls verwenden. Der Standardwert für beide Optionen lautet -1, was bedeutet, dass es keine Einschränkungen gibt. Um also beispielsweise festzulegen, dass der Server Port 8000 oder höher überwachen soll, setzen Sie `min_server_port` auf 8000 und `max_server_port` auf 1.

Beachten Sie, dass Sie zusätzliche Ports über den Haupt-Server-Port öffnen müssen, um einen Stream zu öffnen oder auszuführen, und entsprechend mehr Ports, wenn Sie mehrere Streams gleichzeitig öffnen oder ausführen möchten. Dies dient zur Erfassung des Feedbacks aus der Streamausführung.

Standardmäßig verwendet IBM SPSS Modeler jeden beliebigen offenen Port, der verfügbar ist. Wenn kein offener Port gefunden wird (z. B. wenn alle durch eine Firewall geschlossen sind) wird bei der Ausführung des Streams eine Fehlernachricht angezeigt. Zur Konfiguration des Portbereichs benötigt IBM SPSS Modeler (zusätzlich zum Hauptport des Servers) zwei verfügbare offene Ports pro gleichzeitigem Stream sowie drei zusätzliche Ports für jede ODBC-Verbindung aus allen angeschlossenen Clients (zwei Ports für die ODBC-Verbindung für die Dauer dieser Verbindung und ein zusätzlicher temporärer Port zur Authentifizierung).

Anmerkung: Eine ODBC-Verbindung ist ein Eintrag in der Datenbankverbindungsliste und kann von mehreren Datenbankknoten gemeinsam genutzt werden, die mit derselben Datenbankverbindung angege- ben wurden.

Anmerkung: Es ist möglich, dass die Authentifizierungsports gemeinsam genutzt werden, wenn die Ver- bindungen zu unterschiedlichen Zeiten hergestellt werden.

Anmerkung: Es ist ein bewährtes Verfahren, dieselben Ports für die Kommunikation mit IBM SPSS Colla- boration and Deployment Services und SPSS Modeler Client zu verwenden. Diese können als `max_server_port` und `min_server_port` festgelegt werden.

Anmerkung: Wenn Sie diese Parameter ändern, müssen Sie SPSS Modeler Server erneut starten, damit die Änderungen wirksam werden.

Array-Abbrufoptimierung. (`sql_row_array_size`) Legt fest, wie SPSS Modeler Server Daten von der ODBC-Datenquelle abrufen. Der Standardwert ist 1. Dabei wird jeweils eine einzelne Zeile abgerufen. Bei einer Erhöhung dieses Werts liest der Server die Informationen in größeren Einheiten, indem er jeweils

die angegebene Anzahl an Zeilen in ein Array abrufen. Bei einigen Kombinationen aus Betriebssystem/ Datenbank kann dies zu Verbesserungen bei der Leistungsfähigkeit von SELECT-Anweisungen führen.

SQL

Maximale SQL-Zeichenfolgenlänge. (`max_sql_string_length`) Bei Zeichenfolgen, die mit SQL aus der Datenbank importiert werden, ist dies die maximale Anzahl an Zeichen, die garantiert erfolgreich weitergeleitet werden. Je nach Betriebssystem werden Zeichenfolgewerte, die länger als diese Vorgabe sind, möglicherweise ohne Warnung rechts abgeschnitten. Der zulässige Bereich liegt zwischen 1 und 65.535 Zeichen. Diese Eigenschaft wird auch auf den Datenbankexportknoten angewendet.

Anmerkung: Der Standardwert für diesen Parameter ist 2048. Wenn der Text, den Sie analysieren, länger als 2048 Zeichen ist (dies kann z. B. vorkommen, wenn der SPSS Modeler Text Analytics-Web-Feed-Knoten verwendet wird), empfehlen wir das Erhöhen dieses Werts, wenn im nativen Modus gearbeitet wird, da ansonsten Ihre Ergebnisse abgeschnitten werden können. Wenn Sie eine Datenbank und UDFs (UDF - User-Defined Function, benutzerdefinierte Funktion) verwenden, tritt diese Beschränkung nicht auf. Dies kann zu Unterschieden bei den Ergebnissen zwischen nativem Modus und UDF-Modus führen.

Automatische SQL-Generierung. (`sql_generation_enabled`) Ermöglicht automatische SQL-Generierung für Streams, wodurch die Leistung erheblich verbessert werden kann. Diese Option ist standardmäßig aktiviert. Die Inaktivierung dieser Option wird nur empfohlen, wenn die Datenbank keine Anfragen unterstützen kann, die von SPSS Modeler Server übermittelt wurden. Beachten Sie, dass diese Einstellung die entsprechende Optimierungseinstellung am Client überschreibt. Beachten Sie außerdem, dass die SQL-Generierung für Scoring-Zwecke unabhängig von dieser Einstellung für jeden Modellierungsknoten gesondert aktiviert werden muss. Wenn diese Einstellung auf dem Server inaktiviert ist, kann der Client sie nicht aktivieren. Ist sie auf dem Server aktiviert, kann der Client sie jedoch inaktivieren.

Standardlänge für SQL-Zeichenfolgen. (`default_sql_string_length`). Gibt die Standardbreite der Zeichenfolgenspalten an, die in den Datenbankcachtabeln erstellt werden. Zeichenfolgfelder in Datenbankcachtabeln werden mit einer Standardbreite von 255 erstellt, wenn keine Informationen von einem vorgeordneten Typ zur Verfügung stehen. Wenn Sie in Ihren Daten breitere Werte haben, instanzieren Sie entweder einen vorgeordneten Typknoten mit diesen Werten oder setzen Sie diesen Parameter auf einen Wert, der groß für diese Zeichenfolgewerte ist.

Datenbank für benutzerdefinierte Funktionen aktivieren. (`db_udf_enabled`). Wenn diese Option auf Y (Standard) gesetzt ist, generiert die Option für die SQL-Generierung kein reines SPSS Modeler-SQL, sondern UDF-SQL (User-Defined Function - benutzerdefinierte Funktion). UDF-SQL ist reinem SQL normalerweise überlegen.

SSL

SSL aktivieren. (`ssl_enabled`) Aktiviert die SSL-Verschlüsselung für Verbindungen zwischen SPSS Modeler und SPSS Modeler Server.

Keystore. (`ssl_keystore`) Die Datenbankdatei für SSL-Schlüssel, die geladen werden soll, wenn der Server gestartet wird (ein vollständiger Pfad oder ein relativer Pfad zum SPSS Modeler-Installationsverzeichnis).

Keystore-Stashdatei. (`ssl_keystore_stash_file`) Der Name der Kennwortstashdatei für die Schlüsseldatenbank, die geladen werden soll, wenn der Server gestartet wird (ein vollständiger Pfad oder ein relativer Pfad zum SPSS Modeler-Installationsverzeichnis). Falls Sie diese Einstellung nicht festlegen wollen und beim Starten von SPSS Modeler Server eine Aufforderung zur Kennworteingabe erfolgen soll, beachten Sie die folgenden Anweisungen:

- Unter Windows:
 1. Stellen Sie sicher, dass für die Einstellung `ssl_keystore_stash_file` in der Datei `options.cfg` kein Wert angegeben ist.

2. Starten Sie SPSS Modeler Server neu. Sie werden zur Angabe eines Kennworts aufgefordert. Wenn Sie nun das richtige Kennwort eingeben und auf **OK** klicken, wird der Server gestartet.
- Unter Linux/UNIX:
 1. Stellen Sie sicher, dass für die Einstellung `ssl_keystore_stash_file` in der Datei `options.cfg` kein Wert angegeben ist.
 2. Suchen Sie die folgende Zeile in der Datei `modelersrv.sh`:


```
if "$INSTALLEDPATH/$SCLEMDNAME" -server $ARGS; then
```
 3. Fügen Sie den Schalter `-request_ssl_password` wie folgt hinzu:


```
if "$INSTALLEDPATH/$SCLEMDNAME" -request_ssl_password -server $ARGS; then
```
 4. Starten Sie SPSS Modeler Server neu. Sie werden zur Angabe eines Kennworts aufgefordert. Wenn Sie nun das richtige Kennwort eingeben und auf **OK** klicken, wird der Server gestartet.

Keystorebeschriftung. (`ssl_keystore_label`) Beschriftung für das angegebene Zertifikat.

Anmerkung: Wenn Sie Administration Console mit einer Servereinrichtung für SSL verwenden wollen, müssen Sie alle von SPSS Modeler Server benötigten Zertifikate in den Deployment Manager-Truststore (unter `../jre/lib/security`) importieren.

Anmerkung: Wenn Sie diese Parameter ändern, müssen Sie SPSS Modeler Server erneut starten, damit die Änderungen wirksam werden.

Konfiguration von Coordinator of Processes

Host. (`cop_host`) Der Hostname bzw. die IP-Adresse des Coordinator of Processes-Service. Der Standardname "spssc" ist ein frei wählbarer Name, den die Administratoren als Alias für den IBM SPSS Collaboration and Deployment Services-Host in DNS hinzufügen können.

Portnummer. (`cop_port_number`) Die Portnummer des Coordinator of Processes-Service. Der Standardport, 8080, ist der IBM SPSS Collaboration and Deployment Services-Standardport.

Kontextstammverzeichnis. (`cop_context_root`) Die URL des Coordinator of Processes-Service.

Anmeldename. (`cop_user_name`) Der Benutzername für die Authentifizierung beim Coordinator of Processes-Service. Dies ist ein IBM SPSS Collaboration and Deployment Services-Anmeldename und kann somit ein Präfix für den Sicherheitsprovider enthalten (z. B.: `ad/jsmith`).

Kennwort. (`cop_password`) Das Kennwort für die Authentifizierung beim Coordinator of Processes-Service.

Anmerkung: Wenn Sie die Datei `options.cfg` manuell aktualisieren, statt Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten) zu verwenden, müssen Sie den Wert `cop_password`, den Sie in der Datei angeben, manuell codieren. Klartextkennwörter sind ungültig und bewirken, dass die Registrierung beim Coordinator of Processes fehlschlägt.

Führen Sie die folgenden Schritte aus, um das Kennwort manuell zu codieren:

1. Öffnen Sie eine Eingabeaufforderung und navigieren Sie zum Verzeichnis `./bin` von SPSS Modeler.
2. Geben Sie, wenn dies angefordert wird, den Benutzernamen ein (den Wert für `cop_user_name`, den Sie in `options.cfg` angeben) und drücken Sie die Eingabetaste.
3. Geben Sie das Kennwort für diesen Benutzer ein, wenn dies angefordert wird.
Das codierte Kennwort wird in der Befehlszeile als Teil der zurückgegebenen Zeichenfolge zwischen Anführungszeichen angezeigt. Beispiel:

```
C:\Programme\IBM\SPSS\Modeler\18\bin>pwutil
User name: copuser
Password: Pass1234
copuser, "0Tqb4n.ob0wrs"
```

4. Kopieren Sie das codierte Kennwort ohne die Anführungszeichen und fügen Sie es zwischen den Anführungszeichen ein, die bereits für den Wert `cop_password` in der Datei `options.cfg` vorhanden sind.

Aktiviert. (`cop_enabled`) Legt fest, ob der Server versuchen soll, eine Registrierung beim Coordinator of Processes durchzuführen. Standardmäßig erfolgt *keine* Registrierung, da der Administrator auswählen sollte, welche Services über den Coordinator of Processes angekündigt werden.

SSL aktiviert. (`cop_ssl_enabled`) Legt fest, ob SSL zum Herstellen der Verbindung zum Coordinator of Processes-Server verwendet werden soll. Falls diese Option verwendet wird, müssen Sie die SSL-Zertifikatsdatei in die SPSS Modeler Server-JRE importieren. Hierzu müssen Sie die SSL-Zertifikatsdatei sowie den zugehörigen Aliasnamen und das Kennwort anfordern. Anschließend führen Sie den folgenden Befehl auf SPSS Modeler Server aus:

```
$JAVA_HOME/bin/keytool -import -trustcacerts -alias $ALIAS_NAME -file $CERTIFICATE_FILE_PATH
-keystore $ModelerServer_Install_Path/jre/lib/security/cacerts
```

Servername. (`cop_service_name`) Der Name dieser SPSS Modeler Server-Instanz. Standardmäßig wird der Hostname verwendet.

Beschreibung. (`cop_service_description`) Eine Beschreibung dieser Instanz.

Aktualisierungsintervall (Min.). (`cop_update_interval`) Die Anzahl der Minuten zwischen Keep-Alive-Nachrichten. Der Standardwert ist 2.

Gewichtung. (`cop_service_weight`) Die Gewichtung der betreffenden Instanz als ganze Zahl zwischen 1 und 10. Ein höherer Wert zieht mehr Verbindungen an. Der Standardwert ist 1.

Service-Host. (`cop_service_host`) Der vollqualifizierte Hostname des IBM SPSS Modeler Server-Hosts. Der Standardhostname wird automatisch abgeleitet. Bei mehrfach vernetzten Hosts kann der Administrator den Standardwert überschreiben.

Standarddatenpfad. (`cop_service_default_data_path`) Der Standarddatenpfad für eine für Coordinator of Processes registrierte IBM SPSS Modeler Server-Installation.

In options.cfg sichtbare Optionen

Die meisten Konfigurationsoptionen können über IBM SPSS Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten) geändert werden. Es gibt jedoch einige Ausnahmen, z. B. die in diesem Abschnitt beschriebenen. Die Optionen in diesem Abschnitt müssen durch Bearbeiten der Datei `options.cfg` geändert werden. Weitere Informationen finden Sie in „IBM SPSS Modeler Server-Verwaltung“ auf Seite 38 und „Verwenden der Datei `options.cfg`“ auf Seite 48. Beachten Sie, dass es möglicherweise weitere Einstellungen in der Datei `options.cfg` gibt, die hier nicht aufgelistet sind.

Anmerkung: Diese Informationen gelten nur für einen fernen Server (beispielsweise IBM SPSS Modeler Server).

`administrators`. Geben Sie die Benutzernamen der Benutzer an, denen Sie Administratorzugriff gewähren wollen. Weitere Informationen finden Sie in „Konfigurieren des Zugriffs mithilfe von Modeler Administration Console“ auf Seite 39.

`allow_config_custom_overrides`. Ändern Sie diesen Wert nur, wenn Sie dazu von einem IBM Technical Support-Mitarbeiter angewiesen werden.

fips_encryption. Ermöglicht die mit FIPS kompatible Verschlüsselung. Der Standardwert ist N.

group_configuration. Wenn IBM SPSS Modeler Server aktiviert wird, prüft die Software die Datei groups.cfg, die steuert, wer sich am Server anmelden kann. Weitere Informationen finden Sie in „Konfigurieren von Gruppen“ auf Seite 55.

max_transfer_size. Nur für die systeminterne Verwendung. **Ändern Sie diesen Wert nicht.**

shell. (Nur bei UNIX-Servern) Überschreibt die Standardeinstellung für die UNIX-Shell, beispielsweise shell, "/usr/bin/ksh". Standardmäßig verwendet IBM SPSS Modeler die Shell, die in dem Benutzerprofil des Benutzers definiert wurde, der die Verbindung zu IBM SPSS Modeler Server herstellt.

start_process_as_login_user. Setzen Sie diese Option auf Y, wenn Sie SPSS Modeler Server mit einer privaten Kennwortdatenbank ausführen und den Serverdienst in einem Nicht-Root-Konto starten.

use_bigint_for_count. Wenn die Anzahl der zu zählenden Datensätze höher ist als eine normale ganze Zahl ($2^{31}-1$) angeben kann, setzen Sie diese Option auf Y. Wenn diese Option auf Y gesetzt ist und ein Stream entweder mit DB2, SQL Server oder einer Teradata-, Oracle- oder Netezza-Datenbank verbunden ist, wird eine Funktion verwendet, wenn eine Datensatzanzahl erforderlich ist (z. B. im vom Aggregat-knoten generierten Feld **Datensatzanzahl**).

Wenn diese Option aktiviert ist und mit DB2 oder SQL Server gearbeitet wird, verwendet SPSS Modeler die Funktion COUNT_BIG() zum Zählen der Datensätze. Bei der Arbeit mit Teradata, Oracle oder Netezza verwendet SPSS Modeler die Funktion COUNT(). Für alle anderen Datenbanken gibt es für die Funktion keinen SQL-Pushback. Der Unterschied liegt darin, dass bei Aktivierung von use_bigint_for_count alle Zahlen der Datensätze als Typ BIG INT (oder LONG) (eine ganze Zahl mit Vorzeichen (64 Bit) bis maximal $2^{63}-1$) gespeichert werden, während sie als normale Ganzzahlen (eine ganze Zahl mit Vorzeichen (32 Bit) bis maximal $2^{31}-1$) gespeichert werden, wenn diese Option nicht aktiviert ist.

cop_ssl_enabled. Setzen Sie diese Option auf Y, wenn Sie SSL zum Herstellen der Verbindung zum Coordinator of Processes-Service verwenden. Falls diese Option verwendet wird, müssen Sie die SSL-Zertifikatsdatei in die SPSS Modeler Server-JRE importieren. Hierzu müssen Sie die SSL-Zertifikatsdatei sowie den zugehörigen Aliasnamen und das Kennwort anfordern. Anschließend führen Sie den folgenden Befehl auf SPSS Modeler Server aus:

```
$JAVA_HOME/bin/keytool -import -trustcacerts -alias $ALIAS_NAME -file $CERTIFICATE_FILE_PATH -keystore $ModelerServer_Install_Path/jre/lib/security/cacerts
```

cop_service_default_data_path. Sie verwenden diese Option, um den Standarddatenpfad für eine für Coordinator of Processes registrierte IBM SPSS Modeler Server-Installation festzulegen.

Benutzer können in SPSS Modeler über **Tools > Analytic Server-Verbindungen** ihre eigenen Analytic Server-Verbindungen erstellen. Administratoren können außerdem eine Analytic Server-Standardverbindung mit den folgenden Eigenschaften definieren:

as_ssl_enabled. Y oder N.

as_host. Geben Sie den Hostnamen oder die IP-Adresse von Analytic Server an.

as_port. Geben Sie die Analytic Server-Portnummer an.

as_context_root. Geben Sie das Analytic Server-Kontextstammverzeichnis an.

as_tenant. Geben Sie den Namen des Nutzers an, dem IBM SPSS Modeler Server als Mitglied angehört.

as_prompt_for_password. Y oder N.

Die Analytic Server-Authentifizierung mithilfe der Kerberos-Methode ist standardmäßig nicht aktiviert. Verwenden Sie die folgenden drei Eigenschaften, um die Kerberos-Authentifizierung zu aktivieren:

`as_kerberos_auth_mode`. Setzen Sie diese Option auf `Y`, um die Kerberos-Authentifizierung zu aktivieren.

`as_kerberos_krb5_conf`. Geben Sie den Pfad zur Kerberos-Konfigurationsdatei an, die von Analytic Server verwendet werden soll, z. B. `c:\windows\krb5.conf`.

`as_kerberos_krb5_spn`. Geben Sie den Kerberos-SPN (Name des Service-Principals) für Analytic Server an, z. B. `HTTP/ashost.mydomain.com@MYDOMAIN.COM`.

SPSS Modeler Server-Überwachung

Im Überwachungsbereich von Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten) wird eine Momentaufnahme aller aktuell auf dem SPSS Modeler Server-Computer laufenden Prozesse angezeigt, ähnlich wie im Windows Task-Manager. Sie öffnen den Überwachungsbereich, indem Sie im Serveradministratorbereich auf den Überwachungsknoten unter dem gewünschten Server doppelklicken. Der Bereich wird dann mit einer aktuellen Momentaufnahme der Serverdaten gefüllt. Die Daten werden im angezeigten Intervall aktualisiert (standardmäßig 1 Minute). Um die Daten von Hand zu aktualisieren, klicken Sie auf **Aktualisieren**. Wenn Sie in dieser Liste nur SPSS Modeler Server-Prozesse anzeigen möchten, klicken Sie auf **Nicht-SPSS Modeler-Prozesse herausfiltern**.

Verwenden der Datei `options.cfg`

Die Datei `options.cfg` befindet sich im Verzeichnis `[Serverinstallationspfad]/config`. Jede Einstellung wird durch ein durch Kommas getrenntes Name/Wert-Paar dargestellt, wobei Name der Name der Option und Wert der Wert für die Option ist. Kommentare werden durch Nummernzeichen (`#`) angegeben.

Anmerkung: Die meisten Konfigurationsoptionen können über IBM SPSS Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten) statt über diese Konfigurationsdatei geändert werden, aber es gibt einige Ausnahmen. Weitere Informationen finden Sie in „In `options.cfg` sichtbare Optionen“ auf Seite 46.

Durch Verwendung von IBM SPSS Modeler Administration Console können Sie bei allen Optionen außer der für den Server-Port Serverneustarts vermeiden. Weitere Informationen finden Sie in „IBM SPSS Modeler Server-Verwaltung“ auf Seite 38.

Anmerkung: Diese Informationen gelten nur für einen fernen Server (beispielsweise IBM SPSS Modeler Server).

Konfigurationsoptionen, die der Standarddatei hinzugefügt werden können

Standardmäßig ist das datenbankinterne Caching bei IBM SPSS Modeler Server aktiviert. Sie können diese Funktion durch Hinzufügen der folgenden Zeile zur Datei `options.cfg` inaktivieren:

```
enable_database_caching, N
```

Dadurch werden temporäre Dateien auf dem Server und nicht in der Datenbank erstellt.

So zeigen Sie die Konfigurationsoptionen für IBM SPSS Modeler Server an oder ändern sie:

1. Öffnen Sie die Datei `options.cfg` mit einem Texteditor.
2. Suchen Sie die relevanten Optionen. Eine vollständige Liste von Optionen finden Sie in „SPSS Modeler Server-Konfiguration“ auf Seite 40.
3. Bearbeiten Sie die Werte wie gewünscht. In allen Pfadnamen muss ein normaler Schrägstrich (`/`) statt eines umgekehrten Schrägstrichs als Trennzeichen verwendet werden.
4. Speichern Sie die Datei.

5. Stoppen Sie IBM SPSS Modeler Server und starten Sie die Software neu, damit die Änderungen in Kraft treten. Weitere Informationen finden Sie in „Starten und Stoppen von IBM SPSS Modeler Server“ auf Seite 27.

Schließen nicht verwendeter Datenbankverbindungen

Standardmäßig wird in IBM SPSS Modeler mindestens eine Verbindung in einer Datenbank zwischengespeichert, sobald diese Verbindung hergestellt wurde. Die Datenbanksitzung bleibt auch dann geöffnet, wenn Streams, die für den Datenbankzugriff erforderlich sind, nicht ausgeführt werden.

Durch die Zwischenspeicherung von Datenbankverbindungen kann die Ausführungsdauer verbessert werden, da IBM SPSS Modeler nicht bei jeder Ausführung eines Streams eine neue Verbindung zur Datenbank aufbauen muss. In einigen Umgebungen ist es jedoch wichtig, dass Anwendungen Datenbanksourcen so schnell wie möglich freigeben. Wenn zu viele IBM SPSS Modeler-Sitzungen nicht mehr verwendete Datenbankverbindungen aufrechterhalten, kann dies zu einer Auslastung der Datenbanksourcen führen.

Sie können dies vermeiden, indem Sie die IBM SPSS Modeler-Option `cache_connection` in einer benutzerdefinierten Datenbankkonfigurationsdatei inaktivieren. Dadurch reagiert IBM SPSS Modeler auch stabiler auf Fehler in der Datenbankverbindung (wie etwa Zeitlimitüberschreitungen), die auftreten, wenn Verbindungen über einen längeren Zeitraum von einer IBM SPSS Modeler-Sitzung verwendet werden.

So schließen Sie nicht verwendete Datenbankverbindungen:

1. Suchen Sie das Verzeichnis `[Serverinstallationspfad]/config`.
2. Fügen Sie folgende Datei hinzu (oder öffnen Sie sie, wenn sie bereits vorhanden ist):
`odbc-custom-properties.cfg`
3. Fügen Sie folgende Zeile in die Datei ein:
`cache_connection, N`
4. Speichern und schließen Sie die Datei.
5. Starten Sie IBM SPSS Modeler Server neu.

Anmerkung:

Datenbankinterne Cachespeicher werden je nach Implementierung der jeweiligen Datenbank entweder als reguläre Tabelle oder als temporäre Tabelle in der Datenbank gespeichert. Beispiel: Temporäre Tabellen werden für DB2, HP Neoview, Oracle, Amazon Redshift, Sybase und Teradata verwendet. Für diese Datenbanken funktioniert das Setzen von `cache_connection` auf N nicht wie erwartet, da die temporäre Tabelle nur innerhalb einer Sitzung gültig ist (sie wird automatisch von der Datenbank bereinigt, wenn die Datenbankverbindung geschlossen wird).

Wenn also ein SPSS Modeler-Stream für eine dieser Datenbanken ausgeführt wird und `cache_connection` auf N gesetzt ist, kann ein Fehler wie der folgende auftreten: **Tabelle für datenbankinternes Caching konnte nicht erstellt werden. Stattdessen wird Dateicache verwendet.** Damit wird angezeigt, dass SPSS Modeler den datenbankinternen Cache nicht erstellen konnte. Außerdem wird in einigen Fällen bei einer von SPSS Modeler generierten SQL-Abfrage eine temporäre Tabelle verwendet. Diese Tabelle ist jedoch leer.

Zum Umgehen dieses Problems können Sie eine reguläre Datenbanktabelle für datenbankinterne Cachespeicher verwenden. Erstellen Sie hierzu eine benutzerdefinierte Konfigurationsdatei für Datenbankeigenschaften, die die folgende Zeile enthält:

```
table_create_temp_sql, 'CREATE TABLE <Tabellenname> <(Tabellenspalten)>'
```

Dadurch wird die Verwendung einer regulären Datenbanktabelle für den datenbankinternen Cache erzwungen und die Tabelle wird gelöscht, wenn alle Verbindungen zur Tabelle geschlossen werden oder wenn der aktive Stream geschlossen wird.

Verwenden von SSL zur sicheren Datenübertragung

Secure Sockets Layer (SSL) ist ein Protokoll für die Verschlüsselung von Daten, die zwischen zwei Computern übertragen werden. SSL stellt sicher, dass die Kommunikation zwischen den Computern sicher ist. SSL kann die Authentifizierung von Benutzername/Kennwort sowie den Inhalt eines Austauschs zwischen einem Server und einem Client verschlüsseln.

Funktionsweise von SSL

SSL beruht auf dem öffentlichen und privaten Schlüssel des Servers sowie einem Zertifikat für den öffentlichen Schlüssel, das die Identität des Servers mit seinem öffentlichen Schlüssel verbindet.

1. Wenn ein Client eine Verbindung zu einem Server aufbaut, authentifiziert der Client den Server mit dem Zertifikat für den öffentlichen Schlüssel.
2. Der Client generiert dann eine Zufallszahl, verschlüsselt die Zahl mit dem öffentlichen Schlüssel des Servers und sendet die verschlüsselte Nachricht zurück an den Server.
3. Der Server entschlüsselt die Zufallszahl mit seinem privaten Schlüssel.
4. Aus der Zufallszahl generieren Server und Client die Sitzungsschlüssel, die zur Verschlüsselung und Entschlüsselung nachfolgender Informationen verwendet werden.

Das Zertifikat für den öffentlichen Schlüssel ist in der Regel von einer Zertifizierungsstelle signiert. Zertifizierungsstellen wie VeriSign und Thawte sind Organisationen, die Sicherheitsdaten, die sich in den Zertifikaten für öffentliche Schlüssel befinden, herausgeben, authentifizieren und verwalten. Im Wesentlichen bestätigt die Zertifizierungsstelle die Identität des Servers. Die Zertifizierungsstelle berechnet gewöhnlich eine Gebühr für ein Zertifikat, jedoch können auch selbst signierte Zertifikate generiert werden.

Schützen der Client/Server- und Server/Server-Kommunikation durch SSL

Hauptschritte beim Schützen der Client/Server- und Server/Client-Kommunikation durch SSL:

1. Beziehen und installieren Sie das SSL-Zertifikat und die Schlüssel.
2. Aktivieren und konfigurieren Sie SSL in der Serververwaltungsanwendung (IBM SPSS Deployment Manager).
3. Wenn Sie Verschlüsselungszertifikate verwenden, die eine Verschlüsselung mit einer Stärke von über 2048 Bit aufweisen, installieren Sie Verschlüsselung mit unbegrenzter Stärke auf den -Client-Computern.
4. Weisen Sie Benutzer an, bei der Verbindung zum Server SSL zu aktivieren.

Anmerkung: Gelegentlich fungiert ein Serverprodukt als Client. Ein Beispiel ist ein IBM SPSS Statistics-Server, der eine Verbindung zu IBM SPSS Collaboration and Deployment Services Repository aufbaut. In diesem Fall ist IBM SPSS Statistics-Server der *Client*.

Beziehen und Installieren von SSL-Zertifikat und -Schlüssel

Zur Konfiguration der SSL-Unterstützung müssen Sie zuerst die folgenden Schritte durchführen:

1. Beziehen Sie ein SSL-Zertifikat und eine Schlüsseldatei. Dazu haben Sie drei Möglichkeiten.
 - Erwerben Sie Zertifikat und Schlüsseldatei von einer öffentlichen Zertifizierungsstelle (z. B. VeriSign, Thawte oder Entrust). Die öffentliche Zertifizierungsstelle signiert das Zertifikat, um den Server zu verifizieren, der es verwendet.
 - Generieren Sie Schlüssel- und Zertifikatsdateien mithilfe einer unabhängigen Zertifizierungsstelle. Bei dieser Methode muss das Stammzertifikat der unabhängigen Zertifizierungsstelle in die Keystore-Dateien von Client und Server importiert werden. Weitere Informationen finden Sie in „Importieren eines Stammzertifikats einer unabhängigen Zertifizierungsstelle“ auf Seite 52.
 - Generieren Sie Schlüssel- und Zertifikatsdateien mithilfe einer internen Zertifizierungsstelle für selbst signierte Zertifikate. Hierfür sind die folgenden Schritte erforderlich:

- a. Bereiten Sie eine Schlüsseldatenbank vor. Weitere Informationen finden Sie in „Erstellen einer SSL-Schlüsseldatenbank“ auf Seite 52.
 - b. Erstellen Sie das selbst signierte Zertifikat. Weitere Informationen finden Sie in „Erstellen eines selbst signierten SSL-Zertifikats“ auf Seite 52.
2. Kopieren Sie die in Schritt 1 erstellten .kdb- und .sth-Dateien in ein Verzeichnis, auf das IBM SPSS Modeler Server Zugriff hat, und geben Sie den Pfad zu diesem Verzeichnis in der Datei *options.cfg* an.

Anmerkung: Verwenden Sie im Verzeichnispfad normale Schrägstriche (/) als Trennzeichen.

3. Legen Sie in der Datei *options.cfg* die folgenden Parameter fest:
 - `ssl_enabled, Y`
 - `ssl_keystore, "<Dateiname>.kdb"` (Dabei ist *<Dateiname>* der Name Ihrer Schlüsseldatenbank.)
 - `ssl_keystore_stash_file, "<Dateiname>.sth"` (Dabei ist *<Dateiname>* der Name der Kennwortstashdatei für die Schlüsseldatenbank.)
 - `ssl_keystore_label, <Beschriftung>` (Dabei ist *<Beschriftung>* die Beschriftung Ihres Zertifikats.)
4. Installieren Sie für selbst signierte Zertifikate oder Zertifikate einer unabhängigen Zertifizierungsstelle das Zertifikat auf Clientsystemen. Bei öffentlichen Zertifikaten, die von einer unabhängigen Zertifizierungsstelle erworben wurden, ist dieser Schritt nicht erforderlich. Stellen Sie sicher, dass Zugriffsberechtigungen ein zufälliges Anzeigen des Verzeichnisses verhindern, das das Zertifikat enthält. Weitere Informationen finden Sie in „Installieren eines selbst signierten SSL-Zertifikats“ auf Seite 52.

Konfigurieren der Umgebung für die GSKit-Ausführung: GSKCapiCmd ist ein nicht auf Java basierendes Befehlszeilentool, sodass Java™ nicht auf Ihrem System installiert sein muss, um dieses Tool verwenden zu können; es befindet sich im Ordner *<Modeler-Installationsverzeichnis>/bin*. Der Prozess, mit dem Sie Ihre Umgebung für die Ausführung von IBM Global Security Kit (GSKit) konfigurieren, variiert je nach verwendeter Plattform.

Für eine Konfiguration für Linux/UNIX fügen Sie Ihrer Umgebung das Verzeichnis für die gemeinsam genutzten Bibliotheken *<Modeler-Installationsverzeichnis>/lib* hinzu:

```
$export <Umgebungsvariable_für_Pfad_für_gemeinsam_genutzte_Bibliothek>=<Modeler_Server-Installationspfad>/bin
$export PATH=$PATH:<Modeler_Server-Installationspfad>/bin
```

Der Variablenname für den Pfad für die gemeinsam genutzte Bibliothek hängt von Ihrer Plattform ab:

- AIX verwendet den Variablennamen: LIBPATH
- HP-UX verwendet den Variablennamen: SHLIB_PATH
- Linux verwendet den Variablennamen: LD_LIBRARY_PATH

Um beispielsweise die Umgebung unter Linux festzulegen, verwenden Sie Folgendes:

```
$export LD_LIBRARY_PATH=/path/to/gskit/bin
$export PATH=$PATH:/path/to/gskit/bin
```

Kontozugriff auf Dateien

Stellen Sie sicher, dass Sie die richtigen Berechtigungen für die Konten erteilen, die auf die SSL-Dateien zugreifen:

1. Erteilen Sie allen von SPSS Modeler zur Verbindung verwendeten Konten Lesezugriff auf die SSL-Dateien.

Anmerkung: Dies gilt auch für den Benutzer *Anmelden als*, der im SPSS Modeler Server-Service definiert ist. Unter UNIX oder Linux gilt dies für den Benutzer, als der Sie SPSS Modeler Server starten.

2. Unter Windows reicht es nicht aus, dass sich die Konten in der Gruppe "Administratoren" befinden und dieser Gruppe die Berechtigung erteilt wird, wenn die Benutzerzugriffssteuerung (UAC - User Access Control) aktiviert ist. Zusätzlich müssen Sie eine der folgenden Aktionen ausführen:

- Erteilen Sie den Konten separat Berechtigungen.
- Erstellen Sie eine neue Gruppe, fügen Sie dieser neuen Gruppe Konten hinzu und erteilen Sie der Gruppe die Berechtigung für den Zugriff auf die SSL-Dateien.
- Inaktivieren Sie UAC.

Erstellen einer SSL-Schlüsseldatenbank: Verwenden Sie das GSKCapiCmd-Tool, um Ihre Schlüsseldatenbank zu erstellen. Vor der Verwendung des Tools müssen Sie Ihre Umgebung konfigurieren (weitere Informationen hierzu finden Sie in „Konfigurieren der Umgebung für die GSKit-Ausführung“ auf Seite 51).

Um die Schlüsseldatenbank zu erstellen, führen Sie GSKit aus und geben Sie den folgenden Befehl ein:

```
gsk<Ver>capicmd[_64] -keydb -create -populate -db <Dateiname>.kdb -pw <Kennwort> -stash
```

Dabei steht <Ver> für die GSKit-Versionsnummer, <Dateiname> für den Namen, den Sie für die Schlüsseldatenbankdatei verwenden wollen und <Kennwort> für das Kennwort für die Schlüsseldatenbank.

Die Option *-stash* erstellt im Pfad für die Schlüsseldatenbank eine Stashdatei mit der Dateierweiterung *.sth*. GSKit ruft mithilfe der Stashdatei das Kennwort für die Schlüsseldatenbank ab, sodass es nicht jedes Mal in die Befehlszeile eingegeben werden muss.

Anmerkung: Für die *.sth*-Datei sollten Sie den Schutz des Dateisystems auf hohe Einstellungen setzen.

Erstellen eines selbst signierten SSL-Zertifikats: Mit dem folgenden Befehl können Sie ein selbst signiertes Zertifikat generieren und in der Schlüsseldatenbank speichern:

```
gsk<Ver>capicmd[_64] -cert -create -db <Dateiname>.kdb -stashed -dn "CN=myserver,OU=mynetwork,O=mycompany,C=mycountry" -label <Beschriftung> -expire <Anzahl Tage, die das Zertifikat gültig ist> -default_cert yes
```

Dabei steht <Ver> für die GSKit-Versionsnummer, <Dateiname> für den Namen der Schlüsseldatenbankdatei, <Anzahl Tage, die das Zertifikat gültig ist> für die physische Anzahl der Tage, die das Zertifikat gültig ist und <Beschriftung> für eine beschreibende Beschriftung, die das Identifizieren der Datei erleichtert (eine mögliche Beschriftung wäre z. B.: *meinselbstsigniertesZertifikat*).

Installieren eines selbst signierten SSL-Zertifikats: Für die Clientsysteme, die über SSL eine Verbindung zu Ihrem Server herstellen, müssen Sie den öffentlichen Teil des Zertifikats an die Clients verteilen, damit er in deren Schlüsseldatenbanken gespeichert werden kann. Gehen Sie dazu wie folgt vor:

1. Extrahieren Sie den öffentlichen Teil einer Datei mithilfe des folgenden Befehls:

```
gsk<Ver>capicmd[_64] -cert -extract -db <Dateiname>.kdb -stashed -label <Beschriftung> -format ascii -target mycert.arm
```

2. Verteilen Sie *mycert.arm* an die Clients. Die Datei sollte in das jeweilige Verzeichnis *jre/bin* kopiert werden.
3. Fügen Sie das neue Zertifikat unter Verwendung des folgenden Befehls der Schlüsseldatenbank des Clients hinzu:

```
keytool -import -alias <Beschriftung> -keystore ..\lib\security\cacerts -file mycert.arm
```

Wenn Sie zur Eingabe eines Kennworts aufgefordert werden, verwenden Sie *changeit*. 'keytool' befindet sich im Verzeichnis <Modeller-Installationsverzeichnis>\jre\bin (oder im Verzeichnis <Modeller-Installationsverzeichnis>/SPSSModeler.app/Contents/PlugIns/jre/Contents/Home/bin unter Mac OS).

Importieren eines Stammzertifikats einer unabhängigen Zertifizierungsstelle: Anstatt ein Zertifikat von einer namhaften Zertifizierungsstelle zu erwerben oder ein selbst signiertes Zertifikat zu erstellen, können Sie Ihre Serverzertifikate auch von einer unabhängigen Zertifizierungsstelle signieren lassen. Client und Server müssen Zugriff auf das Stammzertifikat einer unabhängigen Zertifizierungsstelle haben, um die von der unabhängigen Zertifizierungsstelle signierten Serverzertifikate prüfen zu können. Gehen Sie dazu wie folgt vor:

1. Beziehen Sie ein Stammzertifikat von einer unabhängigen Zertifizierungsstelle. Der hierfür erforderliche Prozess variiert je nach Zertifizierungsstelle. Die Stammzertifikate unabhängiger Zertifizierungsstellen sind häufig per Download verfügbar.
2. Fügen Sie das Zertifikat der Schlüsseldatenbank des Servers unter Verwendung des folgenden Befehls hinzu:

```
gsk<Ver>capicmd[_64] -cert -add -db <Dateiname>.kdb -stashed -label <Beschriftung>
-file <Zertifikat einer Zertifizierungsstelle>.crt
-format binary -trust enable
```

3. Fügen Sie das Zertifikat der Schlüsseldatenbank des Clients unter Verwendung des folgenden Befehls hinzu:

Unter Windows:

```
C:> cd <Modeler-Client-Installationspfad>\jre\bin
C:> keytool -import -keystore ..\lib\security\cacerts -file <Zertifikat einer
Zertifizierungsstelle>.crt
-alias <Beschriftung>
```

Unter Mac OS:

```
C:> cd <Modeler-Client-Installationspfad>/SPSSModeler.app/Contents/PlugIns/jre/Contents/Home/bin
C:> keytool -import -keystore ..\lib\security\cacerts -file <Zertifikat einer
Zertifizierungsstelle>.crt
-alias <Beschriftung>
```

Wenn Sie zur Eingabe eines Kennworts aufgefordert werden, verwenden Sie *changeit*. 'keytool' befindet sich im Verzeichnis <Modeler-Installationsverzeichnis>\jre\bin (oder im Verzeichnis <Modeler-Installationsverzeichnis>/SPSSModeler.app/Contents/PlugIns/jre/Contents/Home/bin unter Mac OS).

4. Validieren Sie die Schlüsseldatenbank des Servers mithilfe des Stammzertifikats der unabhängigen Zertifizierungsstelle unter Verwendung des folgenden Befehls:

```
gsk<Ver>capicmd[_64] -cert -validate -db <Dateiname>.kdb -stashed -label <Beschriftung>
```

Wenn die Validierung erfolgreich war, wird dies durch die folgende Nachricht angezeigt: OK.

Anmerkung: Bei den oben beschriebenen Befehle wird ein im Binärformat gehaltenes Stammzertifikat der unabhängigen Zertifizierungsstelle verwendet. Hat das Zertifikat ein ASCII-Format, verwenden Sie die Option `-format ascii`.

Der Parameter `-db` gibt den Namen der Schlüsseldatenbank an, in die Sie das Stammzertifikat der unabhängigen Zertifizierungsstelle importieren wollen.

Der Parameter `-label` gibt die Beschriftung an, die für das Stammzertifikat der unabhängigen Zertifizierungsstelle in der Schlüsseldatenbankdatei verwendet werden soll. Die Beschriftung, die Sie hier verwenden, kann frei gewählt werden, da sie keinen Bezug zu den Beschriftungen hat, die in der Datei IBM SPSS Modeler options.cfg verwendet werden.

Der Parameter `-file` gibt die Datei an, die das Stammzertifikat der unabhängigen Zertifizierungsstelle enthält.

Aktivieren und Konfigurieren von SSL in IBM SPSS Deployment Manager

1. Wenn Sie ein selbst signiertes SSL-Zertifikat installieren, kopieren Sie die von Ihnen erstellte cacerts-Datei in das Verzeichnis <Deployment Manager-Installationsverzeichnis>\jre\lib\security. Weitere Informationen finden Sie in „Installieren eines selbst signierten SSL-Zertifikats“ auf Seite 52.
2. Starten Sie die Serververwaltungsanwendung (IBM SPSS Deployment Manager installiert wird) und stellen Sie eine Verbindung zum Server her.
3. Stellen Sie auf der Konfigurationsseite die Option **Secure Sockets Layer** auf Ja ein.
4. Geben Sie in **SSL-Datei - öffentlicher Schlüssel** den vollständigen Pfad zur Datei mit dem öffentlichen Schlüssel an.

5. Geben Sie in **SSL-Datei - privater Schlüssel** den vollständigen Pfad zur Datei mit dem privaten Schlüssel an.
Hinweis: Wenn öffentlicher und privater Schlüssel in einer Datei gespeichert werden, geben Sie in **SSL-Datei - öffentlicher Schlüssel** und **SSL-Datei - privater Schlüssel** dieselbe Datei an.
6. Wählen Sie in den Menüs Folgendes aus:
Datei > Speichern
7. Starten Sie den Serverdienst oder -dämon neu. Beim Neustart werden Sie zur Eingabe des SSL-Kennworts aufgefordert. Unter Windows können Sie **Kennwort speichern** wählen, um das Kennwort sicher zu speichern. Damit ist es nicht mehr nötig, bei jedem Serverstart das Kennwort einzugeben.

Installieren der Verschlüsselung mit unbegrenzter Stärke

Bei der als Teil des Produkts ausgelieferten Java Runtime Environment ist Verschlüsselung mit US-Exportstärke aktiviert. Zur besseren Sicherheit Ihrer Daten wird ein Upgrade auf eine Verschlüsselung mit unbegrenzter Stärke empfohlen. Dieses Verfahren muss sowohl für die Client- als auch für die Serverinstallation wiederholt werden.

Installieren der Verschlüsselung mit unbegrenzter Stärke

1. Laden Sie nicht eingeschränkte SDK-JCE-Richtliniendateien von der Website IBM.com herunter. (Wählen Sie dabei die für Java 7 gültigen Dateien aus.)
2. Extrahieren Sie die in der komprimierten Datei gepackten Standortrichtliniendateien mit unbegrenzter Stärke. Die komprimierte Datei enthält eine Datei namens `US_export_policy.jar` und eine Datei namens `local_policy.jar`.
3. Erstellen Sie eine Sicherungskopie der vorhandenen Dateien `US_export_policy.jar` und `local_policy.jar` aus dem Verzeichnis `jre/lib/security`.
4. Ersetzen Sie die vorhandenen Dateien `US_export_policy.jar` und `local_policy.jar` durch die beiden Dateien, die Sie heruntergeladen und extrahiert haben.
5. Starten Sie IBM SPSS Modeler Client oder IBM SPSS Modeler Server wie erforderlich neu.

Anweisung an Benutzer, SSL zu aktivieren

Wenn Endbenutzer über ein Clientprodukt eine Verbindung zum Server herstellen, müssen Sie SSL im Dialogfeld für die Verbindung zum Server aktivieren.

Cognos-SSL-Verbindung

Um die Verbindung zum Cognos Analytics-Server mit HTTPS und einem sicheren SSL-Port herstellen zu können, müssen Sie zunächst einige der internen und externen Dispatchereinstellungen von Cognos ändern. Details zu den erforderlichen Änderungen finden Sie im Konfigurations- und Verwaltungshandbuch zu Cognos Server.

Nach dem Ändern der Dispatchereinstellungen importieren Sie die zuvor in Cognos erstellte SSL-Zertifizierung in die SPSS Modeler-JRE, indem Sie die folgenden Schritte ausführen:

1. Legen Sie in Cognos Configuration ein Kennwort für den IBM Cognos-Keystore fest:
 - a. Klicken Sie im Explorerfenster auf **Kryptografie > Cognos**.
 - b. Legen Sie im Eigenschaftenfenster unter **Encryption Key-Einstellungen** das Kennwort für den Keystore für Verschlüsselungsschlüssel fest.
 - c. Wählen Sie **Speichern** im Menü **Datei** aus.
 - d. Wählen Sie **Neustart** im Menü **Aktionen** aus.
2. Wechseln Sie über die Befehlszeile zum Verzeichnis `c10_location\bin`.
3. Setzen Sie die Umgebungsvariable `JAVA_HOME` auf den JRE-Speicherort (Java Runtime Environment), der von dem Anwendungsserver verwendet wird, der Cognos ausführt. Beispiel:
`set JAVA_HOME=c11_location\bin\jre\<Version>`
4. Führen Sie das Zertifikattool über die Befehlszeile aus. Beispiel:

```
ThirdPartyCertificateTool.bat -E -T -r ca.cer -k ..\configuration\encryptkeypair\jEncKeystore
-p <Kennwort>
```

5. Kopieren Sie die Datei ca.cer an den Speicherort von SPSS Modeler Server.
6. Öffnen Sie eine Befehlszeile und wechseln Sie zum Ordner <Modeler-Installationsverzeichnis>\jre\bin.
7. Führen Sie den Befehl zum Importieren des Zertifikats aus. Beispiel:

```
.\keytool -import -alias ca -file <Verzeichnis mit der Datei ca.cer>\ca.cer
-keystore "<Modeler-Installationsverzeichnis>\jre\lib\security\cacerts"
```

Anschließend können Sie HTTPS und den mit SSL gesicherten Dispatcher verwenden, um eine Verbindung zu Cognos herzustellen. Beispiel:

```
https://9.119.83.37:9343/p2pd/servlet/dispatch
```

Cognos TM1-SSL-Verbindung

Führen Sie die folgenden Schritte aus, um die Verbindung zu Cognos TM1 mit HTTPS und einem sicheren SSL-Port herzustellen:

1. Konfigurieren Sie Tomcat SSL. (Weitere Informationen finden Sie zum Beispiel unter <http://tomcat.apache.org/tomcat-7.0-doc/ssl-howto.html>).
 - a. Wechseln Sie über die Befehlszeile zum Verzeichnis C:\Programme\ibm\cognos\tm1_64\bin64\jre\7.0\bin (dies ist der Standardinstallationspfad) und führen Sie den folgenden Befehl aus, um eine Datei mit dem Namen .keystore in Ihrem Ausgangsordner zu generieren:

```
keytool -genkey -alias tomcat -keyalg RSA
```
 - b. Fügen Sie der Datei server.xml im Ordner C:\Programme\ibm\cognos\tm1_64\tomcat\conf die folgenden Connectoreinstellungen hinzu:

```
<Connector SSLEnabled="true" acceptCount="100" clientAuth="false"
  disableUploadTimeout="true" enableLookups="false" maxThreads="25"
  port="8443" keystoreFile="/Users/loiane/.keystore" keystorePass="password"
  protocol="org.apache.coyote.http11.Http11NioProtocol" scheme="https"
  secure="true" sslProtocol="TLS" />
```
 - c. Starten Sie den IBM Cognos TM1 Application Server-Service.
2. Geben Sie folgenden Befehl in der Befehlszeile ein, um die Zertifizierungsdatei für den neu erstellten Keystore zu exportieren:

```
keytool -export -alias tomcat -file certfile.cer -keystore C:\Users\Administrator\.keystore
```
3. Geben Sie folgenden Befehl in der Befehlszeile ein, um die Zertifizierungsdatei in die von SPSS Modeler Server verwendete JRE zu importieren:

```
keytool -import -alias as -file C:\Leon\Temp\certfile.cer -keystore
"c:\Programme\IBM\SPSS\Modeler\<Version>\jre\lib\security\cacerts"
```

Sie können dann HTTPS und die sichere SSL-Portnummer verwenden, um eine Verbindung zu Cognos TM1 herzustellen.

Konfigurieren von Gruppen

Ein authentifizierter Benutzer gehört in der Regel zu mindestens einer Sicherheitsgruppe und wenn die gruppenbasierte Konfiguration für SPSS Modeler Server aktiviert ist, können diese Gruppen verwendet werden, um die Anmeldung am Server zuzulassen oder zu verweigern oder um die Optionseinstellungen für die Benutzersitzung anzupassen.

Die Gruppenkonfiguration wird in den folgenden Szenarios unterstützt:

- Bei einer *Standardinstallation*, bei der der SPSS Modeler Server-Dienst unter dem Konto "Lokales System" oder dem Benutzerkonto mit Rootberechtigung ausgeführt wird und der Benutzer sich über expli-

zite Berechtigungsnachweise oder mithilfe von Single Sign-On (SSO) anmeldet: In diesem Fall handelt es sich bei den Gruppen um die Benutzersicherheitsgruppen des Betriebssystems, die zur Steuerung des Dateizugriffs usw. verwendet werden.

- Bei einer *Installation ohne Rootberechtigung*, bei der der SPSS Modeler Server-Dienst unter einem nicht privilegierten Konto ausgeführt wird und der Benutzer sich mithilfe von SSO anmeldet: In diesem Fall handelt es sich bei den Gruppen um die LDAP-Gruppen, die zum SSO-Principal gehören. Diese Gruppen werden vom LDAP-Sicherheitsprovider in IBM SPSS Collaboration and Deployment Services bezogen, sodass für dieses Szenario einige zusätzliche Konfigurationsschritte erforderlich sind. Weitere Informationen finden Sie in „Abrufen der Gruppenzugehörigkeit des SSO-Benutzers“ auf Seite 20.

Falls keines der beiden Szenarios zutrifft, sind die Gruppen des Benutzers nicht verfügbar und die Gruppenkonfiguration wird nicht unterstützt. Insbesondere bei einer *Installation ohne Rootberechtigung*, bei der der SPSS Modeler Server-Dienst unter einem nicht privilegierten Konto ausgeführt wird und der Benutzer sich über einen Benutzernamen und ein Kennwort anmeldet, sind die Betriebssystemgruppen für den Server nicht verfügbar und die Gruppenkonfiguration wird nicht unterstützt.

Bei der gruppenbasierten Konfiguration können sich die auf die Sitzung eines Benutzers angewendeten Optionseinstellungen je nach Gruppenzugehörigkeit des Benutzers unterscheiden. Dabei handelt es sich um die serverseitigen Einstellungen, die normalerweise aus der Datei `options.cfg` von SPSS Modeler Server gelesen und für alle Sitzungen gleich festgelegt werden. Die Datei `options.cfg` stellt die Standardeinstellungen für alle Sitzungen bereit, es können jedoch gruppenspezifische Konfigurationsdateien vorhanden sein, die für bestimmte Sitzungen ein Subset von Einstellungen überschreiben.

Die Gruppenkonfiguration ermöglicht die Steuerung verschiedener Einstellungen. Beispiele:

- Steuerung des Zugriffs auf Dateien und Datenquellennamen
- Steuerung der Ressourcennutzung

Wenn die Option `group_configuration` in `options.cfg` aktiviert ist, prüft IBM SPSS Modeler Server die Datei `groups.cfg`, die steuert, wer sich am Server anmelden kann. Der Standardwert ist `N`. Es folgt ein Beispiel für `groups.cfg`, das der Gruppe `Test` den Zugriff auf den Server verweigert und der Gruppe `Fraud` den Zugriff mit einer angegebenen Konfiguration ermöglicht. Durch den Stern wird allen anderen Gruppen der Zugriff mit der Standardkonfiguration ermöglicht.

```
Test, DENY
Fraud, "groups/fraud.cfg"
*,
```

Eine bestimmte Gruppenkonfiguration, wie z. B. die für "Fraud" (siehe oben), kann möglicherweise den Zugriff auf bestimmte Datenquellen beschränken oder Ressourceneinstellungen (hinsichtlich SQL-Pushback, Hauptspeicherbelegung, Multithreading usw.) ändern, um die Leistung für Mitglieder dieser Gruppe zu verbessern.

Steuern des DSN-Zugriffs nach Gruppe

Bei der Mehrfaktorauthentifizierung (MFA - Multi-Factor Authentication) muss das Set an ODBC-Datenquellennamen (DSN - Data Source Name), auf das Benutzer je nach ihrer Gruppenzugehörigkeit Zugriff haben, eingeschränkt werden können.

Dies kann nach einem ähnlichen Schema wie beim Dateizugriff realisiert werden. In `options.cfg` stehen zwei Konfigurationseinstellungen zur Verfügung:

```
sql_data_sources_restricted, N
sql_data_source_path, ""
```

Wenn `sql_data_sources_restricted` auf `Y` gesetzt ist, bleibt der Benutzer auf die im zugehörigen Pfad aufgeführten Datenquellennamen beschränkt. Datenquellennamen werden durch das standardmäßig ver-

wendete Pfadtrennzeichen voneinander getrennt, unter Windows ist dies das Semikolon (;) und unter UNIX der Doppelpunkt (:). Auf Windows basierendes Beispiel:

```
sql_data_sources_restricted, Y
sql_data_source_path, "Fraud - Analytic;Fraud - Operational"
```

Wenn diese Einschränkung angewendet wird, führt dies zu folgenden Ergebnissen:

- Wenn ein Benutzer nach Datenquellen sucht (z. B. über das Dialogfeld "ODBC-Verbindung") oder bei Verwendung der API `getServerDataSourceNames` für PSAPI-Sitzungen), werden dem Benutzer nicht alle der auf dem Serversystem definierten Datenquellennamen angezeigt, sondern nur das im Pfad `options.cfg` definierte Subset der Datenquellennamen. Beachten Sie, dass der Pfad möglicherweise Datenquellennamen enthält, die nicht auf dem Server definiert sind, und dass diese ignoriert werden (dem Benutzer werden diese Namen nicht angezeigt).
- Wenn ein Benutzer einen ODBC-Knoten erstellt, der ein Script oder eine PSAPI verwendet, und dabei einen Datenquellennamen angibt, der nicht im Pfad `options.cfg` enthalten ist, dann wird der Knoten nicht ausgeführt und dem Benutzer wird ein Fehler ähnlich dem folgenden angezeigt: **Zugriff verweigert auf Datenquelle: <X>**. (Dies gilt auch für jeden anderen Knoten, der eine ODBC-Verbindung verwendet.)

Der Datenquellenpfad kann die Einfügungen `PATH`, `GROUP` und `USER` enthalten, die an anderer Stelle im Rahmen der Dateipfade beschrieben sind. Die Einfügung `PATH` ermöglicht es, bei Verwendung der gruppenbasierten Konfiguration, den Pfad inkrementell entsprechend der Gruppenzugehörigkeit des Benutzers zu erstellen. Unter bestimmten Umständen kann es auch sinnvoll sein, einen Datenquellennamen nach der Gruppe zu benennen, die dessen Eigner ist.

Wenn in Anknüpfung an das vorherige Beispiel nur Mitgliedern der Gruppe "Betrugsanalysten" (Fraud Analysts) Zugriff auf die Datenquellen `Fraud` gestattet wird, kann die Site die Gruppenkonfiguration aktivieren und eine Konfiguration speziell für die Betrugsanalysten erstellen, die mindestens die folgende Zeile enthält:

```
sql_data_source_path, "${PATH};Fraud - Analytic;Fraud - Operational"
```

Durch Hinzufügen des Präfixes `PATH` in diesem Beispiel wird sichergestellt, dass die Betrugsanalysten weiterhin auf andere Datenquellen, die für jeden zugänglich sind, bzw. auf andere Gruppen, deren Mitglieder sie sind, zugreifen können.

Serverprotokoll

In einer Protokolldatei mit dem Namen `server_logging.log` werden alle wichtigen Aktionen von IBM SPSS Modeler Server erfasst. Unter UNIX befindet sich diese Datei im Ordner `log` im Installationsverzeichnis, unter Windows befindet sie sich im Verzeichnis `%ALLUSERSPROFILE%/IBM/SPSS/Modeler Server/<Version>/log`.

Die Einstellungen, die steuern, wie die Protokollierung in Ihrer Installation ausgeführt wird, befinden sich in der Datei `log4cxx.properties`.

Ändern des Speicherorts der Protokolldatei

Der Standardspeicherort der Protokolldatei ist in der Datei `log4cxx.properties` wie folgt festgelegt:

```
log4j.appender.MainLog.File=${app_log_location}/${PROFILE_NAME}/${app_type}logging.log
```

Wenn Sie den Speicherort der Protokolldatei ändern wollen, bearbeiten Sie diesen Eintrag.

Aktivieren der Traceerstellung

Manchmal benötigen Sie ausführlichere Details als lediglich eine grundlegende Liste von Informationen, die die Hauptaktionen zeigt. Diese Details können z. B. für Ihr Support-Team zur Problembehebung erforderlich sein. In diesen Situationen können Sie die Einstellungen so ändern, dass im Protokoll detailliertere Traceinformationen bereitgestellt werden.

Um die Tracerstellung zu aktivieren, inaktivieren Sie in der Datei `log4cxx.properties` die Zeile `log4j.rootLogger=INFO, MainLog, ConsoleLog` und aktivieren Sie stattdessen die folgende Zeile:
`log4j.rootLogger=TRACE, MainLog, TraceLog`

Bearbeiten Sie den folgenden Eintrag, um den Speicherort des Traceprotokolls zu ändern:

```
log4j.appender.TraceLog.File=${app_log_location}/${PROFILE_NAME}/${app_type}tracing_${PROCESS_ID}.log
```

Ändern der Protokollierungsoptionen

Die Datei `log4cxx.properties` enthält die Einstellungen, die definieren, wie verschiedene Ereignisse protokolliert werden. Diese Einstellungen sind normalerweise auf **INFO** zum Aufzeichnen von Aktionen in der Protokolldatei oder auf **WARN** zum Benachrichtigen des Benutzers bei einem potenziellen Problem gesetzt. Wenn Sie anhand der Protokolldatei potenzielle Fehler identifizieren wollen, können Sie auch einige der Einstellungen auf **TRACE** setzen.

Steuern der Größe der Protokolldatei

Standardmäßig wächst die Protokolldatei bei jeder Verwendung von SPSS Modeler Server. Wenn Sie verhindern wollen, dass das Protokoll zu groß wird, können Sie festlegen, dass es jeden Tag neu erstellt wird, oder eine Größenbeschränkung dafür definieren.

Um festzulegen, dass das Protokoll jeden Tag neu gestartet wird, verwenden Sie in der Datei `log4cxx.properties` die folgenden Einträge:

```
log4j.appender.MainLog=org.apache.log4j.DailyRollingFileAppender
log4j.appender.MainLog.DatePattern='yyyy-MM-dd'
```

Um alternativ hierzu eine Größenbeschränkung für das Protokoll (z. B. 8 MB) festzulegen, verwenden Sie in der Datei `log4cxx.properties` die folgenden Einträge:

```
log4j.appender.MainLog=org.apache.log4j.RollingFileAppender
log4j.appender.MainLog.MaxFileSize=8MB
```

Kapitel 5. Leistungsüberblick

Die tatsächliche Leistung beim Analysieren von Daten wird durch eine Reihe an Faktoren beeinflusst, von der Server- und Datenbankkonfiguration bis hin zum Sortieren von einzelnen Knoten in einem Stream. Im Allgemeinen erzielen Sie die beste Leistung, indem Sie die folgenden Punkte beachten:

- Speichern Sie Ihre Daten in einem Datenbankverwaltungssystem und verwenden Sie die SQL-Generierung, wann immer dies möglich ist.
- Verwenden Sie Hardware, die die Empfehlungen in Kapitel 2, „Architektur- und Hardwareempfehlungen“, auf Seite 7 erfüllt oder übertrifft.
- Stellen Sie sicher, dass die Leistungs- und Optimierungseinstellungen für Client und Server ordnungsgemäß konfiguriert sind. Beachten Sie: Wenn SPSS Modeler mit einer SPSS Modeler Server-Installation verbunden ist, setzen die Leistungs- und Optimierungseinstellungen des Servers die entsprechenden Einstellungen auf dem Client außer Kraft.
- Konzipieren Sie Streams für maximale Leistung.

In den folgenden Abschnitten finden Sie Informationen zu jedem einzelnen dieser Leistungsfaktoren.

Leistungs- und Optimierungseinstellungen für Ihren Server

Bestimmte IBM SPSS Modeler Server-Einstellungen können konfiguriert werden, um die Leistung zu optimieren. Sie können diese Einstellungen mithilfe der Schnittstelle von IBM SPSS Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten) anpassen. Weitere Informationen finden Sie in „IBM SPSS Modeler Server-Verwaltung“ auf Seite 38.

Diese Einstellungen sind unter der Überschrift **Leistung und Optimierung** im Konfigurationsfenster von IBM SPSS Modeler Administration Console zusammengefasst. Die Einstellungen sind bei den meisten Installationen für die optimale Leistung vorkonfiguriert. Je nach Hardware, Dataset und Streaminhalt kann es erforderlich sein, diese Einstellungen anzupassen. Weitere Informationen finden Sie in „Leistung und Optimierung“ auf Seite 42.

Leistungs- und Optimierungseinstellungen für Ihren Client

Die Leistungs- und Optimierungseinstellungen für den Client können über die Registerkarte "Optionen" des Dialogfelds "Streameigenschaften" aufgerufen werden. Wählen Sie zur Anzeige dieser Optionen Folgendes aus dem Client-Menü aus:

Tools > Streameigenschaften > Optionen > Optimierung

Sie können die Optimierungseinstellungen verwenden, um die Streamleistung zu optimieren. Beachten Sie, dass die Leistungs- und Optimierungseinstellungen auf IBM SPSS Modeler Server (sofern verwendet) alle entsprechenden Einstellungen auf dem Client außer Kraft setzen. Wenn diese Einstellungen im Server inaktiviert sind, kann der Client sie nicht aktivieren. Sind sie auf dem Server aktiviert, kann der Client sie jedoch inaktivieren.

Hinweis: Datenbankmodellierung und SQL-Optimierung erfordern, dass auf dem IBM SPSS Modeler-Computer IBM SPSS Modeler Server-Konnektivität aktiviert ist. Wenn diese Einstellung aktiviert ist, können Sie auf Datenbankalgorithmen zugreifen, SQL direkt aus IBM SPSS Modeler per Pushback übertragen und auf IBM SPSS Modeler Server zugreifen. Wählen Sie zur Überprüfung des aktuellen Lizenzstatus Folgendes im Menü von IBM SPSS Modeler aus.

Hilfe > Info... > Weitere Details

Wenn Konnektivität aktiviert ist, wird auf der Registerkarte "Lizenzstatus" die Option **Serveraktivierung** angezeigt.

Weitere Informationen finden Sie in „Verbinden mit IBM SPSS Modeler Server“ auf Seite 15.

Anmerkung: Ob SQL-Pushback und Optimierung unterstützt werden, hängt vom verwendeten Datenbanktyp ab. Neueste Informationen zu Datenbanken und ODBC-Treibern, die für die Verwendung mit IBM SPSS Modeler unterstützt werden und getestet wurden, finden Sie auf der unternehmensweiten Support-Site unter <http://www.ibm.com/support>.

Streamneuerstellung aktivieren. Wählen Sie diese Option aus, um die Streamneuerstellung in IBM SPSS Modeler zu aktivieren. Es stehen vier Neuerstellungsarten zur Verfügung, von denen Sie eine oder mehrere auswählen können. Die Streamneuerstellung ordnet die Knoten in einem Stream für eine effizientere Operation im Hintergrund neu an, ohne die Streamsemantik zu ändern.

- **SQL-Generierung optimieren.** Mit dieser Option können die Knoten im Stream so neu angeordnet werden, dass ein Pushback weiterer Operationen mithilfe der SQL-Generierung für die Ausführung in der Datenbank möglich ist. Wenn die Optimierungsfunktion einen Knoten findet, der nicht in SQL wiedergegeben werden kann, sucht sie nachgeordnete Knoten, die in SQL wiedergegeben und ohne Beeinträchtigung der Streamsemantik vor den problematischen Knoten verschoben werden können. Nicht nur, dass die Datenbank die Operationen effizienter als IBM SPSS Modeler durchführen kann, solche Pushbacks dienen auch dazu, die Größe des Datensets, das an IBM SPSS Modeler zur Verarbeitung zurückgegeben wird, zu reduzieren. So wird wiederum der Netzverkehr reduziert und die Geschwindigkeit der Streamoperationen erhöht. Das Kontrollkästchen **SQL generieren** muss aktiviert sein, damit die SQL-Optimierung funktionieren kann.
- **CLEM-Ausdruck optimieren.** Mit dieser Option kann das Optimierungsprogramm nach CLEM-Ausdrücken suchen, die vor der Ausführung des Streams vorverarbeitet werden können, damit die Verarbeitungsgeschwindigkeit erhöht wird. Ein einfaches Beispiel dafür ist, dass das Optimierungsprogramm bei einem Ausdruck wie $\log(\text{Gehalt})$ den tatsächlichen Gehaltswert berechnen und diesen dann für die Verarbeitung übergeben würde. Dies kann die Leistung von SQL-Pushback und IBM SPSS Modeler Server erhöhen.
- **Syntaxausführung optimieren.** Diese Methode der Streamneuerstellung erhöht die Effizienz von Operationen, die mehr als einen Knoten mit IBM SPSS Statistics-Syntax umfassen. Die Optimierung wird durch die Zusammenfassung der Syntaxbefehle in einer einzigen Operation erreicht, anstatt jeden einzelnen Syntaxbefehl als separate Operation auszuführen.
- **Andere Ausführung optimieren.** Diese Methode der Streamneuerstellung erhöht die Effizienz von Operationen, die nicht an die Datenbank delegiert werden können. Eine Optimierung wird dadurch erreicht, dass die Datenmenge im Stream so früh wie möglich reduziert wird. Während die Datenintegrität gewahrt bleibt, wird der Stream neu erstellt, um die Operationen näher zur Datenquelle zu bringen. So werden die Daten abwärts für kostenintensive Operationen, z. B. Joins, reduziert.

Parallele Verarbeitung aktivieren. Bei Ausführung auf einem Computer mit mehreren Prozessoren ermöglicht diese Option dem System, die Last auf diese Prozessoren zu verteilen, was zu einer schnelleren Leistung führen kann. Durch die Verwendung mehrerer Knoten oder der folgenden Einzelknoten können Sie von der Parallelverarbeitung profitieren: C5.0, Zusammenführen (nach Schlüssel), Sortieren, Klasse (Rangbildungsmethode und N-Perzentil-Methode) sowie Aggregieren (unter Verwendung mindestens eines Schlüsselfelds).

SQL exportieren. Wählen Sie diese Option aus, um die SQL-Generierung zu aktivieren. Dabei wird zugelassen, dass Streamoperationen per Pushback an die Datenbank zurückübertragen werden, indem SQL-Code zur Generierung der Ausführungsprozesse verwendet wird, was zu einer Leistungssteigerung führen kann. Zur weiteren Steigerung der Leistung kann außerdem **SQL-Generierung optimieren** ausgewählt werden. Dadurch wird die Anzahl der Operationen, die per Pushback an die Datenbank zurückübertragen werden, maximiert. Wenn Operationen für einen Knoten per Pushback an die Datenbank zurückübertragen wurden, wird der Knoten bei der Ausführung des Streams violett gekennzeichnet.

- **Datenbankcaching.** Bei Streams, die in der Datenbank auszuführendes SQL generieren, muss nicht das Dateisystem als Cache verwendet werden, sondern die Daten können mitten im Stream in einer temporären Tabelle in der Datenbank zwischengespeichert werden. Bei einer Kombination mit SQL-Optimierung kann dies zu einer erheblichen Leistungssteigerung führen. So kann beispielsweise die Ausgabe aus einem Stream, der mehrere Tabellen zusammenfasst, um eine Data-Mining-Ansicht zu erstellen, im Cache gespeichert und bei Bedarf wieder verwendet werden. Klicken Sie bei aktiviertem Datenbankcaching einfach mit der rechten Maustaste auf einen Knoten (keinen Endknoten), um die Daten an dieser Stelle im Cache zu speichern, und der Cache wird bei der nächsten Ausführung des Streams automatisch unmittelbar in der Datenbank erstellt. Dadurch kann SQL für nachgeordnete Knoten generiert werden, was zu weiteren Leistungsverbesserungen führt. (Alternativ kann diese Option bei Bedarf inaktiviert werden, beispielsweise wenn es aufgrund von Richtlinien oder Berechtigungen nicht möglich ist, Daten in die Datenbank zu schreiben. Wenn Datenbankcaching oder SQL-Optimierung nicht aktiviert ist, wird der Cache stattdessen in das Dateisystem geschrieben.
- **Unproblematische Umwandlung verwenden.** Mit dieser Option können Daten von Zeichenfolgen in Zahlen oder von Zahlen in Zeichenfolgen konvertiert werden, sofern sie in einem geeigneten Format gespeichert sind. Wenn die Daten beispielsweise als Zeichenfolge in der Datenbank gespeichert sind, tatsächlich jedoch eine sinnvolle Zahl enthalten, können die Daten beim Eintreten des Pushbacks zur Verwendung konvertiert werden.

Anmerkung: Aufgrund kleinerer Unterschiede in der SQL-Implementierung weichen die Ergebnisse der in einer Datenbank ausgeführten Streams möglicherweise leicht von den Ergebnissen bei Ausführung in IBM SPSS Modeler ab. Aus ähnlichen Gründen können diese Unterschiede auch je nach Datenbankanbieter variieren.

Datenbankverwendung und -optimierung

Datenbankserver. Wenn möglich, erstellen Sie eine dedizierte Datenbankinstanz für das Data-Mining, so dass der Produktionsserver nicht von IBM SPSS Modeler-Anfragen beeinflusst wird. Die von IBM SPSS Modeler generierten SQL-Anweisungen können einen großen Teil der Ressourcen in Anspruch nehmen. Auf dem IBM SPSS Modeler Server-Computer können mehrere Aufgaben SQL in derselben Datenbank ausführen.

Datenbankinternes Mining. Viele Datenbankanbieter bieten für ihre Produkte Data-Mining-Erweiterungen an. Mit diesen Erweiterungen können Sie Data-Mining-Aufgaben (z. B. Modellerstellung oder Scoring) auf dem Datenbankserver oder einem separaten dedizierten Server ausführen. Die datenbankinternen Mining-Funktionen von IBM SPSS Modeler ergänzen und erweitern die SQL-Generierungsfähigkeiten des Produkts und bieten auf diese Weise eine Möglichkeit, die herstellereigene Datenbankerweiterungen zu betreiben. In manchen Fällen kann diese Herangehensweise den manchmal kostenintensiven Aufwand für die Datenübertragung zwischen IBM SPSS Modeler und der Datenbank reduzieren. Auch das Datenbankcaching kann dazu beitragen, diese Vorteile weiter auszubauen. Weitere Informationen finden Sie in der Datei *DatabaseMiningGuide.pdf*, die über Ihren eImage-Download verfügbar ist.

SQL-Optimierung

Die beste Leistung erzielen Sie, wenn Sie die generierte Menge an SQL voll ausschöpfen und somit die Leistungsfähigkeit und Skalierbarkeit der Datenbank nutzen. Nur die Teile des Streams, die nicht in SQL kompiliert werden können, sollten in IBM SPSS Modeler Server ausgeführt werden. Weitere Informationen finden Sie in Kapitel 6, „SQL-Optimierung“, auf Seite 63.

Hochladen von dateibasierten Daten

Daten, die nicht in einer Datenbank gespeichert sind, können nicht von der SQL-Optimierung profitieren. Wenn sich die zu analysierenden Daten noch nicht in einer Datenbank befinden, können Sie sie mit einem Datenbankausgabeknoten hochladen. Sie können diesen Knoten außerdem verwenden, um Datenzwischenmengen aus der Datenvorbereitung und die Ergebnisse der Bereitstellung zu speichern.

IBM SPSS Modeler kann mit den externen Ladeprogrammen für viele gängige Datenbanksysteme zusammenarbeiten. In der Software sind mehrere Scripts enthalten. Diese stehen (mit Dokumentation) im Unterverzeichnis */scripts* unter dem IBM SPSS Modeler-Installationsordner zur Verfügung.

In der folgenden Tabelle wird der potenzielle Leistungsvorteil des Massenladens gezeigt. In der Abbildung wird dargestellt, wie viel Zeit zum Exportieren von 250.000 Datensätzen und 21 Feldern in eine Oracle-Datenbank erforderlich ist. Das externe Ladeprogramm ist das Oracle-Dienstprogramm `sqlldr`.

Tabelle 2. Leistungsvorteil des Massenladens

Exportoption	Zeit (in Sekunden)
Standard (ODBC)	409
Massenladen über ODBC	52
Massenladen über externes Ladeprogramm	33

Kapitel 6. SQL-Optimierung

Eine der wichtigsten Funktionen von IBM SPSS Modeler besteht in der Möglichkeit, mehrere Datenvorbereitungs- und Mining-Operationen direkt in der Datenbank durchzuführen. Durch die Generierung von SQL-Code, der für die Ausführung zurück in die Datenbank übertragen wird (Pushback), können viele Operationen, wie die Stichprobenziehung, Sortierung, Ableitung neuer Felder und bestimmte Formen der Diagrammerstellung, anstatt auf dem IBM SPSS Modeler- oder IBM SPSS Modeler Server-Computer in der Datenbank durchgeführt werden. Wenn Sie mit großen Datasets arbeiten, können diese **Pushbacks** die Leistung auf unterschiedliche Weise erheblich steigern.

- Durch Reduzierung der Größe des Ergebnisses, das aus dem Datenbankverwaltungssystem nach IBM SPSS Modeler übertragen werden soll. Wenn große Ergebnisssets mithilfe eines ODBC-Treibers gelesen werden, kann es zu Netz-E/A- oder Treiberfehlern kommen. Aus diesem Grund sind Zeilen- und Spaltenauswahl und -aggregation (Auswahl-, Stichproben-, Aggregatknotten), mit denen üblicherweise die Größe des zu übertragenden Datasets verringert wird, die Operationen, die am meisten von der SQL-Optimierung profitieren. Daten können außerdem an kritischen Punkten im Stream (z. B. nach einem Zusammenführungs- oder Auswahlknotten) in einer temporären Tabelle in der Datenbank zwischengespeichert werden, um die Leistung weiter zu steigern.
- Durch Nutzung der Leistung und Skalierbarkeit der Datenbank. Die Effizienz wird erhöht, da ein Datenbankverwaltungssystem häufig die Vorteile paralleler Verarbeitung, leistungsstärkerer Hardware, einer ausgereifteren Verwaltung des Speicherplatzes und des Vorhandenseins von Indizes nutzen kann.

Angesichts dieser Vorteile wurde IBM SPSS Modeler so entworfen, dass es die Menge an SQL, das von den einzelnen Streams generiert wird, maximiert, sodass nur Operationen, die sich nicht in SQL kompilieren lassen, von IBM SPSS Modeler Server ausgeführt werden. Aufgrund der Beschränkungen der Ausdrucksmöglichkeiten in der Standard-SQL (SQL-92) werden bestimmte Operationen nicht unterstützt. Weitere Informationen finden Sie in „Tipps für die Maximierung der SQL-Generierung“ auf Seite 67.

Anmerkung: Beachten Sie folgende Hinweise bezüglich SQL:

- Aufgrund kleinerer Unterschiede in der SQL-Implementierung weichen die Ergebnisse der in einer Datenbank ausgeführten Streams möglicherweise ein wenig von den bei der Ausführung in IBM SPSS Modeler erzielten Ergebnissen ab. Diese Unterschiede können aus ähnlichen Gründen auch je nach Datenbankanbieter variieren. Je nach der Datenbankkonfiguration für die Unterscheidung zwischen Groß- und Kleinschreibung beim Vergleich und der Sortierung von Zeichenfolgen können IBM SPSS Modeler Streams, die unter Verwendung von SQL-Pushback ausgeführt werden, zu anderen Ergebnissen führen als Streams, die ohne SQL-Pushback ausgeführt werden. Wenden Sie sich an den Datenbankadministrator, wenn Sie Unterstützung bei der Konfiguration Ihrer Datenbank benötigen. Für größtmögliche Kompatibilität mit IBM SPSS Modeler sollte beim Vergleich von Datenbankzeichenfolgen zwischen Groß- und Kleinschreibung unterschieden werden.
- Datenbankmodellierung und SQL-Optimierung erfordern, dass IBM SPSS Modeler Server-Konnektivität auf dem IBM SPSS Modeler-Computer aktiviert ist. Wenn diese Einstellung aktiviert ist, können Sie auf Datenbankalgorithmen zugreifen, SQL direkt aus IBM SPSS Modeler per Pushback übertragen und auf IBM SPSS Modeler Server zugreifen. Wählen Sie zur Überprüfung des aktuellen Lizenzstatus die folgenden Optionen aus dem IBM SPSS Modeler-Menü aus.
- Wenn IBM SPSS Modeler zum Generieren von SQL verwendet wird, ist es möglich, dass das Ergebnis bei Verwendung von SQL-Pushback auf einigen Plattformen (z. B. Linux/zLinux) nicht konsistent mit IBM SPSS Modeler (native Version) ist. Die Ursache ist, dass Gleitkommata auf verschiedenen Plattformen unterschiedlich gehandhabt werden.

Hilfe > Info... > Weitere Details

Wenn Konnektivität aktiviert ist, wird auf der Registerkarte "Lizenzstatus" die Option **Serveraktivierung** angezeigt.

Weitere Informationen finden Sie in „Verbinden mit IBM SPSS Modeler Server“ auf Seite 15.

Anmerkung: Wenn Streams in einer Netezza-Datenbank ausgeführt werden, werden die Angaben zu Datum und Uhrzeit aus der betreffenden Datenbank übernommen. Diese Werte können von den lokalen Werten für Datum und Uhrzeit bzw. den Werten in IBM SPSS Modeler Server abweichen, wenn sich die Datenbank beispielsweise auf einem Computer in einem anderen Land bzw. einer anderen Zeitzone befindet.

Datenbankanforderungen

Neueste Informationen zu Datenbanken und ODBC-Treibern, die für die Verwendung mit IBM SPSS Modeler unterstützt werden und getestet wurden, finden Sie in den Produktkompatibilitätsdiagrammen auf der unternehmensweiten Support-Site unter <http://www.ibm.com/support>.

Beachten Sie, dass durch Verwendung von Datenbankmodellierung eventuell weitere Leistungsverbesserungen möglich sind.

ODBC-Treibersetup

Um sicherzustellen, dass Zeitangaben (z. B. HH:MM:SS) bei der Verwendung von SQL 2012 unter Windows 8-Systemen mit 32 Bit ordnungsgemäß verarbeitet werden, sollten Sie beim Konfigurieren Ihres ODBC-Verbindungsprotokolltreibers für SQL Server sowohl die Option **Enable Quoted Identifiers** als auch die Option **Fetch TWFS as Time** auswählen.

Funktionsweise der SQL-Generierung

Die ersten Fragmente eines Streams, der von den Quellenknoten der Datenbank ausgeht, sind die Hauptziele für die SQL-Generierung. Wenn ein Knoten gefunden wird, der sich nicht in SQL kompilieren lässt, werden die Daten aus der Datenbank extrahiert und die weitere Verarbeitung wird von IBM SPSS Modeler Server durchgeführt.

Während der Streamvorbereitung und vor der Ausführung findet die SQL-Generierung wie folgt statt.

- Der Server ordnet Streams neu, um nachgeordnete Knoten in die "SQL-Zone" zu verschieben, wenn dies erwiesenermaßen problemlos möglich ist. (Diese Funktion kann auf dem Server inaktiviert werden.)
- SQL-Ausdrücke werden schrittweise von den Quellenknoten zu den Endknoten hin erstellt. Diese Phase endet, wenn ein Knoten gefunden wird, der nicht in SQL konvertiert werden kann, oder der Endknoten (z. B. Tabellenknoten oder Diagrammknoten) in SQL konvertiert wird. Am Ende dieser Phase wird jeder Knoten mit einer SQL-Anweisung beschriftet, falls der Knoten und seine Vorgänger eine SQL-Entsprechung aufweisen.
- Das SQL wird ausgehend von den Knoten mit den kompliziertesten SQL-Äquivalenten zurück zu den Quellenknoten auf Validität überprüft. Das SQL, das erfolgreich validiert wurde, wird für die Ausführung ausgewählt.
- Knoten, für die alle Operationen SQL generiert haben, werden im Streamerstellungsbereich mit violetter Markierung angezeigt. Auf der Grundlage der Ergebnisse kann es angebracht sein, den Stream an den entsprechenden Stellen weiter zu reorganisieren, um die Datenbankausführung bestmöglich zu nutzen. Weitere Informationen finden Sie in „Tipps für die Maximierung der SQL-Generierung“ auf Seite 67.

Auftreten von Verbesserungen

SQL-Optimierung verbessert die Leistung in einer Reihe von Datenoperationen:

- **Joins (Zusammenführen nach Schlüssel).** Joinoperationen können die Optimierung innerhalb von Datenbanken steigern.

- **Aggregation.** Die Aggregat-, Verteilungs- und Netzdiagrammknoten verwenden jeweils Aggregation zum Erstellen der Ergebnisse. Zusammengefasste Daten benötigen erheblich weniger Bandbreite als die ursprünglichen Daten.
- **Auswahl.** Die Auswahl von Datensätzen auf der Grundlage bestimmter Kriterien verringert die Menge der Datensätze.
- **Sortieren.** Das Sortieren von Datensätzen ist eine ressourcenintensive Aktivität, die sich effizienter in einer Datenbank ausführen lässt.
- **Feldableitung.** Neue Felder lassen sich effizienter in einer Datenbank generieren.
- **Feldprojektion.** IBM SPSS Modeler Server extrahiert nur Felder aus der Datenbank, die für die nachfolgende Verarbeitung erforderlich sind. Dadurch werden die Anforderungen an Bandbreite und Arbeitsspeicher minimiert. Dasselbe gilt auch für überflüssige Felder in Flatfiles: Der Server muss die überflüssigen Felder zwar lesen, ordnet ihnen jedoch keinen Speicher zu.
- **Scoring.** SQL lässt sich aus Entscheidungsbäumen, Regelsets, linearer Regression und faktorgenerierten Modellen generieren.

Beispiel für SQL-Generierung

Mit dem folgenden Stream werden drei Datenbanktabellen mithilfe von Schlüsseloperationen zusammengefügt und dann sowohl eine Aggregation als auch eine Sortierung durchgeführt.

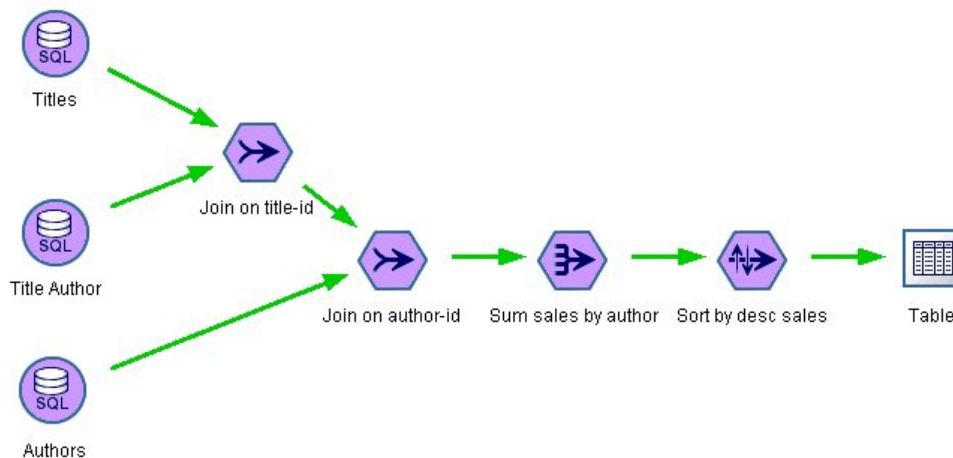


Abbildung 3. Optimierter Stream mit violetten Knoten, die SQL-Pushbacks (in der Datenbank durchgeführte Operationen) anzeigen

Generiertes SQL

Das für diesen Stream generierte SQL lautet wie folgt:

```
SELECT
  T2. au_lname AS C0,
  T2. au_fname AS C1,
  SUM({fn CONVERT(T0. ytd_sales ,SQL_BIGINT)}) AS C2
FROM
  dbo . titles T0,
  dbo . titleauthor T1,
  dbo . authors T2
WHERE
  (T0. title_id = T1. title_id )
  AND (T1. au_id = T2. au_id )
GROUP BY T2. au_lname ,T2. au_fname
ORDER BY 3 DESC
```

Ausführen des Streams

Wenn der Stream mit einem Datenbankexportknoten beendet wird, kann der gesamte Stream in der Datenbank ausgeführt werden.

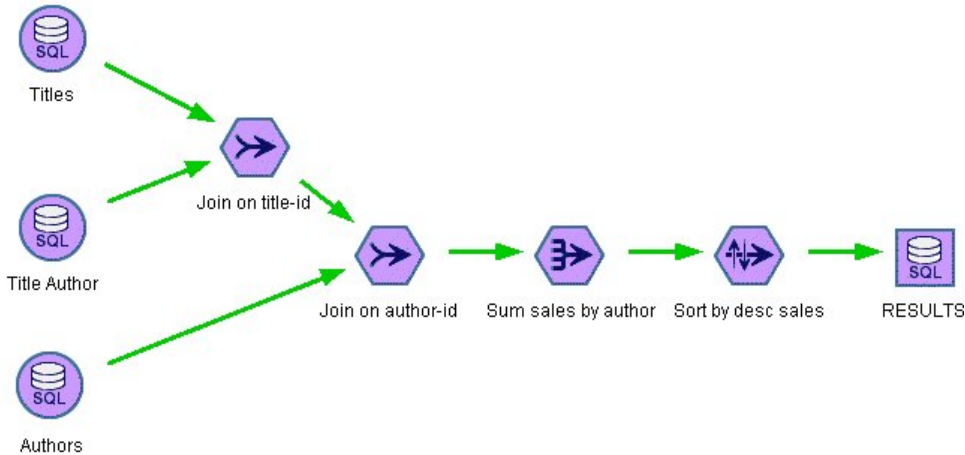


Abbildung 4. Gesamter, in der Datenbank ausgeführter Stream

Konfigurieren der SQL-Optimierung

1. Installieren Sie einen ODBC-Treiber und konfigurieren Sie eine Datenquelle für die zu verwendende Datenbank. Weitere Informationen finden Sie in „Datenzugriff“ auf Seite 11.
2. Erstellen Sie einen Stream, der einen Quellenknoten zum Extrahieren von Daten aus dieser Datenbank verwendet.
3. Prüfen Sie gegebenenfalls, ob die SQL-Generierung auf Client und Server aktiviert ist. Standardmäßig ist sie für beide aktiviert.

So aktivieren Sie SQL-Optimierung auf dem Client

1. Wählen Sie im Menü "Tools" die Option **Streameigenschaften > Optionen** aus.
2. Klicken Sie dann auf die Registerkarte "Optimierung". Wählen Sie **SQL generieren**, um die SQL-Optimierung zu aktivieren. Optional können Sie andere Einstellungen auswählen, um die Leistung zu verbessern. Weitere Informationen finden Sie in „Leistungs- und Optimierungseinstellungen für Ihren Client“ auf Seite 59.

So aktivieren Sie SQL-Optimierung auf dem Server

Da die Servereinstellungen alle auf dem Client angegebenen Spezifikationen überschreiben, müssen die Serverkonfigurationseinstellungen **Streamneuerstellung** und **Automatische SQL-Generierung** jeweils aktiviert sein. Weitere Informationen zum Ändern der IBM SPSS Modeler Server-Einstellungen finden Sie im Abschnitt „Leistung und Optimierung“ auf Seite 42. Wenn diese Einstellungen auf dem Server inaktiviert sind, kann der Client sie nicht aktivieren. Sind sie auf dem Server aktiviert, kann der Client sie jedoch inaktivieren.

So aktivieren Sie die Optimierung beim Scoring von Modellen

Zum Zwecke des Scorings muss die SQL-Generierung für jeden Modellierungsknoten separat aktiviert werden, unabhängig von den Einstellungen auf der Server- bzw. Client-Ebene. Das liegt daran, dass manche Modelle extrem komplexe SQL-Ausdrücke generieren, die in der Datenbank unter Umständen nicht effektiv evaluiert werden. In der Datenbank können Fehler auftreten, wenn das generierte SQL ausgeführt werden soll, weil das SQL zu groß oder zu komplex ist.

Unter Umständen erfordert es einige Anläufe, um festzustellen, ob die SQL-Generierung die Leistung eines bestimmten Modells verbessert. Dies geschieht auf der Registerkarte "Einstellungen", nachdem ein generiertes Modell zu einem Stream hinzugefügt wurde.

Vorschau für generiertes SQL

Sie können das generierte SQL im Nachrichtenprotokoll in der Vorschau anzeigen, bevor sie in der Datenbank ausgeführt wird. Dies ist u. U. für das Debugging nützlich und ermöglicht außerdem den Export des generierten SQL, um diese zu einem späteren Zeitpunkt in der Datenbank zu bearbeiten bzw. auszuführen. Außerdem wird angezeigt, welche Knoten per Pushback an die Datenbank zurückübertragen werden. Dadurch können Sie leichter ermitteln, ob sich durch eine Neuordnung des Streams eine Leistungssteigerung erzielen lässt.

1. Stellen Sie sicher, dass die Optionen **SQL während der Streamausführung im Nachrichtenprotokoll anzeigen** und **Einzelheiten zur SQL-Generierung während der Streamvorbereitung im Nachrichtenprotokoll anzeigen** im Dialogfeld "Benutzeroptionen" ausgewählt sind. Weitere Informationen finden Sie in „Leistungs- und Optimierungseinstellungen für Ihren Client“ auf Seite 59.
2. Wählen Sie im Streamerstellungsbereich den Knoten oder Stream aus, den Sie in der Vorschau anzeigen möchten.
3. Klicken Sie in der Symbolleiste auf die Schaltfläche für die **SQL-Vorschau**.
Alle Knoten, für die SQL generiert wird (und die bei der Ausführung des Streams per Pushback in die Datenbank zurückübertragen werden) werden im Streamerstellungsbereich violett gekennzeichnet.
4. Um das generierte SQL in der Vorschau anzuzeigen, wählen Sie in den Menüs Folgendes aus:
Tools > Streameigenschaften > Nachrichten...

Anzeigen von SQL für Modellnuggets

Bei einigen Modellen kann SQL für das Modellnugget generiert werden, wobei die Modellscoring-Phase per Pushback an die Datenbank zurückgegeben wird. Diese Funktion dient nicht in erster Linie zur Leistungsverbesserung, sondern dazu, dass das gesamte SQL von Streams, die diese Nuggets enthalten, per Pushback zurückgeführt werden kann. Weitere Informationen finden Sie in „Knoten zur Unterstützung der SQL-Generierung“ auf Seite 69.

Führen Sie folgende Schritte aus, um das SQL für ein Modellnugget anzuzeigen, das SQL-Generierung unterstützt:

1. Wählen Sie die Registerkarte "Einstellungen" auf dem Modellnugget.
2. Wählen Sie nach Bedarf eine der Optionen **Mit (Ohne) Unterstützung für fehlende Werte generieren** oder **SQL für dieses Modell generieren**.
3. Wählen Sie im Menü des Modellnuggets:
Datei > SQL exportieren
4. Speichern Sie die Datei.
5. Öffnen Sie die Datei, um das SQL anzuzeigen.

Tipps für die Maximierung der SQL-Generierung

Um die bestmögliche Leistungssteigerung aus der SQL-Optimierung zu erzielen, sollten Sie auf folgende Punkte achten:

Streamreihenfolge. Die SQL-Generierung wird möglicherweise angehalten, wenn die Funktion des Knotens keine semantische Entsprechung in SQL aufweist, da die Data-Mining-Funktion von IBM SPSS Modeler umfassender ist als die herkömmlichen Datenverarbeitungsoperationen, die von Standard-SQL unterstützt werden. In diesem Fall wird die SQL-Generierung auch für alle nachgeordneten Knoten unterdrückt. Daher können Sie die Leistung eventuell signifikant steigern, wenn Sie die Knoten neu anordnen, um die Operationen, die SQL stoppen, so weit unten im Stream platzieren wie möglich. Das

SQL-Optimierungsprogramm kann Neuordnungen bis zu einem bestimmten Grad automatisch durchführen (sie müssen jedoch sicherstellen, dass die Streamneuerstellung aktiviert ist), es sind jedoch eventuell weitere Verbesserungen möglich. Ein guter Kandidat dafür ist der Auswahlknoten, der häufig nach weiter vorne im Stream verschoben werden kann. Weitere Informationen finden Sie in „Knoten zur Unterstützung der SQL-Generierung“ auf Seite 69.

CLEM-Ausdrücke. Wenn ein Stream nicht neu geordnet werden kann, können Sie eventuell Knotenoptionen oder CLEM-Ausdrücke ändern oder anderweitig die Art und Weise, wie die Operation ausgeführt wird, umgestalten, sodass die SQL-Generierung nicht mehr behindert wird. Ableitungsknoten, Auswahlknoten und ähnliche Knoten können allgemein in SQL gerendert werden, vorausgesetzt, dass es für alle CLEM-Ausdrucks-Operatoren SQL-Entsprechungen gibt. Die meisten Operatoren können gerendert werden. Es gibt jedoch eine Reihe von Operationen, die die SQL-Generierung blockieren (insbesondere die Sequenzfunktionen [“@-Funktionen“]). Manchmal wird die Generierung angehalten, da die generierte Abfrage so komplex wurde, sodass die Datenbank nicht mehr damit umgehen kann. Weitere Informationen finden Sie in „CLEM-Ausdrücke und Operatoren, die die SQL-Generierung unterstützen“ auf Seite 73.

Mehrere Quellenknoten. Wenn ein Stream mehrere Datenbankquellenknoten aufweist, wird die SQL-Generierung auf jeden Eingabezweig unabhängig angewendet. Wenn die Generierung in einem Zweig angehalten wird, kann sie in einem anderen fortgesetzt werden. Wenn zwei Zweige zusammengeführt werden (und beide Zweige bis zur Zusammenführung in SQL ausgedrückt werden können), kann die Zusammenführung als solche oft durch eine Datenbankverknüpfung ersetzt werden und die Generierung kann nachgeordnet fortgesetzt werden.

Datenbankalgorithmus. Die Modellschätzung wird fast immer auf IBM SPSS Modeler Server und nicht in der Datenbank durchgeführt, es sei denn, es werden datenbankeigene Algorithmen von Microsoft, IBM bzw. Oracle verwendet.

Scoring-Modelle. Für einige Modelle wird datenbankinternes Scoring durch Rendern des generierten Modells in SQL unterstützt. Einige Modelle generieren jedoch extrem komplexe SQL-Ausdrücke, die nicht immer effektiv in der Datenbank evaluiert werden. Aus diesem Grund muss die SQL-Generierung separat für die einzelnen Modellknoten aktiviert werden. Wenn sich herausstellt, dass der Modellknoten die SQL-Generierung behindert, wechseln Sie zur Registerkarte "Einstellungen" im Dialogfeld des Knotens und wählen Sie **SQL für dieses Modell generieren** aus (bei einigen Modellen stehen Ihnen möglicherweise weitere Optionen zur Steuerung der SQL-Generierung zur Verfügung. Führen Sie Tests aus, um sicherzustellen, dass die Option für Ihre Anwendung nützlich ist. Weitere Informationen finden Sie in „Knoten zur Unterstützung der SQL-Generierung“ auf Seite 69.

Wenn Sie Knoten testen, um festzustellen, ob SQL-Generierung für Modelle effektiv arbeitet, sollten Sie zuvor alle Streams von IBM SPSS Modeler speichern. Einige Datenbanksysteme können hängen bleiben, wenn sie versuchen, das (möglicherweise komplexe) generierte SQL zu verarbeiten, und es kann erforderlich sein, IBM SPSS Modeler über den Windows Task-Manager zu schließen.

Datenbankcaching. Wenn Sie einen Knotencache zum Speichern von Daten an kritischen Punkten im Stream verwenden (beispielsweise nach einem Zusammenführungs- oder Aggregatknoten), müssen Sie sicherstellen, dass neben der SQL-Optimierung auch Datenbankcaching aktiviert ist. Dadurch können die Daten in den meisten Fällen eine temporäre Tabelle in der Datenbank (statt des Dateisystems) als Cache fungieren. Weitere Informationen finden Sie in „Konfigurieren der SQL-Optimierung“ auf Seite 66.

Herstellerspezifisches SQL. Der größte Teil des generierten SQL entspricht den Standards (SQL-92), es werden jedoch, wo angemessen, auch einige nicht standardmäßige, herstellerspezifische Funktionen genutzt. Der Grad der SQL-Optimierung kann, je nach Datenbankquelle, schwanken.

Knoten zur Unterstützung der SQL-Generierung

In den folgenden Tabellen werden die Knoten für die Data-Mining-Operationen beschrieben, die eine SQL-Generierung unterstützen. Mit Ausnahme der Datenbankmodellknoten wird die SQL-Generierung von Knoten, die nicht in diesen Tabellen aufgeführt sind, nicht unterstützt.

Sie können das zu generierende SQL anzeigen, bevor Sie es ausführen. Weitere Informationen finden Sie in „Vorschau für generiertes SQL“ auf Seite 67.

Tabelle 3. Quellen

Knoten zur Unterstützung der SQL-Generierung	Hinweise
Datenbank	Dieser Knoten wird verwendet, um Tabellen und Ansichten für eine weitere Analyse anzugeben. Dieser Knoten ermöglicht die Eingabe von SQL-Abfragen. Vermeiden Sie Ergebnissets mit doppelten Spaltennamen. Weitere Informationen finden Sie in „Schreiben von SQL-Abfragen“ auf Seite 75.

Tabelle 4. Datensatzoperationen

Knoten zur Unterstützung der SQL-Generierung	Hinweise
Auswählen	Unterstützt die Generierung nur, wenn SQL-Generierung für den select-Ausdruck selbst unterstützt wird (siehe Ausdrücke unten). Wenn Felder mit Nullwerten vorliegen, führt die SQL-Generation für "discard" nicht zu denselben Ergebnissen wie in IBM SPSS Modeler.
Stichprobe	Die einfache Stichprobenziehung unterstützt die SQL-Generierung bis zu einem gewissen Grad je nach Datenbank. Siehe Tabelle 5 auf Seite 70.
Aggregieren	Die Unterstützung der SQL-Generierung für die Aggregation hängt vom Datenspeichertyp ab. Siehe Tabelle 6 auf Seite 70.
RFM-Aggregat	Unterstützt die Generierung, außer beim Speichern des Datums der zweit- oder drittaktuellsten Transaktion oder wenn nur aktuelle Transaktionen eingeschlossen werden. Die Aufnahme aktueller Transaktionen ist jedoch möglich, wenn für die Funktion <code>datetime_date(JAHR,MONAT,TAG)</code> ein Pushback-Vorgang durchgeführt wird.
Sortieren	
Zusammenführen	Kein SQL generiert für Zusammenführen nach Schlüssel. Ein Zusammenführen nach Schlüssel mit vollständigem oder partiellem Outer Join wird nur unterstützt, falls sie auch von der Datenbank/dem Treiber unterstützt wird. Nicht übereinstimmende Eingabefelder können mithilfe eines Filterknotens oder über die Registerkarte "Filter" eines Quellenknotens umbenannt werden. Unterstützt SQL-Generierung für die Zusammenführung nach Bedingung. Für alle Zusammenführungstypen gilt, dass <code>SQL_SP_EXISTS</code> nicht unterstützt wird, wenn die Eingaben aus verschiedenen Datenbanken stammen.
Anhängen	Unterstützt die Generierung bei unsortierten Eingaben. Anmerkung: Die SQL-Optimierung ist nur möglich, wenn Ihre Eingaben jeweils dieselbe Anzahl Spalten haben.
Duplikat	Die SQL-Optimierung wird bei einem Duplikatknoten, bei dem der Modus (Standardmodus) Zusammengesetzten Datensatz für jede Gruppe erstellen ausgewählt ist, nicht unterstützt.

Tabelle 5. Unterstützung der SQL-Generierung im Stichprobenknoten für einfache Stichprobenziehung.

Modus	Stichprobe	Max. Größe	Startwert	DB2 for z/OS	DB2 for OS/400	DB2 for Win/UNIX	Netezza	Oracle	SQL Server	Teradata	
Einschließen	Erster	entfällt		J	J	J	J	J	J	J	
	1 in n	Aus		J	J	J	J	J		J	
		max		J	J	J	J	J		J	
	Zufällig %	Aus	Aus	Aus	J		J	J	J		J
			Ein	J		J		J			
		max	Aus	J		J	J	J	J		J
			Ein	J		J		J			
		Verwerfen	Erster	Aus					J	J	
max								J	J		
1 in n	Aus			J	J	J	J	J		J	
	max			J	J	J	J	J		J	
Zufällig %	Aus		Aus	Aus	J		J	J	J		J
			Ein	J		J		J			
	max		Aus	J		J	J	J	J		J
			Ein	J		J		J			

Tabelle 6. Unterstützung der SQL-Generierung im Aggregatknoten.

Speichertyp	Summe	Mittelwert	Min	Max	Std. Abw.	Median	Anzahl	Varianz	Perzentil
Ganzzahl	J	J	J	J	J		J	J	
Reelle Zahl	J	J	J	J	J		J	J	
Datum			J	J			J		
Zeit			J	J			J		
Zeitmarke			J	J			J		
Zeichenfolge			J	J			J		

Tabelle 7. Feldoperationen

Knoten zur Unterstützung der SQL-Generierung	Hinweise
Typ	Unterstützt SQL-Generierung, wenn der Typknoten instanziiert wird und keine ABORT- oder WARN-Typprüfung angegeben ist.
Filter	
Ableitung	Unterstützt SQL-Generierung, wenn das für den Ableitungsausdruck generierte SQL unterstützt wird (siehe Ausdrücke unten).
Ensemble	Unterstützt SQL-Generierung für stetige Ziele. Bei anderen Zielen wird die Generierung nur unterstützt, wenn die Ensemble-Methode "Höchste Konfidenz hat Vorrang" verwendet wird.
Füller	Unterstützt SQL-Generierung, wenn das für den "derive"-Ausdruck generierte SQL unterstützt wird (siehe Ausdrücke unten).

Tabelle 7. Feldoperationen (Forts.)

Knoten zur Unterstützung der SQL-Generierung	Hinweise
Anonymisierung	Unterstützt SQL-Generierung für stetige Ziele und die partielle SQL-Generierung für nominale und Flagziele.
Umcodierung	
Klassierung	Unterstützt SQL-Generierung, wenn die Klassiermethode "n-tile (gleiche Anzahl)" verwendet wird und die Option "Von Registerkarte "Klassenwerte" lesen, sofern verfügbar" ausgewählt ist. Anmerkung: Infolge von (durch die Art der Verteilung von Daten in Klassenfeldern verursachten) Unterschieden in der Art, wie Klassengrenzen berechnet werden, stellen Sie möglicherweise Unterschiede bei der Klassierungsausgabe zwischen Ergebnissen der normalen Streamausführung und SQL-Pushback-Ergebnissen fest. Um dies zu vermeiden, verwenden Sie die Perzentilmethode Datensatzanzahl und N-Perzentile des Typs Zu nächstem hinzu oder In aktuellem beibehalten , um die größte Übereinstimmung zwischen den beiden Methoden der Streamausführung zu erhalten.
RFM-Analyse	Unterstützt SQL-Generierung, wenn die Option "Von Registerkarte "Klassenwerte" lesen, sofern verfügbar" ausgewählt ist, jedoch unterstützen dies nachgeordnete Knoten nicht.
Partitionierung	Unterstützt SQL-Generierung für die Zuweisung von Datensätzen zu Partitionen.
Dichotom	
Umstrukturierung	

Tabelle 8. Grafiken

Knoten zur Unterstützung der SQL-Generierung	Hinweise
Diagrammtafel	SQL-Generierung wird für die folgenden Diagrammtypen unterstützt: Fläche, 3-D-Fläche, Balken, 3-D-Balken, Balken für Häufigkeiten, Heat-Map, Kreis, 3-D-Kreis, Kreis für Häufigkeiten. Bei Histogrammen wird SQL-Generierung nur für kategoriale Daten unterstützt. Die SQL-Generierung wird für die Animation auf der Diagrammtafel nicht unterstützt.
Verteilung	
Internet	
Evaluierung	

Bei einigen Modellen kann SQL für das Modellnugget generiert werden, wobei die Modellscoring-Phase per Pushback an die Datenbank zurückgegeben wird. Diese Funktion dient nicht in erster Linie zur Leistungsverbesserung, sondern dazu, dass das gesamte SQL von Streams, die diese Nuggets enthalten, per Pushback zurückgeführt werden kann. Weitere Informationen finden Sie in „Anzeigen von SQL für Modellnuggets“ auf Seite 67.

Tabelle 9. Modellnuggets

SQL-Generierung unterstützendes Modellnugget	Hinweise
C&R-Baum	Unterstützt SQL-Generierung für die Option für einzelne Bäume, nicht jedoch für Boosting, Bagging oder für große Datasetoptionen.
QUEST	
CHAID	

Tabelle 9. Modellnuggets (Forts.)

SQL-Generierung unterstützendes Modellnugget	Hinweise
C5.0	
Entscheidungsliste	
Linear	Unterstützt SQL-Generierung für die Standardmodelloption, nicht jedoch für Boosting, Bagging oder große Datasetoptionen.
Neuronales Netz	Unterstützt SQL-Generierung für die Standardmodelloption (nur Mehrschicht-Perzeptron), nicht jedoch für Boosting, Bagging oder große Datasetoptionen.
Faktor/PCA	
Logistisch	Unterstützt SQL-Generierung für das multinomiale, nicht jedoch für das binomiale Verfahren. Beim multinomialen Verfahren wird die Generierung nicht unterstützt, wenn Konfidenzen ausgewählt wurden, es sei denn, der Zieltyp ist "Flag".
Generierte Regelsets	
Automatisches Klassifikationsmerkmal	Wenn ein UDF-Scoring-Adapter (UDF - User Defined Function, benutzerdefinierte Funktion) aktiviert ist, unterstützen diese Nuggets ein SQL-Pushback. Wenn entweder die SQL-Generierung für stetige Ziele oder die Ensemblemethode "Höchste Konfidenz hat Vorrang" verwendet wird, unterstützen diese Nuggets außerdem ein nachgeordnetes Pushback.
Autonumerisch	

Tabelle 10. Ausgabe

Knoten zur Unterstützung der SQL-Generierung	Hinweise
Tabelle	Unterstützt die Generierung, wenn die SQL-Generierung für den highlight-Ausdruck (siehe Ausdrücke unten) unterstützt wird.
Matrix	Unterstützt die Generierung, außer wenn "Alle numerischen Werte" für die Option "Felder" ausgewählt wurde.
Analyse	Unterstützt die Generierung je nach den ausgewählten Optionen.
Transformieren	
Grafiken	SQL-Pushback wird nicht unterstützt, wenn die Animation in einem Diagrammtafelknoten aktiviert ist.
Statistik	Unterstützt die Generierung, wenn die Option "Korrelieren" nicht verwendet wird.
Bericht	
Globalwerte	

Tabelle 11. Export

Knoten zur Unterstützung der SQL-Generierung	Hinweise
Datenbank	
Publisher	Der veröffentlichte Stream enthält das generierte SQL.

CLEM-Ausdrücke und Operatoren, die die SQL-Generierung unterstützen

In den folgenden Tabellen erhalten Sie eine Beschreibung der mathematischen Operationen und Ausdrücke, die eine SQL-Generierung unterstützen und die häufig während des Data-Minings verwendet werden. Operationen, die in diesen Tabellen nicht enthalten sind, unterstützen die SQL-Generierung in der aktuellen Version nicht.

Tabelle 12. Operatoren

Operation zur Unterstützung der SQL-Generierung	Hinweise
+	
-	
/	
*	
><	Wird zur Verkettung von Zeichenfolgen verwendet.

Tabelle 13. Relationale Operatoren

Operation zur Unterstützung der SQL-Generierung	Hinweise
=	
/=	Wird verwendet, um "ungleich" anzugeben.
>	
>=	
<	
<=	

Tabelle 14. Funktionen

Operation zur Unterstützung der SQL-Generierung	Hinweise
abs	
allbutfirst	
allbutlast	
and	
arccos	
arcsin	
arctan	
arctanh	
cos	
div	
exp	
fracof	
hasstartstring	
hassubstring	
integer	

Tabelle 14. Funktionen (Forts.)

Operation zur Unterstützung der SQL-Generierung	Hinweise
intof	
isalpha	
islower	
isnumber	
isstartstring	
issubstring	
isupper	
last	
length	
locchar	
log	
log10	
lowertoupper	
max	
member	
min	
negate	
not	
number	
or	
pi	
real	
rem	
round	
sign	
sin	
sqrt	
string	
strmember	
subscrs	
substring	
substring_between	
uppertolower	
to_string	

Tabelle 15. Sonderfunktionen

Operation zur Unterstützung der SQL-Generierung	Hinweise
@NULL	

Table 15. Sonderfunktionen (Forts.)

Operation zur Unterstützung der SQL-Generierung	Hinweise
@GLOBAL_AVE	Die speziellen globalen Funktionen werden zum Abrufen von Globalwerten verwendet, die vom Globalwerteknoten berechnet werden.
@GLOBAL_SUM	
@GLOBAL_MAX	
@GLOBAL_MEAN	
@GLOBAL_MIN	
@GLOBALSDEV	

Table 16. Aggregierungsfunktionen

Operation zur Unterstützung der SQL-Generierung	Hinweise
Summe	
Mittelwert	
Min	
Max	
Anzahl	
Standardabweichung	

Verwenden von SQL-Funktionen in CLEM-Ausdrücken

Mit der Funktion @SQLFN können benannte SQL-Funktionen innerhalb von CLEM-Ausdrücken zum ausschließlichen Zweck der Datenbankausführung hinzugefügt werden. Diese Vorgehensweise kann in bestimmten Fällen sinnvoll sein, wenn proprietäres SQL oder andere herstellerspezifische Anpassungen erforderlich sind.

Die Verwendung dieser Funktion wird nicht durch die standardmäßige IBM SPSS Modeler-Supportvereinbarung abgedeckt, da die Ausführung von externen Datenbankkomponenten abhängt, die sich der Kontrolle von IBM entziehen. Eine Bereitstellung ist jedoch in Sonderfällen möglich, normalerweise im Rahmen einer Dienstleistungsabsprache. Weitere Informationen erhalten Sie, falls erforderlich, unter <http://www.ibm.com/software/analytics/spss/services/>.

Schreiben von SQL-Abfragen

Wenn Sie den Datenbankknoten verwenden, sollten Sie besonders auf alle SQL-Abfragen achten, die zu einem Dataset mit doppelten Spaltennamen führen. Diese doppelten Namen verhindern häufig die SQL-Optimierung für alle nachgeordneten Knoten.

IBM SPSS Modeler verwendet verschachtelte SELECT-Anweisungen, um ein Pushback von SQL für Streams durchzuführen, die eine SQL-Abfrage im Datenbankquellenknoten verwenden. Mit anderen Worten: Der Stream verschachtelt die im Datenbankquellenknoten angegebene Abfrage in eine oder mehrere SELECT-Anweisungen, die während der Optimierung von nachgeordneten Knoten erstellt werden. Wenn also das Ergebnis einer Abfrage doppelte Spaltennamen enthält, kann die Anweisung nicht vom Verwaltungssystem für relationale Datenbanken verschachtelt werden. Verschachtelungsschwierigkeiten treten häufig während einer Tabellenzusammenfügung auf, bei der eine Spalte mit demselben Namen in mehr als einer der zusammengeführten Tabellen ausgewählt ist. Beachten Sie beispielsweise folgende Abfrage im Quellenknoten:

```
SELECT e.ID, e.LAST_NAME, d.*  
FROM EMP e RIGHT OUTER JOIN  
DEPT d ON e.ID = d.ID;
```

Die Abfrage verhindert eine nachfolgende SQL-Optimierung, da diese SELECT-Anweisung zu einem Datensatz mit zwei Spalten mit der Bezeichnung ID führen würde.

Um eine vollständige SQL-Optimierung zu ermöglichen, sollten Sie beim Schreiben von SQL-Abfragen genau sein und die Spaltenaliasnamen angeben, wenn eine Situation mit doppelten Spaltennamen auftritt. Die Anweisung unten zeigt eine genauere Abfrage:

```
SELECT e.ID AS ID1, e.LAST_NAME, d.*  
FROM EMP e RIGHT OUTER JOIN  
DEPT d ON e.ID = d.ID;
```

Scoring-Adapter für Teradata - doppelte Zeilen

Der Scoring-Adapter für Teradata von IBM SPSS Modeler Server erwartet keine identischen Zeilen in seinen Eingabedaten. Teradata lässt in einer Tabelle keine zwei identischen Zeilen zu. Allerdings können doppelte Zeilen beim Verknüpfen von Tabellen auftreten, oder wenn der Benutzer nur einen Teil der Felder einer Tabelle als Eingabe verwendet. Diese doppelten Zeilen führen dann zu einer falschen Anzahl Datensätze nach einem kartesischen Join.

Anhang A. Konfigurieren von Oracle für UNIX-Plattformen

Konfigurieren von Oracle für die SQL-Optimierung

Wenn Sie IBM SPSS Modeler Server auf UNIX-Plattformen ausführen und die Daten aus einer Oracle-Datenbank gelesen werden, sollten Sie die folgenden Tipps beachten, um sicherzustellen, dass das generierte SQL in der Datenbank vollständig optimiert wird.

Richtige Ländereinstellung

Wenn Sie IBM SPSS Modeler Server mit einer Ländereinstellung ausführen, die von der mit den Connect-ODBC-Treibern mitgelieferten abweicht, sollten Sie den Computer neu konfigurieren, um so die SQL-Optimierung zu verbessern. Die Connect-ODBC-Treiber werden lediglich mit den Dateien für die Ländereinstellung *en_US* geliefert. Wenn der IBM SPSS Modeler Server-Computer mit einer anderen Ländereinstellung ausgeführt wird oder wenn für die Shell, in der IBM SPSS Modeler Server gestartet wurde, die Ländereinstellung nicht vollständig definiert wurde, wird das generierte SQL u. U. in Oracle nicht vollständig optimiert. Dafür gibt es folgende Gründe:

- IBM SPSS Modeler Server verwendet die Dateien für die ODBC-Ländereinstellungen entsprechend der Ländereinstellung, mit der sie ausgeführt werden, um die von der Datenbank in Textzeichenfolgen ausgegebenen Codes zu übersetzen. Diese Textzeichenfolgen werden dann verwendet, um zu ermitteln, mit welcher Datenbank eine Verbindung hergestellt wird.
- Wenn die Ländereinstellung (wie von der Systemabfrage `$LANG` an IBM SPSS Modeler Server ausgegeben) nicht *en_US* lautet, kann IBM SPSS Modeler die Codes, die es vom ODBC-Treiber erhält, nicht in Text übersetzen. Mit anderen Worten: Zu Beginn der Datenbankverbindung wird anstelle der Zeichenfolge *Oracle* ein unübersetzter Code an IBM SPSS Modeler Server ausgegeben. Dies bedeutet, dass IBM SPSS Modeler keine Streams für Oracle optimieren kann.

So prüfen Sie die Ländereinstellungen und setzen Sie zurück:

1. Führen Sie in einer UNIX-Shell folgenden Befehl aus:

```
#locale
```

Sie erhalten Informationen zur Ländereinstellung für die Shell. Beispiel:

```
$ locale
LANG=en_US.ISO8859-15
LC_CTYPE="en_US.ISO8859-15"
LC_NUMERIC="en_US.ISO8859-15"
LC_TIME="en_US.ISO8859-15"
LC_COLLATE="en_US.ISO8859-15"
LC_MONETARY="en_US.ISO8859-15"
LC_MESSAGES="en_US.ISO8859-15"
LC_ALL=en_US.ISO8859-15
```

2. Wechseln Sie in das Connect-Verzeichnis ODBC/locale. (Hier sehen Sie das Unterverzeichnis *en_US*.)
3. Erstellen Sie eine Verknüpfung zu diesem Verzeichnis *en_US*, indem Sie den Namen der Ländereinstellung der Shell angeben. Ein Beispiel:

```
#ln -s en_US en_US.ISO8859-15
```

Für eine anderssprachige Ländereinstellung, wie z. B. *fr_FR.ISO8859-1*, sollten Sie die Verknüpfung wie folgt herstellen:

```
#ln -s en_US fr_FR.ISO8859-1
```

4. Nachdem Sie die Verknüpfung erstellt haben, müssen Sie IBM SPSS Modeler Server aus derselben Shell neu starten. (IBM SPSS Modeler Server erhält die jeweiligen Ländereinstellungsinformationen aus der Shell, aus der das Produkt gestartet wurde.)

Hinweise

Wenn Sie ein UNIX-System für SQL-Pushbacks nach Oracle optimieren, sollten Sie folgende Tipps beachten:

- Die vollständige Ländereinstellung muss angegeben werden. Im obigen Beispiel müssen Sie die Verknüpfung in der Form `language_territory.code-page` erstellen. Das bereits vorhandene Verzeichnis für die Ländereinstellung `en_US` reicht nicht aus.
- Um das datenbankinterne Mining zu optimieren, müssen sowohl LANG als auch LC_ALL in der zum Starten von IBM SPSS Modeler Server verwendeten Shell definiert sein. LANG kann in der Shell genauso definiert werden wie jede andere Umgebungsvariable, bevor IBM SPSS Modeler Server neu gestartet wird. Siehe folgende Definition:

```
#LANG=en_US.IS08859-15; export LANG
```

- Bei jedem Start von IBM SPSS Modeler Server müssen Sie prüfen, ob die Shell-Informationen zur Ländereinstellung vollständig angegeben sind und ob die entsprechende Verknüpfung im Verzeichnis ODBC/locale vorhanden ist.

Anhang B. Konfigurieren von UNIX-Startscripts

Einführung

In diesem Anhang werden einige der Scripts beschrieben, die im Lieferumfang der UNIX-Versionen von IBM SPSS Modeler Server enthalten sind. Außerdem erfahren Sie hier, wie Sie die Scripts konfigurieren können. Scripts werden für die folgenden Zwecke verwendet:

- Konfigurieren von IBM SPSS Modeler Server für einen automatischen Start beim Neustart des Server-Computers.
- Manuelles Stoppen und Neustarten von IBM SPSS Modeler Server.
- Ändern der Berechtigungen für die von IBM SPSS Modeler Server erstellten Dateien.
- Konfigurieren von IBM SPSS Modeler Server, um mit den in IBM SPSS Modeler Server enthaltenen ODBC Connect-Treibern zu arbeiten. Weitere Informationen finden Sie in „IBM SPSS Modeler Server und das Data Access Pack“ auf Seite 80.

Scripts

IBM SPSS Modeler Server verwendet verschiedene Scripts, darunter:

- **modelersrv.sh.** Das Script für den manuellen Start von IBM SPSS Modeler Server befindet sich im Installationsverzeichnis von IBM SPSS Modeler Server. Es konfiguriert die Serverumgebung, wenn der Serverdämonprozess *manuell* gestartet wird. Führen Sie es aus, wenn Sie den Server manuell starten und stoppen möchten. Bearbeiten Sie es, wenn Sie die Konfiguration für den manuellen Start ändern möchten.
- **auto.sh.** Mit diesem Script wird Ihr System konfiguriert, um den Serverdämonprozess beim Neustart automatisch zu starten. Führen Sie das Script ein Mal aus, um Ihr System für den automatischen Start zu konfigurieren. Sie müssen es nicht bearbeiten. Das Script befindet sich im Installationsverzeichnis von IBM SPSS Modeler Server.
- **rc.modeler.** Beim Ausführen von *auto.sh* wird dieses Script an einem Speicherort erstellt, der vom Betriebssystem Ihres Servers abhängt. Es konfiguriert die Serverumgebung, wenn der Server *automatisch* gestartet wird. Bearbeiten Sie es, wenn Sie die Konfiguration für den automatischen Start ändern möchten.

Tabelle 17. Speicherort von "rc.modeler" unter den verschiedenen Betriebssystemen.

Betriebssystem	Speicherort
AIX	/etc/rc.modeler

Automatisches Starten und Stoppen von IBM SPSS Modeler Server

IBM SPSS Modeler Server muss als Dämonprozess gestartet werden. Im Installationsprogramm ist das Script *auto.sh* enthalten, mit dem Sie Ihr System für das automatische Stoppen und Neustarten von IBM SPSS Modeler Server konfigurieren können.

So konfigurieren Sie das System für das automatische Starten und Herunterfahren:

1. Melden Sie sich als Root an.
2. Wechseln Sie zum IBM SPSS Modeler Server-Installationsverzeichnis.
3. Führen Sie das Script aus. Geben Sie an der UNIX-Eingabeaufforderung Folgendes ein:
./auto.sh

Das automatische Startscript *rc.modeler* wird an dem in der vorangehenden Tabelle angegebenen Speicherort erstellt. Das Betriebssystem verwendet *rc.modeler*, um den IBM SPSS Modeler Server-Dämonprozess zu starten, wenn der Server-Computer neu gestartet wird. Außerdem verwendet das Betriebssystem *rc.modeler* auch, um den Dämon zu stoppen, wenn der Computer heruntergefahren wird.

Manuelles Starten und Stoppen von IBM SPSS Modeler Server

Sie können IBM SPSS Modeler Server manuell starten und stoppen, indem Sie das Script *modelersrv.sh* ausführen.

So können Sie IBM SPSS Modeler Server manuell starten und stoppen:

1. Wechseln Sie zum IBM SPSS Modeler Server-Installationsverzeichnis.
2. Um den Server zu starten, geben Sie an der UNIX-Eingabeaufforderung Folgendes ein:
`./modelersrv.sh start`
3. Um den Server zu stoppen, geben Sie an der UNIX-Eingabeaufforderung Folgendes ein:
`./modelersrv.sh stop`

Bearbeiten von Scripts

Wenn Sie sowohl den manuellen als auch den automatischen Start verwenden, nehmen Sie die Änderungen parallel in *modelersrv.sh* und *rc.modeler* vor. Wenn Sie nur den manuellen Start verwenden, bearbeiten Sie nur *modelersrv.sh*. Wenn Sie nur den automatischen Start verwenden, bearbeiten Sie nur *rc.modeler*.

So bearbeiten Sie Scripts:

1. Stoppen Sie IBM SPSS Modeler Server. (Weitere Informationen finden Sie in „Manuelles Starten und Stoppen von IBM SPSS Modeler Server“.)
2. Wählen Sie das entsprechende Script aus. (Weitere Informationen finden Sie in „Scripts“ auf Seite 79.)
3. Öffnen Sie das Script in einem Texteditor, nehmen Sie die Änderungen vor und speichern Sie die Datei.
4. Starten Sie IBM SPSS Modeler Server, entweder automatisch (indem Sie den Server-Computer neu starten) oder manuell.

Steuern der Berechtigungen für die Dateierstellung

IBM SPSS Modeler Server erstellt temporäre Dateien mit Lese-, Schreib- und Ausführungsberechtigungen für alle Benutzer. Sie können diese Standardeinstellung überschreiben, indem Sie im Startscript *modelersrv.sh*, *rc.modeler* oder in beiden die Einstellung *UMASK* bearbeiten. (Weitere Informationen finden Sie in „Bearbeiten von Scripts“.) Wir empfehlen als restriktivste *UMASK*-Einstellung 077. Restriktivere Einstellungen könnten zu Berechtigungsproblemen für IBM SPSS Modeler Server führen.

IBM SPSS Modeler Server und das Data Access Pack

Wenn Sie die ODBC-Treiber mit IBM SPSS Modeler Server verwenden möchten, muss die ODBC-Umgebung mit *odbc.sh* konfiguriert werden, wenn der IBM SPSS Modeler Server-Prozess startet. Sie erreichen dies, indem Sie das entsprechende IBM SPSS Modeler-Startscript in *modelersrv.sh* und/oder in *rc.modeler* bearbeiten. (Weitere Informationen finden Sie in „Bearbeiten von Scripts“.)

Weitere Informationen finden Sie auf der Website des Technical Support unter <http://www.ibm.com/support>. Wenn Sie Fragen zur Erstellung oder Einstellung von Berechtigungen für ODBC-Datenquellen haben, wenden Sie sich an Ihren Datenbankadministrator.

So konfigurieren Sie ODBC für den Start mit IBM SPSS Modeler Server:

1. Stoppen Sie den IBM SPSS Modeler Server-Host, falls er gerade ausgeführt wird.

2. Laden Sie das relevante komprimierte TAR-Archiv für die Plattform herunter, auf der IBM SPSS Modeler Server installiert ist. Achten Sie darauf, die richtigen Treiber für installierte Version von IBM SPSS Modeler Server herunterzuladen. Kopieren Sie die Datei an den Speicherort, an dem Sie die ODBC-Treiber installieren möchten (z. B. /usr/spss/odbc).
3. Extrahieren Sie die TAR-Archivdatei mit **tar -xvof**.
4. Führen Sie das setodbcpath.sh-Script aus, das aus dem Archiv extrahiert wird.
5. Bearbeiten Sie das Script odbc.sh, um die Definition von ODBCINI unten zu diesem Script hinzuzufügen und sie zu exportieren, z. B.:

```
ODBCINI=/usr/spss/odbc/odbc.ini;
export ODBCINI
```

ODBCINI muss auf den vollständigen Pfadnamen der odbc.ini-Datei verweisen, die von IBM SPSS Modeler gelesen werden soll, um eine Liste der ODBC-Datenquellen, die Sie definieren, zu erhalten (eine Standard-odbc.ini wurde mit den Treibern installiert).

6. Speichern Sie odbc.sh.
7. (Nur 64-Bit-IBM SPSS Modeler Server-Installationen, fahren Sie für alle anderen Installationen beim nächsten Schritt fort) Definieren und exportieren Sie LD_LIBRARY_PATH_64 in odbc.sh:

```
if [ "$LD_LIBRARY_PATH_64" = "" ]; then
    LD_LIBRARY_PATH_64=<Bibliothekspfad>
else
    LD_LIBRARY_PATH_64=<Bibliothekspfad>:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
```

Dabei ist library_path identisch mit dem Wert der LD_LIBRARY_PATH-Definition, die bereits in dem Script enthalten ist, das mit Ihrem Installationspfad initialisiert wurde (z. B. /usr/spss/odbc/lib). Der einfachste Weg ist, die Anweisungen if und export für LD_LIBRARY_PATH in Ihre odbc.sh-Datei zu kopieren, sie an das Ende der Datei anzuhängen und dann die LD_LIBRARY_PATH-Zeichenfolgen in den neu angehängten if- und export-Anweisungen durch LD_LIBRARY_PATH_64 zu ersetzen.

Ihre resultierende odbc.sh-Datei in einer 64-Bit-IBM SPSS Modeler Server-Installation könnte zum Beispiel wie folgt aussehen:

```
if [ "$LD_LIBRARY_PATH" = "" ]; then
    LD_LIBRARY_PATH=/usr/spss/odbc/lib
else
    LD_LIBRARY_PATH=/usr/spss/odbc/lib:$LD_LIBRARY_PATH
fi
export LD_LIBRARY_PATH
if [ "$LD_LIBRARY_PATH_64" = "" ]; then
    LD_LIBRARY_PATH_64=/usr/spss/odbc/lib
else
    LD_LIBRARY_PATH_64=/usr/spss/odbc/lib:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
ODBCINI=/usr/spss/odbc/odbc.ini;
export ODBCINI
```

Denken Sie daran, LD_LIBRARY_PATH_64 zu exportieren und mit der if-Schleife zu definieren.

8. Bearbeiten Sie die odbc.ini-Datei, die Sie vorher mit \$ODBCINI definiert haben. Definieren Sie die Datenquellennamen, die Sie benötigen (diese hängen von der Datenbank ab, auf die sie zugreifen).
9. Speichern Sie die odbc.ini-Datei.
10. Konfigurieren Sie IBM SPSS Modeler Server, um diese Treiber zu verwenden. Bearbeiten Sie dazu modelersrv.sh und fügen Sie die folgende Zeile direkt unter die Zeile ein, die SCLEMDNAME definiert:


```
. <odbc.sh_path>
```

Dabei ist `odbc.sh_path` der vollständige Pfad zur Datei `odbc.sh`, die Sie zu Beginn dieses Vorgangs bearbeitet haben, z. B.:

```
. /usr/spss/odbc/odbc.sh
```

Anmerkung: Hier ist die Syntax wichtig. Achten Sie darauf, dass zwischen dem ersten Punkt und dem Pfad zur Datei ein Leerzeichen ist.

11. Speichern Sie `modelersrv.sh`.

Wichtig: Damit der Treiber für SDAP 7.1 in DB2 unter z/OS funktioniert, müssen Sie Zugriff auf `SYS-IBM.SYSPACKSTMT` erteilen.

So testen Sie die Verbindung:

1. Starten Sie IBM SPSS Modeler Server neu.
2. Stellen Sie über einen Client eine Verbindung zu IBM SPSS Modeler Server her.
3. Fügen Sie auf dem Client einen Datenbankquellenknoten zum Erstellungsbereich hinzu.
4. Öffnen Sie den Knoten und stellen Sie sicher, dass Sie die Datenquellennamen sehen können, die Sie zu einem früheren Zeitpunkt des Konfigurationsvorgangs in der `odbc.ini`-Datei definiert haben.

Wenn die Namen nicht angezeigt werden oder wenn bei dem Versuch, eine Verbindung zu einer von Ihnen definierten Datenquelle herzustellen, Fehlermeldungen angezeigt werden, folgen Sie dem Verfahren zur Fehlerbehebung. Weitere Informationen finden Sie in „Fehlerbehebung bei der ODBC-Konfiguration“ auf Seite 83.

So konfigurieren Sie ODBC für den Start mit IBM SPSS Modeler Solution Publisher Runtime:

Wenn Sie erfolgreich von IBM SPSS Modeler Server eine Verbindung zur Datenbank herstellen können, können Sie eine IBM SPSS Modeler Solution Publisher Runtime-Installation auf dem gleichen Server konfigurieren, indem Sie das gleiche `odbc.sh`-Script aus dem Startscript von IBM SPSS Modeler Solution Publisher Runtime referenzieren.

1. Bearbeiten Sie das `modelerrun`-Script in IBM SPSS Modeler Solution Publisher Runtime, um die folgende Zeile direkt unterhalb der letzten Zeile des Scripts einzufügen:

```
. <odbc.sh_path>
```

Dabei ist `odbc.sh_path` der vollständige Pfad zur Datei `odbc.sh`, die Sie zu Beginn dieses Vorgangs bearbeitet haben, z. B.:

```
. /usr/spss/odbc/odbc.sh
```

Anmerkung: Hier ist die Syntax wichtig. Achten Sie darauf, dass zwischen dem ersten Punkt und dem Pfad zur Datei ein Leerzeichen steht.

2. Speichern Sie die `modelerrun`-Scriptdatei.
3. Standardmäßig ist der DataDirect-Treibermanager nicht für die Verwendung von ODBC durch IBM SPSS Modeler Solution Publisher Runtime auf UNIX-Systemen konfiguriert. Geben Sie folgende Befehle ein, um UNIX so zu konfigurieren, dass der DataDirect-Treibermanager geladen wird (wobei `Solution_Publisher-Installationsverzeichnis` für das Installationsverzeichnis von Solution Publisher Runtime steht):

```
cd Solution_Publisher-Installationsverzeichnis
rm -f libspssodbc.so
ln -s libspssodbc_datadirect.so libspssodbc.so
```

So konfigurieren Sie ODBC für den Start mit IBM SPSS Modeler Batch:

Für ODBC ist keine Konfiguration des IBM SPSS Modeler Batch-Scripts erforderlich. Das liegt daran, dass Sie die Verbindung zu IBM SPSS Modeler Server über IBM SPSS Modeler Batch herstellen, um Streams

auszuführen. Stellen Sie sicher, dass die IBM SPSS Modeler Server-ODBC-Konfiguration durchgeführt wurde und ordnungsgemäß funktioniert, wie weiter oben in diesem Abschnitt beschrieben.

So fügen Sie einen Datenquellennamen hinzu oder bearbeiten ihn:

1. Bearbeiten Sie die `odbc.ini`-Datei, um den neuen oder geänderten Namen aufzunehmen.
2. Testen Sie die Verbindung wie weiter oben in diesem Abschnitt beschrieben.

Wenn die Verbindung mit IBM SPSS Modeler Server ordnungsmäßig funktioniert, sollte die neue oder geänderte Datenquelle auch korrekt mit IBM SPSS Modeler Solution Publisher Runtime und IBM SPSS Modeler Batch funktionieren.

SQL Server-Unterstützung mit dem Data Access Pack-Treiber

Für die ODBC-Konfiguration für SQL Server muss das ODBC-Verbindungsattribut `Enable Quoted Identifiers` auf `Yes` gesetzt werden (der Standardwert für diesen Treiber ist `No`). Unter UNIX wird dieses Attribut in der Systeminformationsdatei (`odbc.ini`) über die Option `QuotedId` konfiguriert.

Fehlerbehebung bei der ODBC-Konfiguration

Keine Datenquellen aufgelistet oder zufälliger Text angezeigt

Wenn Sie einen Datenbankquellenknoten öffnen und die Liste der verfügbaren Datenquellen leer ist oder unerwartete Einträge enthält, kann dies an einem Problem mit dem Startscript liegen.

1. Überprüfen Sie, ob `$ODBCINI` in `modelersrv.sh` definiert ist, entweder ausdrücklich im Script selbst oder im `odbc.sh`-Script, das in `modelersrv.sh` referenziert ist.
2. Stellen Sie in letzterem Fall sicher, dass `ODBCINI` auf den vollständigen Pfad zur `odbc.ini`-Datei verweist, die Sie verwendet haben, um Ihre ODBC-Datenquellen zu definieren.
3. Wenn die Pfadangabe in `ODBCINI` richtig ist, überprüfen Sie den `$ODBCINI`-Wert, der in der IBM SPSS Modeler Server-Umgebung verwendet wird, indem Sie die Variable aus `modelersrv.sh` wiederholen. Fügen Sie dazu die folgende Zeile zu `modelersrv.sh` hinzu, und zwar nach der Position, an der Sie `ODBCINI` definieren:

```
echo $ODBCINI
```
4. Speichern Sie `modelersrv.sh` und führen Sie das Script dann aus. Der in der IBM SPSS Modeler Server-Umgebung festgelegte `$ODBCINI`-Wert wird zur Verifizierung in die Standardausgabe geschrieben.
5. Wenn kein Wert an die Standardausgabe zurückgegeben wird und Sie `$ODBCINI` im `odbc.sh`-Script, das Sie von `modelersrv.sh` referenzieren, definieren, überprüfen Sie, ob die Referenzierungssyntax korrekt ist. Sie sollte wie folgt aussehen:

```
. <odbc.sh_path>
```

Dabei ist `odbc.sh_path` der vollständige Pfad zur Datei `odbc.sh`, die Sie zu Beginn dieses Vorgangs bearbeitet haben, z. B.:

```
. /usr/spss/odbc/odbc.sh
```

Hinweis: Hier ist die Syntax wichtig. Achten Sie darauf, dass zwischen dem ersten Punkt und dem Pfad zu Datei ein Leerzeichen ist.

Wenn der korrekte Wert bei ausgeführter `modelersrv.sh` in der Standardausgabe wiederholt wird, sollten Sie die Datenquellennamen im Datenbankquellenknoten sehen können, wenn Sie IBM SPSS Modeler Server neu starten und vom Client eine Verbindung herstellen.

Der IBM SPSS Modeler-Client hängt beim Klicken auf das Dialogfeld "In Datenbankverbindungen verbinden"

Dieses Verhalten kann von Ihrem Bibliothekspfad verursacht werden, der nicht korrekt festgelegt wurde, sodass er den Pfad zu den ODBC-Bibliotheken enthält. Der Bibliothekspfad wird definiert von \$LD_LIBRARY_PATH (und \$LD_LIBRARY_PATH_64 bei 64-Bit-Versionen).

Um den Wert des Bibliothekspfads in der IBM SPSS Modeler Server-Dämonumgebung anzuzeigen, wiederholen Sie den Wert der entsprechenden Umgebungsvariable aus *modelersrv.sh* nach der Zeile, in der Sie den ODBC-Bibliothekspfad an den Bibliothekspfad anhängen, und führen Sie das Script aus. Der Wert des Bibliothekspfads wird im Terminal wiederholt, wenn Sie das Script das nächste Mal ausführen.

Wenn Sie *odbc.sh* aus *modelersrv.sh* referenzieren, um Ihre IBM SPSS Modeler Server-ODBC-Umgebung einzurichten, wiederholen Sie den Bibliothekspfadwert aus der Zeile nach der, in der Sie das *odbc.sh*-Script referenzieren. Um den Wert zu wiederholen, fügen Sie die folgende Zeile zum Script hinzu, speichern Sie dann und führen Sie die Scriptdatei aus:

```
echo <library_path_variable>
```

Dabei ist *<library_path_variable>* die geeignete Bibliothekspfadvariable für Ihr Serverbetriebssystem.

Der zurückgegebene Wert Ihres Bibliothekspfads muss den Pfad zum *lib*-Unterverzeichnis Ihrer ODBC-Installation enthalten. Wenn dies nicht der Fall ist, hängen Sie diesen Speicherort an die Datei an.

Wenn Sie die 64-Bit-Version von IBM SPSS Modeler Server ausführen, überschreibt \$LD_LIBRARY_PATH_64 \$LD_LIBRARY_PATH, wenn diese Variable festgelegt ist. Wenn dieses Problem bei einer der 64-Bit-Plattformen auftritt, wiederholen Sie LD_LIBRARY_PATH_64 und \$LD_LIBRARY_PATH aus *modelersrv.sh* und, falls erforderlich, legen Sie \$LD_LIBRARY_PATH_64 so fest, dass diese Variable das *lib*-Unterverzeichnis Ihrer ODBC-Installation enthält, und exportieren Sie die Definition.

Datenquellename nicht gefunden und kein Standardtreiber angegeben

Wenn dieser Fehler beim Klicken auf das Dialogfeld "In Datenbankverbindungen verbinden" angezeigt wird, bedeutet dies meist, dass Ihre *odbc.ini* nicht richtig definiert wurde. Stellen Sie sicher, dass der Datenquellename (DSN), wie er im Abschnitt [ODBC Data Sources] oben in der Datei definiert ist, mit der Zeichenfolge zwischen den eckigen Klammern weiter unten in *odbc.ini* übereinstimmt, um den DSN zu definieren. Wenn diese sich in irgendeiner Art und Weise unterscheiden, wird diese Fehlernachricht angezeigt, wenn Sie eine Verbindung unter Verwendung des DSN von IBM SPSS Modeler herstellen möchten. Es folgt ein Beispiel einer *falschen* Spezifikation:

```
[ODBC Data Sources]
Oracle=Oracle Wire Protocol

....
....
[Oracle Driver]
Driver=/usr/ODBC/lib/XEora22.so
Description=SPSS 5.2 Oracle Wire Protocol
AlternateServers=
....
```

Sie müssen eine der beiden fett markierten Zeichenfolgen ändern, damit sie genau übereinstimmen. Dies sollte den Fehler beheben.

Angegebener Treiber konnte nicht geladen werden

Dieser Fehler gibt auch an, dass die *odbc.ini*-Datei falsch definiert ist. Eine Möglichkeit ist, dass der Treiberparameter in der Treiberstanza falsch festgelegt ist, z. B.:

```
[ODBC Data Sources]
Oracle=Oracle Wire Protocol
```

```

....
....
[Oracle]
Driver=/nosuchpath/ODBC/lib/XEora22.so
Description=SPSS 5.2 Oracle Wire Protocol
AlternateServers=

```

1. Überprüfen Sie, ob das vom Treiberparameter angegebene gemeinsam genutzte Objekt vorhanden ist.
2. Korrigieren Sie den Pfad zum gemeinsam genutzten Objekt, falls dieser falsch ist.
3. Wenn der Treiberparameter im Format

```
Driver=ODBCHOME/lib/XEora22.so
```

spezifiziert ist, bedeutet dies, dass Sie Ihre ODBC-bezogenen Scripts nicht initialisiert haben. Führen Sie das *setodbcpath.sh*-Script aus, das mit den Treibern installiert wurde. Weitere Informationen finden Sie in „IBM SPSS Modeler Server und das Data Access Pack“ auf Seite 80. Wenn Sie dieses Script ausgeführt haben, sollten Sie sehen, dass die Zeichenfolge "ODBCHOME" durch den Pfad zu Ihrer ODBC-Installation ersetzt wurde. Dies sollte das Problem lösen.

Ein weiterer Grund könnte ein Problem mit der Bibliothek des Treibers sein. Verwenden Sie das in ODBC enthaltene Tool *ivtestlib*, um zu bestätigen, dass der Treiber nicht geladen werden kann. Bei Connect64 verwenden Sie das Tool *ddtestlib*. Beheben Sie das Problem, indem Sie die Variable für den Bibliothekspfad im Startscript festlegen.

Wenn beispielsweise der Oracle-Treiber für eine 32-Bit-Installation nicht geladen werden kann, führen Sie folgende Schritte durch:

1. Verwenden Sie *ivtestlib*, um zu bestätigen, dass der Treiber nicht geladen werden kann. Geben Sie beispielsweise an der UNIX-Eingabeaufforderung Folgendes ein:

```

sh
cd ODBCDir
. odbcs.sh
./bin/ivtestlib MFor815

```

Dabei wird *ODBCDir* durch den Pfad zum ODBC-Installationsverzeichnis ersetzt.

2. Lesen Sie die Nachricht, um festzustellen, ob ein Fehler vorliegt. Die Nachricht *Load of MFor815.so failed: ld.so.1: bin/ivtestlib: fatal: libcIntsh.so: open failed: No such file or directory* gibt z. B. an, dass die Oracle-Client-Bibliothek *libcIntsh.so* nicht vorhanden ist oder sich nicht im Bibliothekspfad befindet.
3. Stellen Sie fest, ob die Bibliothek vorhanden ist. Falls nicht, installieren Sie den Oracle-Client erneut. Ist die Bibliothek vorhanden, geben Sie an der UNIX-Eingabeaufforderung die folgende Befehlssequenz ein:

```

LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/bigdisk/oracle/product/8.1.6/lib
export LD_LIBRARY_PATH
./bin/ivtestlib Mfor815

```

Dabei wird */bigdisk/oracle/product/8.1.6/lib* durch den Pfad zu *libcIntsh.so* ersetzt. *LD_LIBRARY_PATH* ist die Variable für den Bibliothekspfad für Ihr Betriebssystem.

Beachten Sie: Wenn Sie die 64-Bit-Version von IBM SPSS Modeler unter Linux ausführen, enthält die Variable für den Bibliothekspfad das Suffix *_64*. Die ersten beiden Zeilen aus dem vorstehenden Beispiel lauten hier also entsprechend:

```

LD_LIBRARY_PATH_64=$LD_LIBRARY_PATH_64:/bigdisk/oracle/product/8.1.6/lib
export LD_LIBRARY_PATH_64

```

4. Lesen Sie die Nachricht, um festzustellen, ob der Treiber nun geladen werden kann. Die Nachricht *Load of MFor815.so successful, qehandle is 0xFF3A1BE4* gibt an, dass die Oracle-Client-Bibliothek geladen werden kann.
5. Berichten Sie den Bibliothekspfad im Startscript von IBM SPSS Modeler.
6. Starten Sie IBM SPSS Modeler Server mit dem Startscript neu, das Sie bearbeitet haben (*modelersrv.sh* bzw. *rc.modeler*).

Bibliothekspfade

Der Name der Variablen für den Bibliothekspfad hängt von Ihrem Betriebssystem ab. Die folgende Tabelle dient als Hilfestellung bei den entsprechenden Einstellungen, wenn Sie Ihr System konfigurieren oder Fehler beheben.

Tabelle 18. Bibliothekspfade (nach Betriebssystem geordnet).

Betriebssystem	Name der Bibliothekspfadvariablen
AIX	LIBPATH
Linux (64-Bit)	LD_LIBRARY_PATH_64

Anhang C. Konfigurieren und Ausführen von SPSS Modeler Server als Nicht-Root-Prozess unter UNIX

Einführung

In den folgenden Anweisungen finden Sie Informationen zur Ausführung von IBM SPSS Modeler Server als Nicht-Root-Prozess unter UNIX.

Ausführen als Root. Die Standardinstallation von IBM SPSS Modeler Server geht davon aus, dass der Serverdämonprozess als Root ausgeführt wird. Bei der Ausführung als Root kann IBM SPSS Modeler jede Benutzeranmeldung zuverlässig authentifizieren und jede Benutzersitzung mit dem entsprechenden UNIX-Benutzerkonto starten. Auf diese Weise wird sichergestellt, dass Benutzer ausschließlich Zugriff auf ihre eigenen Dateien und Verzeichnisse erhalten.

Ausführen als Nicht-Root. Um IBM SPSS Modeler Server als Nicht-Root-Prozess auszuführen, werden die echten und effektiven Benutzer-IDs des Serverdämonprozesses auf ein Konto Ihrer Wahl gesetzt. Alle von SPSS Modeler Server gestarteten Benutzersitzungen verwenden dasselbe UNIX-Konto. Dies bedeutet, dass alle Dateidaten, die von SPSS Modeler gelesen oder geschrieben werden, für alle SPSS Modeler-Benutzer freigegeben werden. Dies gilt nicht für den Zugriff auf Datenbankdaten, da sich Benutzer unabhängig hiervon bei allen den von ihnen verwendeten Datenbank-Datenquellen authentifizieren müssen. Ohne Rootberechtigung funktioniert IBM SPSS Modeler auf eine der beiden folgenden Weisen:

- **Ohne private Kennwortdatenbank.** Bei dieser Methode verwendet SPSS Modeler die bestehende UNIX-Kennwortdatenbank, NIS oder den LDAP-Server, der normalerweise für die Authentifizierung auf dem UNIX-System verwendet wird. Weitere Informationen finden Sie in „Konfigurieren als Nicht-Root-Benutzer ohne private Kennwortdatenbank“.
- **Mit privater Kennwortdatenbank.** Bei dieser Methode authentifiziert SPSS Modeler die Benutzer anhand einer privaten Kennwortdatenbank, die von der bestehenden UNIX-Kennwortdatenbank, NIS bzw. dem LDAP-Server, der normalerweise für die Authentifizierung auf dem UNIX-System verwendet wird, getrennt ist. Weitere Informationen finden Sie in „Konfigurieren als Nicht-Root-Benutzer mit privater Kennwortdatenbank“ auf Seite 88.

Konfigurieren als Nicht-Root-Benutzer ohne private Kennwortdatenbank

So konfigurieren Sie IBM SPSS Modeler Server für die Ausführung mit einem Nicht-Root-Konto, ohne dass eine private Kennwortdatenbank erforderlich ist:

1. Öffnen Sie die SPSS Modeler Server-Datei *options.cfg* zum Bearbeiten.
2. Setzen Sie die Option `start_process_as_login_user` auf Y.
3. Speichern und schließen Sie die Datei *options.cfg*.

Standardmäßig probiert SPSS Modeler Server jede Authentifizierungsmethode aus, bis eine funktionierende Methode gefunden wird. Sie können über die Option `authentication_methods` in *options.cfg* den Server aber auch so konfigurieren, dass er nur eine bestimmte Authentifizierungsmethode ausprobiert. Mögliche Werte für die Option sind `pasw_modeler`, `gss`, `pam`, `sspi`, `unix` oder `windows`.

Bei der Ausführung als Nicht-Root-Benutzer sind wahrscheinlich Konfigurationsaktualisierungen erforderlich. Weitere Informationen finden Sie in „Fehlerbehebung bei fehlgeschlagener Benutzerauthentifizierung“ auf Seite 90.

Vorsicht:

Sie sollten nicht die Einstellung `start_process_as_login_user` aktivieren und dann IBM SPSS Modeler Server als `Root` starten. Dadurch würden für alle Benutzer, die mit dem Server verbunden sind, die Serverprozesse als `Root` ausgeführt, was ein Sicherheitsrisiko darstellen würde. Beachten Sie, dass der Server möglicherweise automatisch gestoppt wird, wenn Sie versuchen, dies zu tun.

Konfigurieren als Nicht-Root-Benutzer mit privater Kennwortdatenbank

Wenn Sie auswählen, dass die Benutzer mithilfe einer privaten Kennwortdatenbank authentifiziert werden sollen, werden alle Sitzungen mit demselben Nicht-Root-Benutzerkonto gestartet.

So konfigurieren Sie IBM SPSS Modeler Server für eine solche Ausführung als Nicht-Root-Konto:

1. Erstellen Sie eine Gruppe, die alle Benutzer enthält. Sie können dieser Gruppe einen beliebigen Namen geben. Für dieses Beispiel wird `modelerusers` verwendet.
2. Erstellen Sie das Benutzerkonto, auf dem IBM SPSS Modeler Server ausgeführt werden soll. Dieses Konto dient zur ausschließlichen Verwendung für den IBM SPSS Modeler Server-Dämonprozess. Für dieses Beispiel wird der Name `modelerserv` verwendet.

Beachten Sie beim Erstellen des Kontos Folgendes:

- Die Hauptgruppe sollte die soeben erstellte Gruppe `<modelerusers>` sein.
 - Das Basisverzeichnis kann das Installationsverzeichnis von IBM SPSS Modeler sein oder jedes andere geeignete Standardverzeichnis (wenn das Konto auch nach Aufrüstungen erhalten bleiben soll, sollten Sie nicht das Installationsverzeichnis verwenden).
3. Konfigurieren Sie anschließend die Startscripts, um IBM SPSS Modeler Server mit dem neu erstellten Konto zu starten. Öffnen Sie das entsprechende Startscript in einem Texteditor. Weitere Informationen finden Sie in „Scripts“ auf Seite 79.

- a. Ändern Sie die Einstellung `umask`, um mindestens Gruppenlesezugriff für erstellte Dateien zu erteilen:

```
umask 027
```

4. Bearbeiten Sie die Datei mit den Serveroptionen `config/options.cfg`, um die Authentifizierung mit der privaten Kennwortdatenbank einzustellen, indem Sie die folgende Zeile anhängen:

```
authentication_methods, "pasw_modeler"
```

5. Setzen Sie die Option `start_process_as_login_user` auf `Y`.

6. Anschließend erstellen Sie eine private Kennwortdatenbank, die in der Datei `config/passwords.cfg` gespeichert wird. Die Kennwortdatei definiert die Kombinationen aus Benutzername/Kennwort, die für die Anmeldung bei IBM SPSS Modeler zulässig sind. *Hinweis:* Diese Daten sind nur für IBM SPSS Modeler bestimmt (privat) und stehen nicht in Zusammenhang mit den Benutzernamen und Kennwörtern, die für die Anmeldung an UNIX verwendet werden. Sie können dieselben Benutzernamen verwenden, wenn dies praktikabel ist, die Verwendung derselben Kennwörter ist jedoch nicht zulässig. Zum Erstellen der Kennwortdatei benötigen Sie das Dienstprogramm für Kennwörter (`pwutil`), das sich im Verzeichnis `bin` der IBM SPSS Modeler Server-Installation befindet. Die Syntax dieses Programms lautet:

```
pwutil [ Benutzername [ Kennwort ] ]
```

Das Programm nimmt einen Benutzernamen und ein Nur-Text-Kennwort und schreibt den Benutzernamen und das verschlüsselte Kennwort in einem für die Kennwortdatei geeigneten Format in die Standardausgabe. Wenn Sie beispielsweise den Benutzernamen `modeler` mit dem Kennwort "data mining" definieren wollen, geben Sie Folgendes ein:

```
bin/pwutil modeler "data mining" > config/passwords.cfg
```

Das Definieren eines einzigen Benutzernamens ist in den meisten Fällen, in denen sich alle Benutzer mit demselben Namen und Kennwort anmelden, ausreichend. Mit dem Operator `>>`, der an die einzelnen Dateien angehängt wird, können Sie jedoch noch weitere Benutzer erstellen, beispielsweise:

```
bin/pwutil modeler "data miner2" >> config/passwords.cfg
```

Hinweis: Mit einem einzigen > wird der Inhalt von *passwords.cfg* jedes Mal überschrieben und die zuvor festgelegten Benutzer werden ersetzt. Beachten Sie, dass unabhängig hiervon alle Benutzer dasselbe UNIX-Konto verwenden.

Hinweis: Wenn Sie der privaten Kennwortdatenbank neue Benutzer hinzufügen, während SPSS Modeler Server ausgeführt wird, müssen Sie SPSS Modeler Server neu starten, damit die neu definierten Benutzer erkannt werden. Ansonsten schlagen Anmeldeversuche aller neuen Benutzer fehl, die seit dem letzten Neustart von SPSS Modeler Server mit *pwutil* hinzugefügt wurden.

7. Ändern Sie die Besitzerrechte für das IBM SPSS Modeler-Installationsverzeichnis und des gesamten Inhalts rekursiv für den Benutzer *<modelerserv>* und die Gruppe *<modelerusers>*, wobei die zuvor von Ihnen definierten Namen referenziert werden. Beispiel:

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

8. Erstellen Sie Unterverzeichnisse im Datenverzeichnis für Ihre IBM SPSS Modeler-Benutzer, sodass sie einen Ort haben, an dem sie störungsfrei Ihre Arbeitsdaten speichern können. Diese Verzeichnisse sollten im Besitz der Gruppe *<modelerusers>* sein, die über Lese-, Schreib- und Suchberechtigungen für das Verzeichnis verfügen sollte. So erstellen Sie beispielsweise ein Arbeitsverzeichnis für den Benutzer *bob*

```
mkdir data/bob
chown bob:modelerusers data/bob
chmod ug=rwx,o= data/bob
```

Außerdem können Sie für das Verzeichnis das Bit "set-group-ID" festlegen, sodass alle Datendateien, die in das Verzeichnis kopiert werden, automatisch in den Besitz der Gruppe *<modelerusers>* übergehen:

```
chmod g+s data/bob
```

Ausführen von SPSS Modeler Server als Nicht-Root-Benutzer

So führen Sie SPSS Modeler Server als Nicht-Root-Benutzer aus:

1. Melden Sie sich mit dem zuvor erstellten Nicht-Root-Benutzerkonto an.
2. Wenn Sie bei der Ausführung die Konfigurationsdateioption *start_process_as_login_user* aktiviert haben, können Sie SPSS Modeler Server starten, stoppen und den Status überprüfen. Weitere Informationen finden Sie in „Starten, Stoppen und Statusprüfung unter UNIX“ auf Seite 27.

Endbenutzer stellen eine Verbindung mit SPSS Modeler Server her, indem sie sich über die Client-Software anmelden. Sie müssen Endbenutzern die nötigen Informationen zum Herstellen einer Verbindung übermitteln, inklusive der IP-Adresse oder dem Hostnamen des Server-Computers.

Ausführen von IBM SPSS Modeler Server als Nicht-Root-Benutzer unter AIX

Beim Start von IBM SPSS Modeler Server unter AIX als Non-Root-Benutzer besteht die Möglichkeit, dass der Fehler *lsattr not found* an der Konsole ausgegeben wird. Das liegt daran, dass *lsattr* sich im Verzeichnis */usr/sbin* befindet, aber wenn IBM SPSS Modeler Server einen neuen Benutzer (als Non-Root) anmeldet, wird der Pfad so zurückgesetzt, dass er nur */usr/bin* und */bin* enthält.

Wenn dieser Fehler auftritt, legen Sie die folgende Verknüpfung unter */usr/bin* an:

```
su
cd /usr/bin
ln -s /usr/sbin/lsattr lsattr
```

Starten Sie IBM SPSS Modeler Server nach dem Anlegen der Verknüpfung neu.

Fehlerbehebung bei fehlgeschlagener Benutzerauthentifizierung

Je nachdem, wie das Betriebssystem für die Durchführung der Authentifizierung konfiguriert ist, kann es bei der Ausführung in einer Nicht-Root-Konfiguration Probleme mit der Anmeldung bei SPSS Modeler Server geben. Dies kann beispielsweise dann vorkommen, wenn das Betriebssystem (mit der Datei */etc/nsswitch.conf* o. ä.) so konfiguriert ist, dass die lokale Kennwortspiegeldatei überprüft wird, anstatt NIS oder LDAP zu verwenden. Dies geschieht, da SPSS Modeler Server über Lesezugriff auf die Dateien verfügen muss, die zur Durchführung der Authentifizierung verwendet werden, einschließlich der Datei */etc/shadow* bzw. ihrer Entsprechung (z. B. */etc/security/passwd* unter AIX), in der sichere Informationen zum Benutzerkonto gespeichert werden. Die Dateiberechtigungen des Betriebssystems sind jedoch im Allgemeinen so festgelegt, dass nur der Root-Benutzer auf die Datei */etc/shadow* zugreifen kann. Unter diesen Umständen kann ein Nicht-Root-Prozess */etc/shadow* nicht zur Überprüfung der Benutzerkennwörter lesen, was zu einem Authentifizierungsfehler führt.

Es gibt mehrere Möglichkeiten zur Lösung dieses Problems:

- Bitten Sie den Systemadministrator, das Betriebssystem so zu konfigurieren, dass NIS oder LDAP für die Authentifizierung verwendet wird.
- Ändern Sie die Dateiberechtigungen für die geschützten Dateien, beispielsweise indem Sie Lesezugriff auf die Datei */etc/shadow* erteilen, sodass das lokale Benutzerkonto, das zur Ausführung von SPSS Modeler Server verwendet wird, auf die Datei zugreifen kann. Diese Problemumgehung mag in Produktionsumgebungen unpraktisch sein, sie könnte jedoch vorübergehend in einer Testumgebung angewendet werden, um zu überprüfen, ob das Fehlschlagen der Autorisierung mit der Betriebssystemkonfiguration zu tun hat.
- Geben Sie eine Zugriffssteuerungsliste (ACL) für die Datei */etc/shadow* an.
- Führen Sie SPSS Modeler Server als Root aus, damit die Serverprozesse die Datei */etc/shadow* lesen können.

Vorsicht:

Stellen Sie in diesem Fall sicher, dass die Datei *options.cfg* für SPSS Modeler Server die Option *start_process_as_login_user*, N beinhaltet, um die weiter oben erläuterten Sicherheitsprobleme zu vermeiden.

Anhang D. Konfigurieren und Ausführen von SPSS Modeler Server mit privater Kennwortdatei unter Windows

Einführung

In den folgenden Anweisungen finden Sie Informationen zur Ausführung von IBM SPSS Modeler Server mit einer privaten Kennwortdatei auf Windows-Systemen. Bei dieser Methode authentifiziert IBM SPSS Modeler die Benutzer anhand einer privaten Kennwortdatenbank, die sich von der Systemauthentifizierung unter Windows unterscheidet.

Konfigurieren einer privaten Kennwortdatenbank

Wenn Sie auswählen, dass die Benutzer über eine private Kennwortdatenbank authentifiziert werden sollen, werden alle Sitzungen mit demselben Benutzerkonto gestartet.

So konfigurieren Sie SPSS Modeler Server hierzu:

1. Erstellen Sie das Benutzerkonto, auf dem SPSS Modeler Server ausgeführt werden soll. Dieses Konto dient zur ausschließlichen Verwendung für den SPSS Modeler Server-Dämonprozess. Sie müssen den Dämonprozess mit diesem Benutzerkonto auf der Registerkarte zur Anmeldung des Dienstes für SPSS Modeler Server 18.1 starten. Für dieses Beispiel wird der Name *modelerserv* verwendet.
2. Bearbeiten Sie die Datei mit den Serveroptionen (*config/options.cfg*), um die Option *start_process_as_login_user* auf *Y* zu setzen und die Authentifizierung mit der privaten Kennwortdatenbank einzustellen, indem Sie die folgende Zeile anhängen:
`authentication_methods, "pasw_modeler"`
3. Anschließend erstellen Sie eine private Kennwortdatenbank, die in der Datei *config/passwords.cfg* gespeichert wird. Die Kennwortdatei definiert die Kombinationen aus Benutzername/Kennwort, die für die Anmeldung bei SPSS Modeler zulässig sind. Beachten Sie, dass diese Kombinationen für SPSS Modeler bestimmt (privat) sind und nicht in Zusammenhang mit den Benutzernamen und Kennwörtern stehen, die für die Anmeldung an Windows verwendet werden. Sie können dieselben Benutzernamen verwenden, wenn dies praktikabel ist, die Verwendung derselben Kennwörter ist jedoch nicht zulässig.

Zum Erstellen der Kennwortdatei benötigen Sie das Dienstprogramm für Kennwörter (*pwutil*) im Verzeichnis *bin* der SPSS Modeler Server-Installation. Die Syntax dieses Programms lautet:

```
pwutil [ Benutzername [ Kennwort ] ]
```

Das Programm nimmt einen Benutzernamen und ein Nur-Text-Kennwort und schreibt den Benutzernamen und das verschlüsselte Kennwort in einem für die Kennwortdatei geeigneten Format in die Standardausgabe. Zum Definieren eines Benutzers mit dem Namen *modeler* und dem Kennwort *data mining* würden Sie zum Beispiel in einer DOS-Eingabeaufforderung zum SPSS Modeler Server-Installationsverzeichnis navigieren und dann Folgendes eingeben:

```
bin\pwutil modeler "data mining" > config\passwords.cfg
```

Anmerkung: Stellen Sie sicher, dass es nur 1 Instanz jedes Benutzers in der Datei gibt; Duplikate verhindern, dass SPSS Modeler Server gestartet wird.

Das Definieren eines einzigen Benutzernamens ist in den meisten Fällen, in denen sich alle Benutzer mit demselben Namen und Kennwort anmelden, ausreichend. Mit dem Operator *>>*, der an die einzelnen Dateien angehängt wird, können Sie jedoch weitere Benutzer erstellen. Beispiel:

```
bin\pwutil modeler "data miner2" >> config\passwords.cfg
```

Anmerkung:

Wird nur ein einziges > verwendet, wird der Inhalt von passwords.cfg jedes Mal überschrieben und die zuvor festgelegten Benutzer werden ersetzt. Beachten Sie, dass unabhängig hiervon alle Benutzer dasselbe UNIX-Konto verwenden.

Wenn Sie der privaten Kennwortdatenbank neue Benutzer hinzufügen, während SPSS Modeler Server ausgeführt wird, müssen Sie SPSS Modeler Server neu starten, damit die neu definierten Benutzer erkannt werden. Ansonsten schlagen Anmeldeversuche aller neuen Benutzer fehl, die seit dem letzten Neustart von SPSS Modeler Server mit pwutil hinzugefügt wurden.

4. Geben Sie dem in Schritt 1 erstellten Benutzer uneingeschränkten Zugriff auf die Datei mit den Serveroptionen config\options.cfg und das Verzeichnis %ALLUSERSPROFILE%\IBM\SPSS.
5. Stoppen Sie in den Systemdiensten den **IBM SPSS Modeler Server**-Dienst und ändern Sie über **Lokales Systemkonto** die Anmeldung in das Benutzerkonto, das in Schritt 1 erstellt wurde. Starten Sie den Dienst anschließend erneut.

Anhang E. Lastverteilung mit Server-Clustern

Mit IBM SPSS Collaboration and Deployment Services kann das Plug-in Coordinator of Processes verwendet werden, um Services im Netz zu verwalten. Der Coordinator of Processes stellt Serververwaltungsfunktionen bereit, mit denen die Kommunikation und Verarbeitung zwischen Client und Server optimiert werden können.

Die zu verwaltenden Services, beispielsweise IBM SPSS Statistics Server oder IBM SPSS Modeler Server registrieren sich beim Start beim Coordinator of Processes und senden in regelmäßigen Abständen aktualisierte Statusnachrichten. Die Services können außerdem alle notwendigen Konfigurationsdateien in IBM SPSS Collaboration and Deployment Services Repository speichern und bei der Initialisierung abrufen.

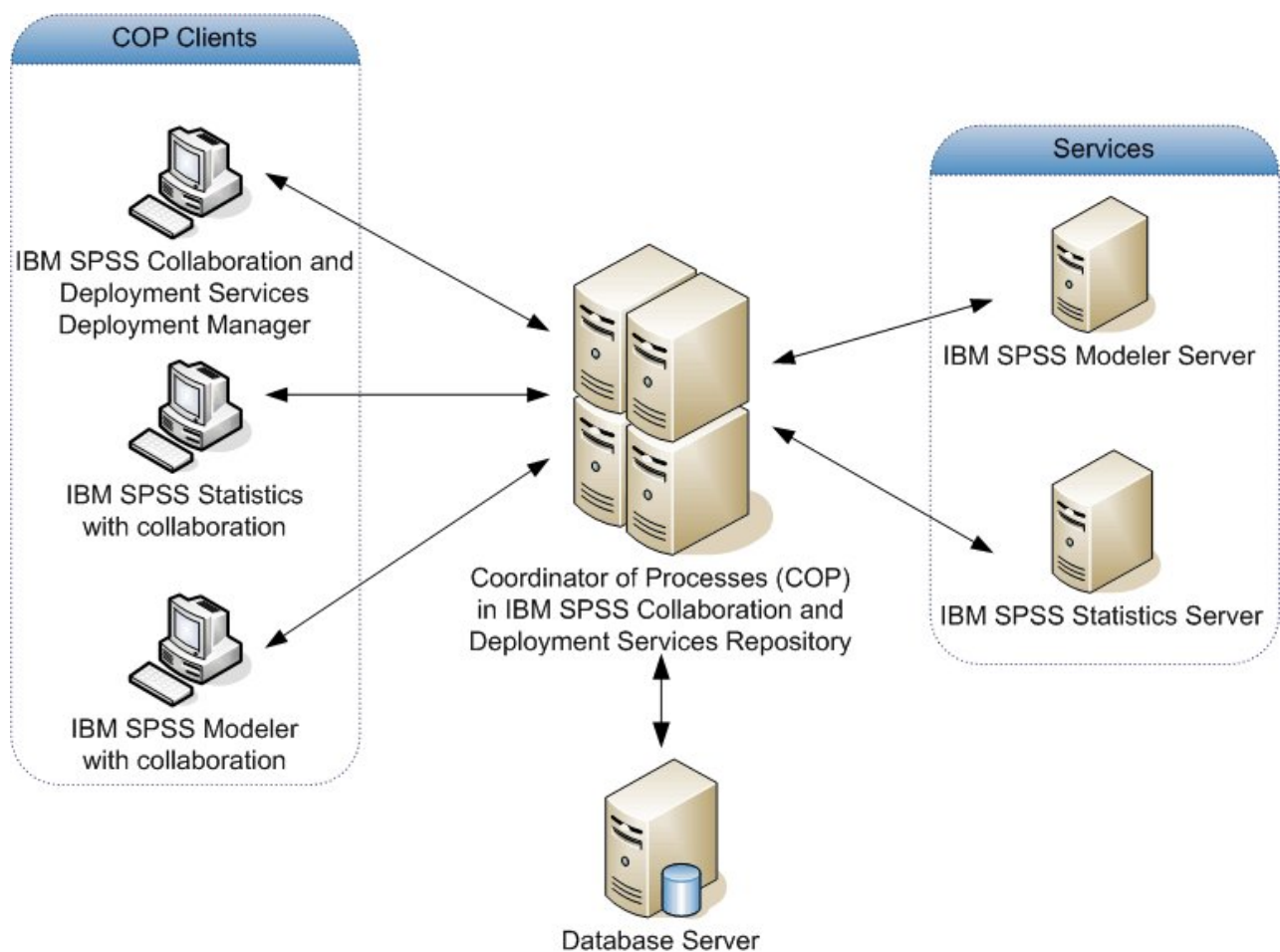


Abbildung 5. Architektur des Coordinator of Processes

Durch die Ausführung von IBM SPSS Modeler-Streams auf einem Server kann die Leistungsfähigkeit erhöht werden. In einigen Fällen haben Sie nur die Wahl zwischen einem Server und zwei Servern. In anderen Fällen kann eine größere Auswahl von Servern zur Verfügung stehen, da es einen substantziellen Unterschied zwischen den einzelnen Servern gibt, wie beispielsweise Inhaber, Zugriffsrechte, Serverdaten, ob es sich um einen Testserver oder einen Produktionsserver handelt usw. Wenn Coordinator of Processes in Ihrem Netz installiert ist, wird Ihnen möglicherweise zusätzlich ein Server-Cluster angeboten.

Ein Server-Cluster ist eine Gruppe von Servern, die hinsichtlich Konfiguration und Ressourcen austauschbar sind. Coordinator of Processes ermittelt, welcher Server am besten für die Antwort auf eine Verarbeitungsanforderung geeignet ist. Hierfür verwendet er einen Algorithmus, der die Last anhand mehrerer Kriterien ausgleicht, wie beispielsweise die Gewichtung der einzelnen Server, Benutzerprioritäten und die aktuelle Verarbeitungsauslastungen. Weitere Informationen finden Sie im Handbuch *Coordinator of Processes Service Developer Guide*, das in der Dokumentationssuite für IBM SPSS Collaboration and Deployment Services zur Verfügung gestellt wird.

Wenn Sie in IBM SPSS Modeler eine Verbindung zu einem Server oder Server-Cluster herstellen, können Sie den Server manuell aufrufen oder mithilfe von Coordinator of Processes nach einem bestimmten Server oder Cluster suchen. Weitere Informationen finden Sie in „Verbinden mit IBM SPSS Modeler Server“ auf Seite 15.

Anhang F. LDAP-Authentifizierung

Die folgenden Anweisungen stellen grundlegende Richtlinien dar, anhand derer Sie SPSS Modeler Server unter UNIX für die Verwendung der LDAP-Authentifizierung konfigurieren können, wobei die Identitäten der Benutzer, die sich am Server anmelden, in einem LDAP-Verzeichnis gespeichert werden.

Anmerkung: Voraussetzung hierfür ist eine ordnungsgemäße Konfiguration der LDAP-Client-Software auf dem Hostbetriebssystem. Weitere Informationen finden Sie in der Originaldokumentation des Herstellers.

Eine weitere Konfiguration ist in der Regel nicht erforderlich und die Verwendung von LDAP ist für den Server nicht ersichtlich. Unter den folgenden Umständen sind beispielsweise keine weiteren Änderungen erforderlich:

- Die LDAP-Client-Software und LDAP-Server-Software sind gemäß RFC 2307 konfiguriert.
- Zugriff auf die Datenbank passwd (und soweit zutreffend shadow) wird an LDAP umgeleitet (z. B. in `nsswitch.conf`).
- Für jeden gültigen Benutzer von SPSS Modeler Server ist ein Eintrag passwd (und shadow) vorhanden, der im LDAP-Verzeichnis gespeichert wird.
- Der SPSS Modeler Server-Service wird über das Benutzerkonto mit Rootberechtigung gestartet.

Unter den folgenden beiden Bedingungen muss SPSS Modeler Server möglicherweise für LDAP konfiguriert werden:

- Wenn der Service mit einem anderen Konto als dem Benutzerkonto mit Rootberechtigung gestartet wird, ist der Service möglicherweise nicht berechtigt, die Authentifizierung nach der Standardmethode durchzuführen. Grund hierfür ist in der Regel der eingeschränkte Zugriff auf die Spiegeldatenbank.
- Wenn für die Benutzer kein Eintrag passwd (oder shadow) im Verzeichnis gespeichert ist, das heißt, sie über keine für die Anmeldung am Hostsystem gültigen Benutzeridentitäten verfügen.

Das LDAP-Authentifizierungsverfahren verwendet das PAM-Subsystem und erfordert das Vorhandensein eines für das Hostbetriebssystem ordnungsgemäß konfigurierten PAM-LDAP-Moduls. Weitere Informationen finden Sie in der Originaldokumentation des Herstellers.

Führen Sie die folgenden Schritte aus, um SPSS Modeler Server für die ausschließliche Verwendung der LDAP-Authentifizierung zu konfigurieren.

Anmerkung: Mit den folgenden Schritten wird die Mindestkonfiguration vorgenommen, die zur Funktionsfähigkeit ausreichen sollte. Abhängig von Ihrem Betriebssystem und Ihrer lokalen Sicherheitsrichtlinie können weitere Optionen oder andere Einstellungen erforderlich sein. Weitere Informationen finden Sie in der Originalbetriebsdokumentation.

1. Bearbeiten Sie die Servicekonfigurationsdatei (`options.cfg`) und fügen Sie die folgende Zeile hinzu (bzw. bearbeiten Sie die Zeile entsprechend): `authentication_methods, pam`. Diese Zeile weist den Server an, die PAM-Authentifizierung gegenüber der Standardauthentifizierung bevorzugt zu verwenden.
2. Stellen Sie für den SPSS Modeler Server-Service eine PAM-Konfiguration bereit; hierfür sind häufig Rootberechtigungen erforderlich. Der Service wird durch den Namen `modelerserver` identifiziert.
3. Erstellen Sie auf einem Linux-/UNIX-System, das das Verzeichnis `/etc/pam.d` verwendet, in diesem Verzeichnis eine Datei mit dem Namen `modelerserver` und fügen Sie Inhalt wie den im folgenden Beispiel dargestellten hinzu:

```
# IBM SPSS Modeler Server
auth required pam_ldap.so
account required pam_ldap.so
password required pam_deny.so
session required pam_deny.so
```

4. Bearbeiten Sie bei älteren Systemen, die `/etc/pam.conf` verwenden, diese Datei und fügen Sie Zeilen wie die im folgenden Beispiel dargestellten hinzu:

```
# IBM SPSS Modeler Server
modelerserver auth required pam_ldap.so
modelerserver account required pam_ldap.so
modelerserver password required pam_deny.so
modelerserver session required pam_deny.so
```

5. Die Namen der referenzierten PAM-Module variieren je nach Betriebssystem. Stellen Sie fest, welche Module für Ihr Hostbetriebssystem erforderlich sind.

Anmerkung: Die Zeilen in Schritt 3 und 4 geben an, dass SPSS Modeler Server zur Authentifizierung und Kontoverwaltung das PAM-LDAP-Modul heranziehen muss. Kennwortänderungen und das Sitzungsmanagement werden jedoch nicht unterstützt, sodass diese Aktionen nicht zulässig sind. Wenn das Kontomanagement nicht erforderlich oder nicht anwendbar ist, ändern Sie wie im folgenden Beispiel dargestellt die betreffende Zeile, um alle Anforderungen zuzulassen:

```
# IBM SPSS Modeler Server
auth required pam_ldap.so
account required pam_permit.so
password required pam_deny.so
session required pam_deny.so
```

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

Sonderzeichen

@SQLFN-Funktion 75

Numerische Stichwörter

64-Bit-Betriebssysteme 9

A

Administratorzugriff
für IBM SPSS Modeler Server 39
mit Benutzerzugriffssteuerung (UAC) 39
allow_modelling_memory_override options.cfg (Datei) 42
Anmeldung bei IBM SPSS Modeler Server 15
Anwendungsbeispiele 3
Arbeitsspeicher 11
Architektur
Komponenten 7
Authentifizierung 23
auto.sh (UNIX)
Speicherort 79
Automatischer Serverstart
Konfiguration unter UNIX 79

B

Beispiele
Anwendungshandbuch 3
Übersicht 4
Benutzer-ID
IBM SPSS Modeler Server 15
Benutzerauthentifizierung 23
Benutzerkonten
Berechtigungen 24
IBM SPSS Modeler Server 23
Berechtigungen 24
Betriebssysteme
64 Bit 9

C

cache_compression options.cfg (Datei) 42
cache_connection, Option 49
Cachekomprimierung 42
Caching, datenbankinternes 48
chemsrv.sh (UNIX)
Speicherort 79
CLEM-Ausdrücke
SQL-Generierung 73
Client
Single Sign-on 19
Cognos-SSL-Verbindung 54
Cognos TM1-SSL-Verbindung 55

Coordinator of Processes 22
Lastverteilung 93
Server-Cluster 93
Coordinator of Processes, Konfiguration für IBM SPSS Modeler Server 45
COP 22
Lastverteilung 93
Server-Cluster 93
cop_enabled options.cfg (Datei) 45
cop_host options.cfg (Datei) 45
COP-Konfiguration für IBM SPSS Modeler Server 45
cop_password options.cfg (Datei) 45
cop_port_number options.cfg (Datei) 45
cop_service_description options.cfg (Datei) 45
cop_service_host options.cfg (Datei) 45
cop_service_name options.cfg (Datei) 45
cop_service_weight options.cfg (Datei) 45
cop_update_interval options.cfg (Datei) 45
cop_user_name options.cfg (Datei) 45

D

Data Access Pack
Fehlerbehebung für ODBC unter UNIX 83
Konfigurieren von UNIX 80
ODBC, Konfiguration unter UNIX und UNIX-Bibliothekspfade 86
data_file_path options.cfg (Datei) 41
data_files_restricted options.cfg (Datei) 41
Dateiberechtigungen
IBM SPSS Modeler Server 25
Konfiguration unter UNIX 80
Dateinamen
UNIX 12
Windows 12
Datenbankcaching
SQL-Generierung 67
über options.cfg steuern 48
Datenbanken
Zugriff 11
Datenbankinternes Caching 48
Datenbankinternes Mining 61
Datenbankserver 61
Datenbankverbindungen
schließen 49
Datendateien
IBM SPSS Statistics 13

Datendateien (*Forts.*)
Import und Export 13
Datenquellen
Single Sign-on 21
Datenzugriff 11
DB2
SQL-Optimierung 63, 64
Dokumentation 3
Domänenname (Windows)
IBM SPSS Modeler Server 15

E

Ergebnisse
Datensatzreihenfolge 25
Runden von 25
Unterschiede zwischen Client und Server 25

F

Fehler bei der Streamausführung 42
Festplatten 10
FIPS-Verschlüsselung 46
Firewall-Einstellungen
options.cfg (Datei) 42

G

Geisterprozesse, IBM SPSS Modeler Server 28
group_configuration 46

H

Hardwareempfehlungen für IBM SPSS Modeler Server 9
Hostname
IBM SPSS Modeler Server 15, 22

I

IBM SPSS Analytic Server
Konfigurationsoptionen 41
IBM SPSS Modeler 1
Dokumentation 3
IBM SPSS Modeler Administration Console 38
Administratorzugriff 39
UAC-Zugriff 39
IBM SPSS Modeler-Client
Single Sign-on 19
IBM SPSS Modeler Server 1
Administratorzugriff 39
andere Ergebnisse als der Client 25
Benutzer-ID 15
Benutzerauthentifizierung 23
Benutzerkonten 23
Berechtigungen 24

- IBM SPSS Modeler Server (*Forts.*)
 - Coordinator of Processes, Konfiguration 45
 - COP-Konfiguration 45
 - Dateierstellung 25
 - Domänenname (Windows) 15
 - Hostname 15, 22
 - Informationen für Endbenutzer 23
 - Kennwort 15
 - Konfigurationsoptionen 40
 - nicht reagierende Prozesse 28
 - Portnummer 15, 22, 40
 - Serverprozesse 48
 - Single Sign-on 16, 19
 - Single Sign-on für Datenquellen 21
 - temporäres Verzeichnis 41
 - UAC-Zugriff 39
 - Überwachung 48
 - Verwaltung 38
 - Verwaltungsoptionen 38
- IBM SPSS Modeler Server starten
 - unter UNIX 27
 - unter Windows 27
- IBM SPSS Modeler Server stoppen
 - unter UNIX 27
 - unter Windows 27
- IBM SPSS Modeler Server-Verbindungen
 - hinzufügen 22
- IBM SPSS Statistics-Datendateien
 - Import und Export 13
- IBM SPSS Statistics-Datenzugriffstechnologie 11
- io_buffer_size
 - options.cfg (Datei) 42

K

- Kennwort
 - IBM SPSS Modeler Server 15
- Kerberos 46
- Kernelgrenzwerte unter UNIX 27
- Knoten
 - Unterstützung von SQL-Generierung 69, 76
- Knotencaching
 - in Datenbank schreiben 67
 - SQL-Generierung 67
- Konfigurationsoptionen
 - Anmeldeversuche 40
 - automatische SQL-Generierung 44
 - Coordinator of Processes 45
 - COP 45
 - Leistung und Optimierung 42
 - parallele Verarbeitung 42
 - Portnummer 40
 - Speicherverwaltung 42
 - SQL-Zeichenfolgenlänge 44
 - SSL-Datenverschlüsselung 44
 - Streamneuerstellung 42
 - temporäres Verzeichnis 41
 - Übersicht 40, 41
 - Verbindungen und Sitzungen 40
 - von IBM SPSS Modeler Server 38
 - Zugriff auf Datendateien 41

L

- LDAP
 - Authentifizierung 95
- Leistung
 - von IBM SPSS Modeler Server 59
- Linux
 - Single Sign-on 19

M

- max_file_size
 - options.cfg (Datei) 41
- max_login_attempts
 - options.cfg (Datei) 40
- max_parallelism
 - options.cfg (Datei) 42
- max_sessions
 - options.cfg (Datei) 40
- max_sql_string_length
 - options.cfg (Datei) 44
- Mehrere Streams, Ausführung 42
- memory_usage
 - options.cfg (Datei) 42
- Microsoft SQL Server
 - SQL-Optimierung 63, 64
- Modellierung
 - Speicherverwaltung 42
- modelling_memory_limit_percentage
 - options.cfg (Datei) 42
- Modellnuggets
 - SQL anzeigen für 67

N

- Nachrichten
 - generiertes SQL anzeigen 67
- Neustart des Web-Service 38

O

- ODBC
 - Konfiguration unter UNIX 80
- ODBC-Datenquellen
 - ODBC- und UNIX-Scripts 80
 - und UNIX 80
- ODBC-Treibersetup 63
- Operatoren
 - SQL-Generierung 73
- Optimierung
 - SQL-Generierung 63, 64, 66
- options.cfg 46
- options.cfg (Datei) 48
- Oracle
 - SQL-Optimierung 63, 64, 77

P

- PAM
 - Authentifizierung 95
- Parallele Verarbeitung
 - Steuerung 42
- Pfade 12
- port_number
 - options.cfg (Datei) 40

- Porteinstellungen
 - options.cfg (Datei) 42
- Portnummer
 - IBM SPSS Modeler Server 15, 22, 40
- program_file_path
 - options.cfg (Datei) 41
- program_files_restricted
 - options.cfg (Datei) 41
- Protokolldateien
 - für IBM SPSS Modeler Server 57
 - generiertes SQL anzeigen 67
- Prozesse, nicht reagierend 28
- Prozessoren 9
 - mehrfach 42
- Pushbacks 63, 64, 66, 69, 76
 - CLEM-Ausdrücke 73

R

- RAM 11
- rc.modeler (UNIX)
 - Speicherort 79

S

- Secure Sockets Layer 50
- Server
 - Anmeldung 15
 - Single Sign-on 16, 19
 - Suche nach COP für Server 22
 - Verbindungen hinzufügen 22
- server_logging.log 57
- Server-Port, Einstellungen
 - options.cfg (Datei) 42
- Sicherheit
 - Dateierstellung 25
 - Dateierstellung unter UNIX konfigurieren 80
 - SSL 50
- Single Sign-on 15
- Speicherplatz 10
 - berechnen 10
- Speicherverwaltung
 - Verwaltungsoptionen 42
- SQL
 - Abfragen 75
 - Anzeige für Modellnuggets 67
 - doppelte Spaltennamen 75
 - Oracle optimieren 77
 - Vorschau für generiertes 67
- sql_generation_enabled
 - options.cfg (Datei) 44
- SQL-Generierung 63, 64, 69, 76
 - Aktivierung 66
 - Anzeige für Modellnuggets 67
 - CLEM-Ausdrücke 67, 73
 - für IBM SPSS Modeler Server aktivieren 44
 - Protokollierung 67
 - Streamneuerstellung 67
 - Tipps 67
 - Vorschau 67
- SQL-Pushback. Siehe auch "SQL-Generierung" 63
- SQL Server
 - SQL-Optimierung 63, 64

- SSL 50
 - Cognos TM1-Verbindung 55
 - Cognos-Verbindung 54
 - Kommunikation schützen 50
 - Übersicht 50
- ssl_certificate_file
 - options.cfg (Datei) 44
- SSL-Datenverschlüsselung
 - für IBM SPSS Modeler Server aktivieren 44
- ssl_enabled
 - options.cfg (Datei) 44
- ssl_private_key_file
 - options.cfg (Datei) 44
- ssl_private_key_password
 - options.cfg (Datei) 44
- Status
 - von IBM SPSS Modeler Server unter UNIX 27
 - von IBM SPSS Modeler Server unter Windows 27
- stream_rewriting_enabled
 - options.cfg (Datei) 42
- Streamneuerstellung 67
- Suche nach COP für Verbindungen 22

T

- temp_directory
 - options.cfg (Datei) 41
- Temporäre Dateien 10
 - Berechtigungen für (IBM SPSS Modeler Server) 25
- Temporäres Verzeichnis
 - für IBM SPSS Modeler Server 41

U

- UNC-Dateinamen 12
- UNIX
 - Benutzerauthentifizierung 23
 - Berechtigungen 24
 - Bibliothekspfade 86
 - Dateiberechtigungen konfigurieren 80
 - Neustart des Web-Service 38
 - Single Sign-on 19
- UNIX-Kernelgrenzwerte 27
- UNIX-Scripts
 - auto.sh 79
 - bearbeiten 80
 - modelersrv.sh 79
 - rc.modeler 79
- UNIX-Shell 46

V

- Verbindungen
 - Server-Cluster 22
 - zu IBM SPSS Modeler Server 15, 22
- Verschlüsselung
 - FIPS 46
 - SSL 50
- Verwaltung
 - von IBM SPSS Modeler Server 38

- Violette Knoten
 - SQL-Optimierung 65
- Vorschau
 - SQL-Generierung 67

W

- Web-Service, Neustart 38
- Windows
 - Neustart des Web-Service 38

