

IBM SPSS Modeler Server
14.2-Verwaltungs- und
Leistungshandbuch



Hinweis: Lesen Sie vor der Verwendung dieser Informationen und des zugehörigen Produkts die allgemeinen Informationen unter Hinweise auf S. .

Dieses Dokument enthält eigentumsrechtlich geschützte Informationen von SPSS Inc., an IBM Company. Sie werden im Rahmen einer Lizenzvereinbarung bereitgestellt und sind durch Copyright-Gesetze geschützt. Die in dieser Publikation enthaltenen Informationen umfassen keinerlei Produktgewährleistungen und keine der Aussagen in diesem Handbuch darf als solche ausgelegt werden.

Wenn Sie Informationen an IBM bzw. SPSS senden, räumen Sie IBM und SPSS das nicht ausschließliche Recht ein, die Informationen in jeglicher Form zu verwenden bzw. weiterzugeben, die dem Unternehmen geeignet erscheint, ohne dass ihm daraus Verbindlichkeiten Ihnen gegenüber entstehen.

© Copyright IBM Corporation 1994, 2011..

Vorwort

IBM® SPSS® Modeler ist die auf Unternehmensebene einsetzbare Data-Mining-Workbench von IBM Corp.. 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 benutzen 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 visuelle Benutzeroberfläche von SPSS Modeler erleichtert die Anwendung des spezifischen Geschäftswissens der Benutzer, was zu leistungsstärkeren Vorhersagemodellen führt und die Zeit bis zur Lösungserstellung verkürzt. 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 für Entscheidungsträger oder in einer Datenbank.

Über IBM Business Analytics

IBM Business Analytics-Software bietet vollständige, einheitliche und genaue Informationen, auf die Entscheidungsträger vertrauen, um die Unternehmensleistung zu steigern. Ein umfassendes Portfolio von Anwendungen für [Unternehmensinformationen](#), [Vorhersageanalysen](#), [Verwaltung der Finanzleistung und Strategie](#) sowie [Analysen](#) bietet sofort klare und umsetzbare Einblicke in die aktuelle Leistung und ermöglicht die Vorhersage zukünftiger Ergebnisse. In Kombination mit umfassenden Branchenlösungen, bewährten Vorgehensweisen und professionellen Dienstleistungen können Unternehmen jeder Größe optimale Produktivität erreichen, die Entscheidungsfindung zuverlässig automatisieren und bessere Ergebnisse erzielen.

Als Teil dieses Portfolios unterstützt die IBM SPSS Predictive Analytics-Software Unternehmen dabei, zukünftige Ereignisse vorherzusagen und aktiv auf diese Erkenntnisse zu reagieren, um bessere Geschäftsergebnisse zu erzielen. Kunden aus den Bereichen Wirtschaft, Behörden und Bildung aus aller Welt verlassen sich auf die IBM SPSS-Technologie. Sie bringt Ihnen beim Gewinnen, Halten und Ausbauen neuer Kundenbeziehungen einen Wettbewerbsvorteil und verringert gleichzeitig das Betrugs- sowie andere Risiken. Durch Integration der IBM SPSS-Software in den täglichen Betrieb können diese Unternehmen qualifizierte Vorhersagen treffen und dadurch die Entscheidungsfindung so ausrichten und automatisieren, dass Geschäftsziele erreicht werden und ein messbarer Wettbewerbsvorteil entsteht. Wenn Sie weitere Informationen wünschen oder einen Mitarbeiter kontaktieren möchten, ist dies unter <http://www.ibm.com/spss> möglich.

Technischer Support

Kunden mit Wartungsvertrag können den technischen Support in Anspruch nehmen. Kunden können sich an den technischen Support wenden, wenn sie Hilfe bei der Arbeit mit IBM Corp.-Produkten oder bei der Installation in einer der unterstützten Hardware-Umgebungen benötigen. Die Kontaktdaten des Technischen Supports finden Sie auf der IBM Corp.-Website

unter <http://www.ibm.com/support>. Sie müssen bei der Kontaktaufnahme Ihren Namen, Ihre Organisation und Ihre Supportvereinbarung angeben.

Inhalt

1	Informationen zu IBM SPSS Modeler	1
	IBM SPSS Modeler Server	1
	IBM SPSS Modeler-Optionen	2
	IBM SPSS Text Analytics	2
	IBM SPSS Modeler-Dokumentation	3
	Anwendungsbeispiele	4
	Ordner "Demos"	4
2	Architektur- und Hardwareempfehlungen	6
	IBM SPSS Modeler-Architektur	6
	Architekturbeschreibung	6
	Hardwareempfehlungen	8
	Temporäre Festplattenspeicher- und Arbeitsspeicheranforderungen	9
	Datenzugriff	11
	Referenzieren von Datendateien	13
	Importieren von IBM SPSS Statistics-Datendateien	13
	Installationsanweisungen	13
3	Support für IBM SPSS Modeler	14
	Verbindung mit IBM SPSS Modeler Server	14
	Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung.	16
	Suchen nach Servern in IBM SPSS Collaboration and Deployment Services	17
	Daten- und Dateisysteme	18
	Benutzerauthentifizierung	19
	Dateierstellung	19
	Unterschiede bei den Ergebnissen	20
4	IBM SPSS Modeler Server-Verwaltung	21
	Starten und Stoppen von IBM SPSS Modeler Server	21
	So funktioniert das Starten, Stoppen bzw. die Statusprüfung unter Windows:	21
	So funktioniert das Starten, Stoppen bzw. die Statusprüfung unter UNIX:	21
	Umgang mit nicht reagierenden Server-Prozessen (UNIX-Systeme)	22

Verwaltung	23
Verwendung von IBM SPSS Modeler Administration Console	24
Starten von Modeler Administration Console	24
Konfigurieren von Zugriffsrechten mit Modeler Administration Console	25
SPSS Modeler Server-Verbindungen	26
SPSS Modeler Server-Konfiguration	27
SPSS Modeler Server-Überwachung	34
Verwenden der Datei options.cfg	34
Nicht verwendete Datenbankverbindungen schließen	35
IBM SPSS Statistics-Lizenzspeicherort	36
Verwenden von SSL zur sicheren Datenübertragung	36
Wie SSL funktioniert	37
Schützen der Client-Server- und Server-Server-Kommunikation durch SSL	37
Konfiguration des URL-Präfix	41
Schützen von LDAP mit SSL	42
Server-Protokoll	43

5 Leistungsblick 45

Leistungs- und Optimierungseinstellungen für Ihren Server	45
Leistungs- und Optimierungseinstellungen für Ihren Client	45
Datenbankverwendung und -optimierung	48
SQL-Optimierung	49

6 SQL-Optimierung 50

Funktionsweise der SQL-Erzeugung	51
Beispiel für SQL-Erzeugung	52
Konfigurieren der SQL-Optimierung	53
Vorschau für erzeugte SQL	54
SQL für Modell-Nuggets anzeigen	56
Tipps für die Maximierung der SQL-Erzeugung	56
Knoten zur Unterstützung der SQL-Erzeugung	58
CLEM-Ausdrücke und Operatoren, die die SQL-Erzeugung unterstützen	61
Verwenden von SQL-Funktionen in CLEM-Ausdrücken	63
Schreiben von SQL-Abfragen	64

Anhänge

A Konfigurieren von Oracle für UNIX-Plattformen 65

Konfigurieren von Oracle für die SQL-Optimierung 65

B Konfigurieren von UNIX-Startskripts 67

Einführung 67

Skripts 67

Automatisches Starten und Stoppen von IBM SPSS Modeler Server 68

Manuelles Starten und Stoppen von IBM SPSS Modeler Server 68

Bearbeiten von Skripts 68

Steuern der Berechtigungen für die Dateierstellung 69

IBM SPSS Modeler Server und das Data Access-Pack 69

 Fehlerbehebung für die Connect for ODBC-Konfiguration 72

 Bibliothekspfade 76

Konfigurieren eines Treibers für den Enterprise-Ansichts-Knoten 76

C Ausführen unter UNIX als Non-Root-Prozess 77

Einführung 77

Konfigurieren von IBM SPSS Modeler Server als Nicht-Root-Prozess 77

D Lastenausgleich mit Server-Clustern 80

E Hinweise 82

Index 85

Informationen zu IBM SPSS Modeler

IBM® SPSS® Modeler ist ein Set von Data-Mining-Tools, mit dem Sie auf der Grundlage Ihres Geschäftswissens schnell und einfach Vorhersagemodelle erstellen und zur Erleichterung der Entscheidungsfindung in die Betriebsabläufe einbinden können. 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.

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 besitzt ihre Stärken und eignet sich besonders für bestimmte Problemtypen.

SPSS Modeler kann als Standalone-Produkt oder in Verbindung mit SPSS Modeler Server erworben werden. Außerdem ist eine Reihe von Zusatzoptionen verfügbar, die in den folgenden Abschnitten kurz dargelegt werden. Weitere Informationen finden Sie unter <http://www.ibm.com/software/analytics/spss/products/modeler/>.

IBM SPSS Modeler Server

SPSS Modeler verwendet eine Client/Server-Architektur zur Verteilung von Anforderungen für ressourcenintensive Vorgänge an leistungsstarke Serversoftware, wodurch bei größeren Daten-Sets eine schnellere Leistung erzielt werden kann. Neben den hier aufgeführten Produkten können auch weitere Produkte bzw. Aktualisierungen verfügbar sein. Weitere Informationen finden Sie unter <http://www.ibm.com/software/analytics/spss/products/modeler/>.

SPSS Modeler. SPSS Modeler Clementine Client ist eine im Funktionsumfang vollständige Version des installierten Produkts und kann auf dem Desktop-Computer des Benutzers ausgeführt werden. Es kann im lokalen Modus als Standalone-Produkt oder im verteilten Modus zusammen mit IBM® SPSS® Modeler Server verwendet werden, um im Falle von großen Daten-Sets die Leistung zu verbessern.

SPSS Modeler Server. SPSS Modeler Server wird ständig im verteilten Analysemodus zusammen mit einer oder mehreren IBM® SPSS® Modeler-Installationen ausgeführt, wodurch eine herausragende Leistung bei großen Daten-Sets erzielt werden kann, da speicherintensive Operationen auf dem Server durchgeführt werden können, ohne Daten auf den Client-Computer herunterzuladen. SPSS Modeler Server bietet außerdem Unterstützung für SQL-Optimierung sowie Funktionen zur Modellierung innerhalb der Datenbank, wodurch Leistungsfähigkeit und Automatisierung weiter verbessert werden. Es muss mindestens eine SPSS Modeler-Installation vorhanden sein, um eine Analyse durchzuführen.

IBM SPSS Modeler-Optionen

Die folgenden Komponenten und Funktionen können separat erworben und für die Verwendung mit SPSS Modeler lizenziert werden. Beachten Sie, dass zu einem späteren Zeitpunkt möglicherweise noch weitere Produkte und Updates erhältlich sind. Weitere Informationen finden Sie unter <http://www.ibm.com/software/analytics/spss/products/modeler/>.

- SPSS Modeler Server-Zugriff, der eine bessere Skalierbarkeit und bessere Leistungsfähigkeit bei großen Daten-Sets bietet sowie Unterstützung für SQL-Optimierung und Funktionen zur Modellierung innerhalb der Datenbank.
- SPSS Modeler Solution Publisher, für Scoring in Echtzeit oder automatisiertes Scoring außerhalb der SPSS Modeler-Umgebung. Für weitere Informationen siehe Thema **IBM SPSS Modeler Solution Publisher** in Kapitel 2 in *IBM SPSS Modeler 14.2 Solution Publisher*.
- Adapter zum Deployment für IBM SPSS Collaboration and Deployment Services oder die Thin-Client-Anwendung IBM SPSS Modeler Advantage. Für weitere Informationen siehe Thema **Speichern und Bereitstellen von IBM SPSS Collaboration and Deployment Services Repository-Objekten** in Kapitel 9 in *IBM SPSS Modeler 14.2- Benutzerhandbuch*.

IBM SPSS Text Analytics

IBM® SPSS® Text Analytics ist ein vollständig integriertes Zusatzprodukt für SPSS Modeler, das hoch entwickelte linguistische Technologien und die Verarbeitung natürlicher Sprache (Natural Language Processing, NLP) benutzt, 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 bestehenden 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.

- Der Text-Mining-Knoten bietet die Modellierung von Konzepten und Kategorien sowie eine interaktive Workbench, in der Sie eine erweiterte Untersuchung von Textlinks und Clustern durchführen, Ihre eigenen Kategorien erstellen und die Vorlagen für linguistische Ressourcen verfeinern können.
- Eine Reihe von Importformaten wird unterstützt, darunter Blogs und andere webbasierte Quellen.
- Benutzerdefinierte Vorlagen, Bibliotheken und Wörterbücher für bestimmte Domänen, wie CRM und Genomforschung, sind ebenfalls eingeschlossen.

Anmerkung: Für den Zugriff auf diese Komponente ist eine separate Lizenz erforderlich. Weitere Informationen finden Sie unter <http://www.ibm.com/software/analytics/spss/products/modeler/>.

IBM SPSS Modeler-Dokumentation

Die vollständige Dokumentation im Online-Hilfe-Format finden Sie im Hilfe-Menü von SPSS Modeler. Dazu gehören die Dokumentation für SPSS Modeler, SPSS Modeler Server und SPSS Modeler Solution Publisher sowie das Anwendungshandbuch und weiteres Material zur Unterstützung.

Die vollständige Dokumentation für die einzelnen Produkte im PDF-Format finden Sie im Ordner *Documentation* auf der jeweiligen Produkt-DVD.

- **IBM SPSS Modeler-Benutzerhandbuch.** Allgemeine Einführung in die Verwendung von SPSS Modeler, in der u. a. die Erstellung von Daten-Streams, der Umgang mit fehlenden Werten, die Erstellung von CLEM-Ausdrücken, die Arbeit mit Projekten und Berichten sowie das Packen von Streams für das Deployment in IBM SPSS Collaboration and Deployment Services, Predictive Applications (Prognoseanwendungen) oder IBM SPSS Modeler Advantage beschrieben werden.
- **Quellen-, Prozess- und Ausgabeknoten in IBM SPSS Modeler.** 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. [Für weitere Informationen siehe Thema Überblick über Modellierungsknoten in Kapitel 3 in IBM SPSS Modeler 14.2-Modellierungsknoten.](#)
- **IBM SPSS Modeler-Algorithmushandbuch.** Beschreibung der mathematischen Grundlagen der in SPSS Modeler verwendeten Modellierungsmethoden.
- **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 Hilfe-Menü aufgerufen werden. [Für weitere Informationen siehe Thema Anwendungsbeispiele in IBM SPSS Modeler 14.2-Benutzerhandbuch.](#)
- **Skripterstellung und Automatisierung in IBM SPSS Modeler.** Informationen zur Automatisierung des Systems über Skripterstellung, einschließlich der Eigenschaften, die zur Bearbeitung von Knoten und Streams verwendet werden können.
- **IBM SPSS Modeler Deployment-Handbuch.** Informationen zum Ausführen von SPSS Modeler-Streams und -Szenarien als Schritte bei der Verarbeitung von Jobs im IBM® SPSS® Collaboration and Deployment Services Deployment Manager.
- **IBM SPSS Modeler CLEF-Entwicklerhandbuch.** CLEF bietet die Möglichkeit, Drittanbieterprogramme, wie Datenverarbeitungsroutinen oder Modellierungsalgorithmen, als Knoten in SPSS Modeler zu integrieren.
- **In-Database Mining-Handbuch für IBM SPSS Modeler.** 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- und -Leistungshandbuch.** Informationen zur Konfiguration und Verwaltung von IBM® SPSS® Modeler Server.

- **IBM SPSS Modeler Administration Console – Benutzerhandbuch.** Informationen zur Installation und Nutzung der Konsolen-Benutzeroberfläche zur Überwachung und Konfiguration von SPSS Modeler Server. Die Konsole ist als Plugin für die Deployment Manager-Anwendung implementiert.
- **IBM SPSS Modeler Solution Publisher-Handbuch.** SPSS Modeler Solution Publisher ist eine Zusatzkomponente, mit der Unternehmen Streams zur Verwendung außerhalb der SPSS Modeler-Standardumgebung veröffentlichen können.
- **IBM SPSS Modeler-Handbuch zu CRISP-DM.** Schritt-für-Schritt-Anleitung für das Data-Mining mit SPSS Modeler unter Verwendung der CRISP-DM-Methode.

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 Daten-Sets sind viel kleiner als die riesigen Datenbestände, die von einigen Data-Mining-Experten verwaltet werden müssen, die zugrunde liegenden Konzepte und Methoden sollten sich jedoch auch auf reale Anwendungen übertragen lassen.

Sie können auf die Beispiele zugreifen, indem Sie im Menü “Hilfe” in SPSS Modeler auf die Option Anwendungsbeispiele klicken. Die Datendateien und Beispiel-Streams wurden im Ordner *Demos*, einem Unterordner des Produktinstallationsverzeichnis, installiert. [Für weitere Informationen siehe Thema Ordner “Demos” in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)

Beispiele für die Datenbank-Modellierung. Die Beispiele finden Sie im *IBM SPSS Modeler In-Database Mining-Handbuch*.

Skriptbeispiele. Die Beispiele finden Sie im *IBM SPSS Modeler Handbuch für die Skripterstellung und Automatisierung*.

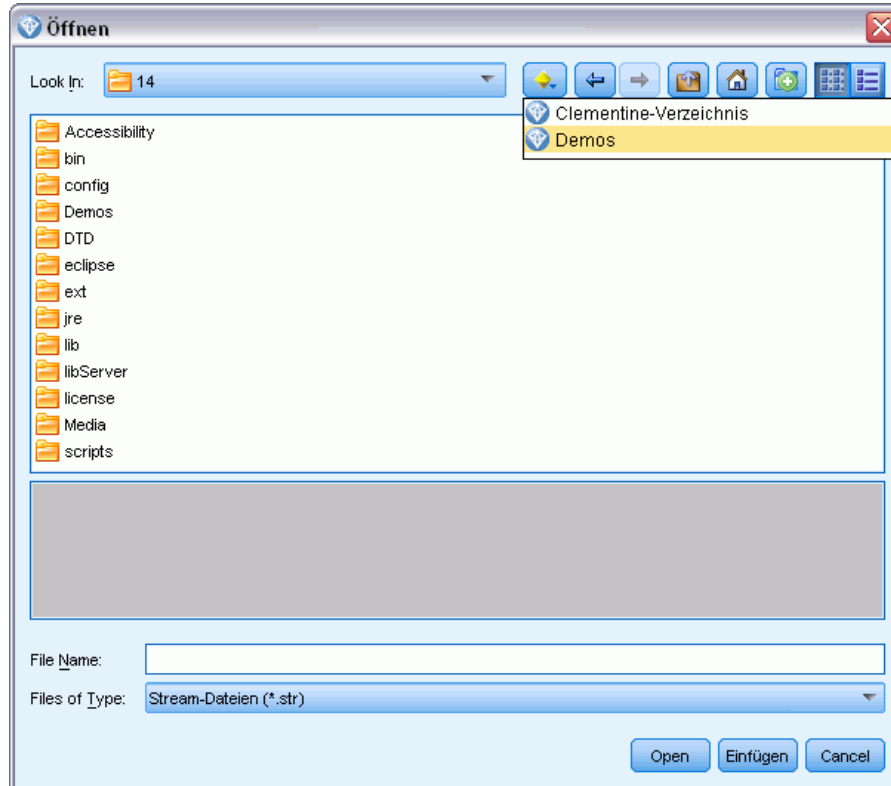
Ordner “Demos”

Die in den Anwendungsbeispielen verwendeten Datendateien und Beispiel-Streams wurden im Ordner *Demos*, einem Unterordner des Produktinstallationsverzeichnis, installiert. Auf diesen Ordner können Sie auch über die Programmgruppe IBM SPSS Modeler 14.2 im

Windows-Startmenü oder durch Klicken auf *Demos* in der Liste der zuletzt angezeigten Verzeichnisse im Dialogfeld "Datei öffnen" zugreifen.

Abbildung 1-1

Auswahl des Ordners "Demos" in der Liste der zuletzt angezeigten Verzeichnisse



Architektur- und Hardwareempfehlungen

IBM SPSS Modeler-Architektur

In diesem Abschnitt wird die Architektur von IBM® SPSS® Modeler Server beschrieben, inklusive der Serversoftware, der Client-Software und der Datenbank. Es umfasst Informationen zum Design von 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 nutzt eine verteilte 3-Tier-Architektur. Softwareoperationen werden zwischen den Client- und Servercomputern aufgeteilt. Die Vorteile von SPSS Modeler Server gegenüber der Standalone-Version IBM® SPSS® Modeler, vor allem beim Verarbeiten großer Daten-Sets, sind zahlreich:

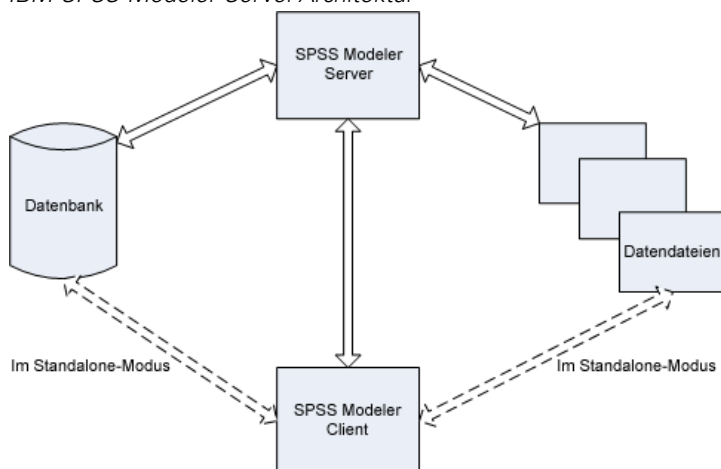
- SPSS Modeler Server kann neben Windows auch unter UNIX ausgeführt werden, sodass Sie eine größere Flexibilität hinsichtlich des Installationsortes besitzen. Sie können auf jeder Plattform einen schnelleren, größeren Servercomputer für Data Mining-Prozesse verwenden.
- SPSS Modeler Server wurde vorwiegend für Schnelligkeit optimiert. Wenn Vorgänge nicht an die Datenbank übertragen werden können, speichert SPSS Modeler Server die Zwischenergebnisse als temporäre Dateien auf der Festplatte und nicht im Arbeitsspeicher. Da Server normalerweise über ausreichend Festplattenspeicher verfügen, kann SPSS Modeler Server Sortier-, Zusammenführungs- und Aggregationsvorgänge mit sehr großen Daten-Sets 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® Collaboration and Deployment Services Deployment Manager (in SPSS Modeler Server enthalten) sowie IBM® SPSS® Collaboration and Deployment Services und IBM® SPSS® Collaboration and Deployment Services Repository (separat erhältlich) können Sie Data-Mining-Prozesse überwachen und sicherstellen, dass ausreichend Computerressourcen verfügbar sind. Sie können bestimmte Data-Mining-Aufgaben automatisieren, den Zugriff auf Datenmodelle verwalten und Ergebnisse unternehmensweit verfügbar machen.

Die Komponenten der verteilten Architektur von IBM® SPSS® Modeler sehen Sie in der Grafik "IBM SPSS Modeler Server-Architektur" auf S. 7.

- **SPSS Modeler.** Die Client-Software wird auf dem Computer des Endbenutzers installiert. Sie stellt die Benutzeroberfläche zur Verfügung und zeigt Data-Mining-Ergebnisse an. Der Client ist eine vollständige Installation der SPSS Modeler-Software. Wird der Client jedoch für die verteilte Analyse mit SPSS Modeler Server verbunden, wird seine Ausführungs-Engine deaktiviert. SPSS Modeler kann nur unter Windows ausgeführt werden.
- **SPSS Modeler Server.** Die Serversoftware wird auf einem Servercomputer mit einer Netzwerkverbindung zu dem/den SPSS Modeler(s) und der Datenbank installiert. SPSS Modeler Server wird als Dienst (unter Windows) oder als Daemonprozess (unter UNIX) ausgeführt und wartet auf die Verbindung von Clients. Er verarbeitet die Ausführung von Streams und Skripts, die mit SPSS Modeler erstellt wurden.
- **Datenbankserver.** Der Datenbankserver könnte als Live Data Warehouse (z. B. Oracle auf einem großen UNIX-Server) genutzt werden oder als Data Mart auf einem lokalen Server/Abteilungsserver (z. B. SQL Server unter Windows), um die Auswirkungen auf andere Betriebssysteme zu verringern.

IBM SPSS Modeler Server-Architektur

Abbildung 2-1
IBM SPSS Modeler Server-Architektur



Bei der verteilten Architektur findet ein Großteil der Verarbeitung auf dem Servercomputer statt. Wenn der Endbenutzer einen Stream ausführt, sendet SPSS Modeler eine Beschreibung des Streams an den Server, 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 SPSS Modeler Server alle 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 SPSS Modeler Server vollständig kompatibel mit Textdateien.

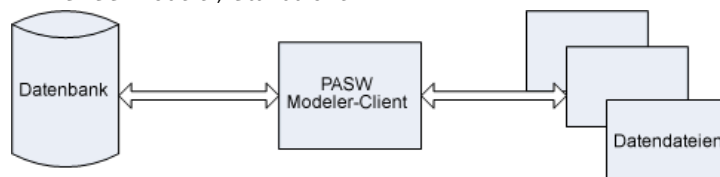
Außerdem steht Lastenausgleich durch die Verwendung von Server-Clustern für die Verarbeitung zur Verfügung. Die Clusterbildung ist ab IBM SPSS Collaboration and Deployment Services 3.5 verfügbar. Dazu wird das Plugin SPSS Coordinator of Processes verwendet. Für

weitere Informationen siehe Thema Lastenausgleich mit Server-Clustern in Anhang D auf S. 80. Die Verbindung mit einem im Coordinator of Processes verwalteten Server oder Cluster ist direkt über das SPSS Modeler-Dialogfeld "Anmelden beim Server" möglich. Für weitere Informationen siehe Thema Verbindung mit IBM SPSS Modeler Server in Kapitel 3 in *IBM SPSS Modeler 14.2- Benutzerhandbuch*.

Standalone-Client

SPSS Modeler kann auch als eigenständige Desktop-Anwendung konfiguriert werden, wie in der Grafik "IBM SPSS Modeler, Standalone" unten. Für weitere Informationen siehe Thema Support für IBM SPSS Modeler in Kapitel 3 auf S. 14.

Abbildung 2-2
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 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 SPSS Modeler Server auf einem dedizierten Servercomputer, dessen Ressourcen allein der Serversoftware zur Verfügung stehen. Dies betrifft auch alle Datenbanken, auf die 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 SPSS Modeler Server und der Datenbank auf demselben Computer die Übertragungszeit zwischen diesen beiden Komponenten aufgrund der verringerten Netzlast reduziert, 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 SPSS Modeler Server akzeptiert Verbindungen von mehreren Clients (Benutzern). Jede Client-Verbindung kann mehrere Stream-Ausführungen 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.

64-Bit-Plattformen. Wenn Sie Modelle mit großen Datenmengen erstellen oder verarbeiten möchten, verwenden Sie 64-Bit-Solaris, -Windows oder -Linux als Plattform für SPSS Modeler Server und verwenden Sie die maximal mögliche Menge an RAM für diesen Computer. Bei großen Daten-Sets 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. 64-Bit-Unterstützung ist für Solaris-, Windows- und Linux-Plattformen verfügbar.

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 SPSS Modeler Server-Versionen gerecht werden zu können.

Temporäre Festplattenspeicher- und Arbeitsspeicheranforderungen

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 leistungsstarke SQL-Optimierungsfunktion von IBM® SPSS® Modeler Server ermöglicht die Verarbeitung innerhalb der Datenbank (anstelle des Servers). Unter den folgenden Umständen kann die SQL-Optimierung jedoch nicht verwendet werden:

- Die zu verarbeitenden Daten befinden sich in einer Textdatei und nicht in einer Datenbank.
- Die SQL-Optimierung ist deaktiviert.
- 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.

- Aggregatknoten
- Duplikatknoten
- Klassierknoten
- Mergeknoten bei Verwendung der Option “Merge nach Schlüssel”
- Jeder Modellierungsknoten
- Sortierknoten
- Tabellenausgabeknoten
- @OFFSET-Funktionen, bei denen die Suchbedingung die Funktion @THIS verwendet.
- Sämtliche @ -Funktionen, 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 Daten-Set. Wenn die Datendatei beispielsweise 2 GB groß ist und die SQL-Generierung nicht verwendet wird, benötigt 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 physikalische Festplatten umfasst, idealerweise mit jeder Festplatte im Striped-Dateisystem auf einem separaten Festplattencontroller.

Arbeitsspeichieranforderungen

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 Neural Net-, Kohonen- und K-Means-Knoten 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:

$(\text{Anzahl_an_Datensätzen} * \text{Anzahl_an_Zellen_pro_Datensatz}) * \text{Anzahl_an_Byte_pro_Zelle}$.

Hierbei kann `number_of_cells_per_record` 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

Um in einer Datenbank zu lesen oder in ihr zu schreiben, muss eine ODBC-Datenquelle für die entsprechende Datenbank mit den erforderlichen Lese- und Schreibberechtigungen installiert und konfiguriert sein. Das IBM® SPSS® Data Access Pack umfasst eine Reihe von ODBC-Treibern, die zu diesem Zweck verwendet werden können. Diese Treiber stehen auf dem zu dieser Version gehörenden IBM SPSS Data Access Pack-Installationsmedium zur Verfügung. Wenn Sie Fragen zur Erstellung oder Einstellung von Berechtigungen für ODBC-Datenquellen haben, wenden Sie sich an Ihren Datenbankadministrator.

Die Datenbankunterstützung in IBM® SPSS® Modeler wird in drei Stufen eingeteilt, wobei jede Stufe je nach Datenbankanbieter für einen unterschiedlichen Unterstützungsgrad für SQL-Pushback und -Optimierung steht. Die unterschiedlichen Unterstützungsebenen werden durch eine Reihe von Systemeinstellungen implementiert, die als Teil einer Dienstleistungsabgabe angepasst werden können.

Die drei Stufen der Datenbankunterstützung sind:

Tabelle 2-1
Stufen der Datenbankunterstützung

Unterstützungsstufe	Beschreibung
Stufe 1	Vollständiger SQL-Pushback verfügbar, mit datenbankspezifischer SQL-Optimierung.
Stufe 2	Teilweiser SQL-Pushback verfügbar, mit datenbankspezifischer SQL-Optimierung.
Stufe 3	Kein SQL-Pushback oder -Optimierung, Daten können nur von der Datenbank gelesen oder in die Datenbank geschrieben werden.

Unterstützte ODBC-Treiber

Neueste Informationen zu Datenbanken und ODBC-Treibern, die für die Verwendung mit SPSS Modeler 14.2 getestet wurden und unterstützt werden, finden Sie in den Produktkompatibilitätsdiagrammen auf der unternehmenseigenen Support-Site unter <http://www.ibm.com/support>.

Installationsort der Treiber

Beachten Sie, dass die ODBC-Treiber auf jedem Computer installiert und konfiguriert werden müssen, auf dem eine Verarbeitung erfolgt.

- Wenn Sie IBM® SPSS® Modeler im lokalen (Standalone-) Modus ausführen, müssen die Treiber auf dem lokalen Computer installiert sein.
- Wenn Sie SPSS Modeler im verteilten Modus mit einem Remote-IBM® SPSS® Modeler Server ausführen, müssen die ODBC-Treiber auf dem Computer installiert sein, auf dem SPSS Modeler Server installiert ist.
- Wenn Sie von SPSS Modeler und SPSS Modeler Server auf die gleichen Datenquellen zugreifen müssen, müssen die ODBC-Treiber auf beiden Computern installiert sein.
- Wenn Sie SPSS Modeler über Terminaldienste ausführen, müssen die ODBC-Treiber auf dem Terminaldienste-Server installiert sein, auf dem Sie SPSS Modeler installiert haben.
- Wenn Sie IBM® SPSS® Modeler Solution Publisher Runtime verwenden, um veröffentlichte Streams auf einem separaten Computer auszuführen, müssen Sie die ODBC-Treiber auch auf diesem Computer installieren und konfigurieren.

Hinweis: Wenn Sie SPSS Modeler Server unter UNIX zum Zugriff auf eine Teradata-Datenbank verwenden, müssen Sie den ODBC-Treiber-Manager verwenden, der mit dem Teradata-ODBC-Treiber installiert wurde. Um diese Änderung an SPSS Modeler Server vorzunehmen, geben Sie für ODBC_DRIVER_MANAGER_PATH einen Wert in der Nähe des oberen Bereichs des Skripts modelersrv.sh ein, wo dies durch die Kommentare angegeben wurde. Diese Umgebungsvariable muss auf den Speicherort des ODBC-Treiber-Managers eingestellt werden, der mit dem Teradata ODBC-Treiber ausgeliefert wird (/usr/odbc/lib in einer Standardinstallation eines Teradata ODBC-Treibers). Sie müssen SPSS Modeler Server neu starten, damit die Änderung wirksam wird. Weitere Informationen zu den SPSS Modeler Server-Plattformen, die Teradata-Zugriff unterstützen, sowie über die unterstützte Teradata ODBC-Treiberversion finden Sie auf der unternehmenseigenen Support-Site unter <http://www.ibm.com/support>.

Hinweis: Die vorangehenden Regeln betreffen ausdrücklich den Zugriff auf Daten in einer Datenbank. Andere Dateizugriffe, wie das Öffnen und Speichern von Stream-, Projekt-, Modell-, Knoten-, PMMI-, Ausgabe- und Skriptdateien, 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* Client-Objekten (z. B. Streams) fest, er hat jedoch keinerlei Auswirkung auf das Arbeitsverzeichnis des Servers.

UNIX. Informationen zum Konfigurieren von SPSS Modeler Server unter UNIX, um mit der IBM® SPSS® Statistics Data Access-Technologie zu arbeiten, finden Sie unter [Konfigurieren von UNIX-Startskripts](#) auf S. 67.

Referenzieren von Datendateien

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

Sind die Daten auf einem anderen Host gespeichert, empfehlen wir die Verwendung von UNC-Dateiverweisen (z. B. `\\Mein Datenserver\ServerData\Umsatz 1998.csv`). Beachten Sie, dass UNC-Namen nur funktionieren, wenn der Pfad den Namen der gemeinsamen Netzwerkressource enthält. Der referenzierende Computer muss die Berechtigung zum Lesen der angegebenen Datei haben. Wenn Sie häufig vom verteilten in den lokalen Analysemodus 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 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 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 SPSS Statistics Client arbeitet auf die gleiche Weise. Zum Importieren und Exportieren zwischen den beiden Anwendungen müssen die beiden Clients im selben Modus arbeiten. Ist dies nicht der Fall, weichen die Ansichten der Dateisysteme voneinander ab und die gemeinsame Nutzung der Dateien ist nicht möglich. Die SPSS Statistics-Knoten von IBM® SPSS® Modeler können den SPSS Statistics Client automatisch starten, die Benutzer müssen jedoch zunächst sicherstellen, dass der SPSS Statistics Client und SPSS Modeler im selben Modus arbeiten.

Installationsanweisungen

Informationen zur Installation von IBM® SPSS® Modeler Server finden Sie unter `\documentation\installation\` auf der Server-CD. Weitere Dokumente sind für Windows und UNIX verfügbar.

Vollständige Informationen zum Installieren und Arbeiten mit IBM® SPSS® Modeler finden Sie auf der Client-CD.

Support für IBM SPSS Modeler

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

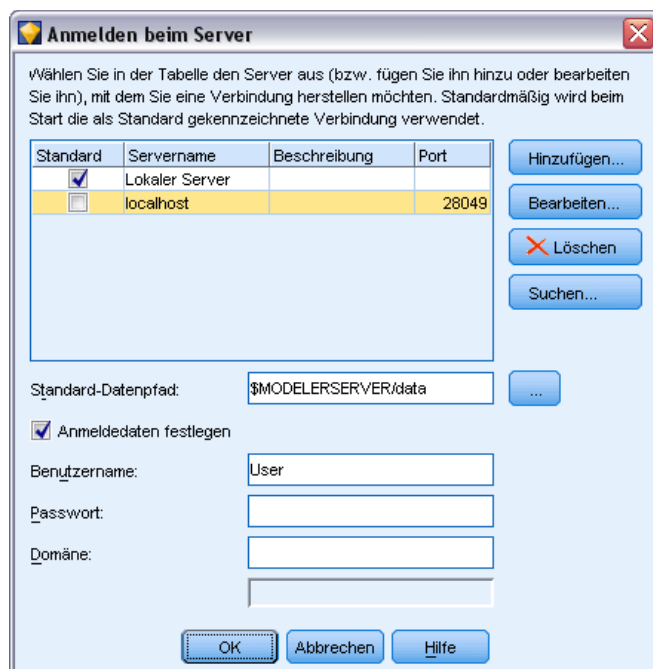
- Anmeldung bei IBM® SPSS® Modeler Server (oder Standalone-Ausführung durch Trennen der Verbindung zu einem Server)
- Vom Benutzer benötigte Daten- und Dateisysteme
- Benutzerkonten und Dateiberechtigungen für SPSS Modeler Server
- Unterschiede bei den Ergebnissen, die Benutzer beim Wechsel zwischen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

Verbindung 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 Plugin Coordinator of Processes von IBM® SPSS® Collaboration and Deployment Services mit einem SPSS Modeler Server oder Server-Cluster verbunden ist. Der aktuelle Verbindungsstatus wird unten links im 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 Netzwerk ausgeführten Statistics-Dienste zu durchsuchen, wird über den Coordinator of Processes bereitgestellt. [Für weitere Informationen siehe Thema Lastenausgleich mit Server-Clustern in Anhang D auf S. 80.](#)

Abbildung 3-1
Dialogfeld "Anmelden beim Server"



So stellen Sie eine Verbindung mit einem Server her:

- ▶ Klicken Sie im Menü "Extras" auf die Option Anmelden beim Server. Das Dialogfeld "Anmelden beim Server" wird geöffnet. Alternativ können Sie auf den Bereich des Verbindungsstatus im SPSS Modeler-Fenster doppelklicken.
- ▶ Legen Sie in diesem Dialogfeld die Optionen zum Verbinden mit dem lokalen Servercomputer fest oder wählen Sie eine Verbindung in der Tabelle aus.
 - Klicken Sie auf Hinzufügen bzw. Bearbeiten, um eine Verbindung hinzuzufügen bzw. zu bearbeiten. [Für weitere Informationen siehe Thema Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)
 - Klicken Sie auf Suche, um auf einen Server bzw. Server-Cluster in Coordinator of Processes zuzugreifen. [Für weitere Informationen siehe Thema Suchen nach Servern in IBM SPSS Collaboration and Deployment Services in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)

Servertabelle. Diese Tabelle enthält die Menge der definierten Serververbindungen. In der Tabelle werden Standardverbindung, Servername, 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".

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

Anmeldedaten einstellen. Belassen Sie dieses Kontrollkästchen nicht aktiviert, um die Funktion für die **Einzelanmeldung** zu aktivieren. Diese versucht, Sie mithilfe Ihres lokalen Benutzernamens und Passworts beim Server anzumelden. Falls die Einzelanmeldung nicht möglich ist oder Sie das Kontrollkästchen zur Deaktivierung der Einzelanmeldung aktivieren (z. B. zur Anmeldung in ein Administratorkonto), wird ein weiteres Fenster angezeigt, in dem Sie aufgefordert werden, Ihre Anmeldedaten einzugeben.

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

Paßwort. Geben Sie das Passwort 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.

- ▶ Klicken Sie auf OK, um die Verbindung herzustellen.

So trennen Sie eine Verbindung mit einem Server:

- ▶ Klicken Sie im Menü “Extras” auf die Option Anmelden beim Server. Das Dialogfeld “Anmelden beim Server” wird geöffnet. Alternativ können Sie auf den Bereich des Verbindungsstatus im SPSS Modeler-Fenster doppelklicken.
- ▶ Wählen Sie im Dialogfeld den lokalen Server aus und klicken Sie auf OK.

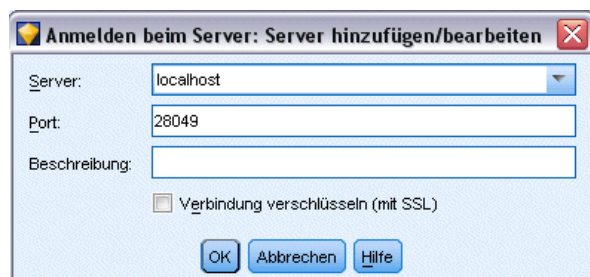
Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung

Serververbindungen können manuell im Kontrollkästchen “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.

Hinweis: Serververbindungen, die aus IBM® SPSS® Collaboration and Deployment Services hinzugefügt wurden, können nicht bearbeitet werden, da Name, Port und sonstige Details in IBM SPSS Collaboration and Deployment Services definiert sind.

Abbildung 3-2

Dialogfeld “Anmeldung beim Server: Server hinzufügen/bearbeiten”



So fügen Sie Serververbindungen hinzu:

- ▶ Klicken Sie im Menü “Extras” auf die Option Anmelden beim Server. Das Dialogfeld “Anmelden beim Server” wird geöffnet.
- ▶ Klicken Sie in diesem Dialogfeld auf Hinzufügen. Das Dialogfeld “Anmeldung beim Server: Server hinzufügen/bearbeiten” wird angezeigt.
- ▶ Geben Sie die Details für die Serververbindung ein und klicken Sie auf OK, um die Verbindung zu speichern und zum Dialogfeld “Anmeldung 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 Netzwerk 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:

- ▶ Klicken Sie im Menü “Extras” auf die Option Anmelden beim Server. Das Dialogfeld “Anmelden beim Server” wird geöffnet.
- ▶ Wählen Sie in diesem Dialogfeld die zu bearbeitende Verbindung aus und klicken Sie dann auf Bearbeiten. Das Dialogfeld “Anmeldung beim Server: Server hinzufügen/bearbeiten” wird angezeigt.
- ▶ Ändern Sie die Details für die Serververbindung und klicken Sie auf OK, um die Änderungen zu speichern und zum Dialogfeld “Anmeldung beim Server” zurückzukehren.

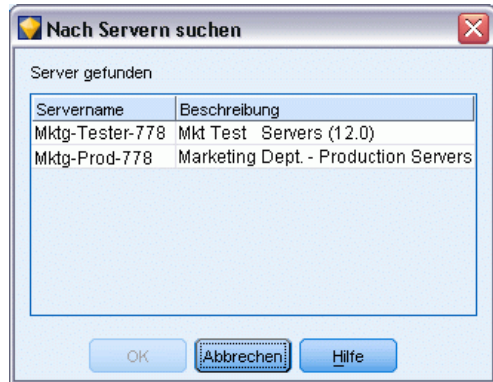
Suchen nach Servern in IBM SPSS Collaboration and Deployment Services

Anstatt eine Serververbindung manuell einzugeben, können Sie einen im Netzwerk 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. [Für weitere Informationen siehe Thema Lastenausgleich mit Server-Clustern in Anhang D auf S. 80.](#)

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 richtigen 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 Passwort.

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. [Für weitere Informationen siehe Thema Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)

Abbildung 3-3
Dialogfeld "Nach Servern suchen"



So suchen Sie nach Servern und Clustern:

- ▶ Klicken Sie im Menü "Extras" auf die Option Anmelden beim Server. Das Dialogfeld "Anmelden beim Server" wird geöffnet.
- ▶ 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. [Für weitere Informationen siehe Thema Verbindung mit dem IBM SPSS Collaboration and Deployment Services Repository in Kapitel 9 in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)
- ▶ Wählen Sie den Server bzw. Server-Cluster in der Liste aus.
- ▶ Klicken Sie auf OK, um das Dialogfeld zu schließen und diese Verbindung zu 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 Netzwerk zugreifen sowie Dateien auf dem Netzwerk speichern. Je nachdem benötigen sie die folgenden Informationen:

- **ODBC-Datenquelle.** Wenn Benutzer auf ODBC-Datenquellen zugreifen müssen, die auf dem Servercomputer definiert sind, benötigen Sie die Namen, Beschreibungen und Anmeldeinformationen (inklusive Datenanmelde-IDs und Passwörter) für die Datenquellen.

- **Datendateizugriff.** Wenn Benutzer auf dem Servercomputer oder an einem anderen Standort im Netzwerk auf Datendateien zugreifen wollen, müssen Sie die Namen und Speicherorte der Datendateien kennen.
- **Speicherort zum Speichern von Dateien.** Wenn Benutzer während einer Verbindung mit SPSS Modeler Server Daten speichern, kann es vorkommen, dass sie die Dateien auf dem Servercomputer 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 Servercomputers 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 Passwortberechtigungen für die Datenbank.

Windows. Unter Windows kann sich jeder Benutzer anmelden, der im Hostnetzwerk über ein gültiges Konto verfügt. Bei der Standardauthentifizierung benötigen Benutzer Zugriff einschließlich Änderungsrecht auf das Verzeichnis `<modeler_server_install>\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 Hostnetzwerk 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 "root"-Berechtigung ausgeführt werden kann. In diesem Fall müssen Sie eine private Passwortdatenbank erstellen, die für die Authentifizierung verwendet wird. Alle IBM® SPSS® Modeler-Benutzer müssen ein gemeinsames UNIX-Benutzerkonto verwenden (und somit gemeinsamen Zugriff auf Datendateien haben). [Für weitere Informationen siehe Thema Konfigurieren von IBM SPSS Modeler Server als Nicht-Root-Prozess in Anhang C auf S. 77.](#)

Unter Solaris, HP-UX, Linux und AIX-Plattformen verwendet SPSS Modeler Server PAM zur Authentifizierung. Sie können den Dienstnamen `modelerserver` zum Konfigurieren der PAM-Module für SPSS Modeler Server verwenden, falls erforderlich.

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. [Für weitere Informationen siehe Thema Temporäre Festplattenspeicher- und Arbeitsspeicheranforderungen in Kapitel 2 auf S. 9.](#)

UNIX. Die UNIX-Versionen von SPSS Modeler Server verwenden den UNIX-Befehl `umask`, um für die temporären Dateien Dateiberechtigungen festzulegen. Sie können die Standardberechtigungen des Servers überschreiben. [Für weitere Informationen siehe Thema Steuern der Berechtigungen für die Dateierstellung in Anhang B auf S. 69.](#)

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 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 SPSS Modeler Server. Aufgrund der Unterschiede beim Runden können die Ergebnisse leicht voneinander abweichen.

IBM SPSS Modeler Server-Verwaltung

Dieses Kapitel enthält Informationen zum Starten und Stoppen von IBM® SPSS® Modeler Server, zum Konfigurieren verschiedener Serveroptionen und zur Interpretation der Protokolldatei. Es beschreibt die Verwendung von IBM® SPSS® Collaboration and Deployment Services Deployment Manager, einer Anwendung zur Konfiguration und Überwachung des Servers. Installationsanweisungen zu dieser Komponente finden Sie in den Installationsanweisungen zu SPSS Modeler Server, die auf der Produkt-CD gespeichert sind.

Starten und Stoppen von IBM SPSS Modeler Server

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

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

So funktioniert das Starten, Stoppen bzw. die Statusprüfung unter Windows:

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

- ▶ **Windows XP.** Öffnen Sie das Windows-Startmenü. Wählen Sie Einstellungen und dann Systemsteuerung. Doppelklicken Sie auf Verwaltung und dann auf Dienste.

Windows 2003 oder 2008. Öffnen Sie das Windows-Startmenü. Wählen Sie Systemsteuerung, dann Verwaltung und dann Dienste.

- ▶ Wählen Sie den IBM SPSS Modeler Server <nn.n>-Service 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.

So funktioniert das Starten, Stoppen bzw. die Statusprüfung unter UNIX:

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

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

```
cd /usr/modelersrv
```

Dabei steht *modelersrv* für das Installationsverzeichnis von SPSS Modeler Server.

- ▶ Um den Server zu starten, geben Sie Folgendes an der Befehlszeilen-Eingabeaufforderung ein:
`./modelersrv.sh start`
- ▶ Um den Server zu stoppen, geben Sie Folgendes an der Befehlszeilen-Eingabeaufforderung ein:
`./modelersrv.sh stop`
- ▶ Um den Status von SPSS Modeler Server zu überprüfen, geben Sie an einer UNIX-Eingabeaufforderung Folgendes ein:
`./modelersrv.sh list`

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

Das SPSS Modeler Server-Installationsprogramm beinhaltet ein Skript (*auto.sh*), das das System so konfiguriert, dass der Server-Daemon automatisch beim Booten gestartet wird. Wenn Sie dieses Skript ausgeführt haben und dann den Server stoppen, wird der Daemon automatisch neu gestartet, wenn der Computer erneut gebootet wird. [Für weitere Informationen siehe Thema Automatisches Starten und Stoppen von IBM SPSS Modeler Server in Anhang B auf S. 68.](#)

UNIX-Kernel-Limits

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

- ▶ Fügen Sie *modelersrv.sh* die folgenden Befehle hinzu:
`ulimit -d unlimited`
`ulimit -m unlimited`
`ulimit -f unlimited`
- ▶ Starten Sie SPSS Modeler Server neu.

Umgang mit nicht reagierenden Server-Prozessen (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 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.

Auf Systemen, auf denen sich oft unbenutzbare (“Zombie-”) Server-Prozesse anhäufen, empfehlen wir, SPSS Modeler Server zu beenden und in regelmäßigen Abständen durch folgende Befehlssequenz neu zu starten:

```
cd modeler_server_install_directory
./modelersrv.sh stop
./modelersrv.sh kill
```

Diejenigen SPSS Modeler-Prozesse, die mithilfe des `modelersrv.sh kill`-Befehls 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 Benutzer-Interrupts, Systemabstürzen oder aus anderen Gründen. Wir empfehlen daher, neben dem regelmäßigen Neustart von SPSS Modeler Server, alle verbleibenden Dateien aus dem temporären Verzeichnis von SPSS Modeler zu entfernen.

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

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 Anwendung IBM® SPSS® Collaboration and Deployment Services Deployment Manager, die aktuellen IBM® SPSS® Modeler-Kunden kostenlos zur Verfügung steht. [Für weitere Informationen siehe Thema Verwendung von IBM SPSS Modeler Administration Console auf S. 24.](#)
- Verwenden Sie die Textdatei `options.cfg`, die sich im Verzeichnis `[Serverinstallationspfad]/config` befindet. [Für weitere Informationen siehe Thema Verwenden der Datei options.cfg auf S. 34.](#)

Wir empfehlen die Installation und Verwendung von Deployment Manager als Verwaltungstool und nicht die Bearbeitung der Datei `options.cfg`. Für die Bearbeitung der Datei ist Zugriff auf das Dateisystem von SPSS Modeler Server erforderlich, mit Deployment Manager dagegen können Sie jede Person mit einem Benutzerkonto zur Anpassung dieser Optionen berechtigen. Außerdem bietet Deployment Manager weitere Informationen zu den Serverprozessen, sodass Sie Nutzung und Leistung überwachen können. Außerdem können, anders als bei der Bearbeitung der Konfigurationsdatei, die meisten Konfigurationsoptionen ohne Neustart von SPSS Modeler Server geändert werden.

Weitere Informationen zur Verwendung von Deployment Manager und der Datei `options.cfg` finden Sie in folgenden Abschnitten.

Verwendung von IBM SPSS Modeler Administration Console

Die Anwendung Modeler Administration Console bietet eine Konsolen-Benutzeroberfläche 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.

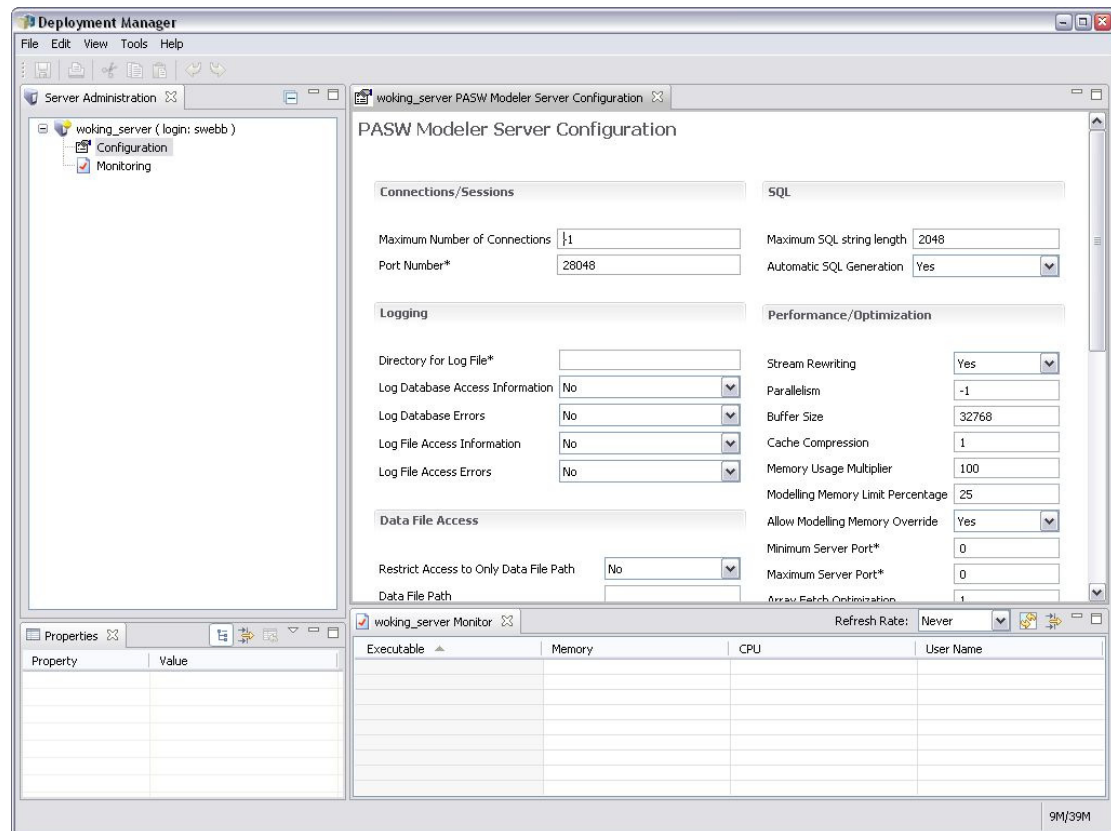
Viele der über Modeler Administration Console verfügbaren Optionen können auch in der Datei *options.cfg* angegeben werden, die sich im Installationsverzeichnis von SPSS Modeler Server unter */config* befindet. Die Modeler Administration Console-Anwendung bietet jedoch eine für die gemeinsame Verwendung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

Starten von Modeler Administration Console

- ▶ Zeigen Sie im Startmenü von Windows auf [Alle] Programme, IBM SPSS Collaboration and Deployment Services, Deployment Manager, IBM SPSS Modeler Administration Console <nn.n>.

Wenn Sie die Anwendung zum ersten Mal ausführen, sehen sie leere Server Administration- und Eigenschaften-Bereiche (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 Server Administrator-Bereich 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](#).

Abbildung 4-1
Modeler Administration Console-Benutzeroberfläche



Konfigurieren von Zugriffsrechten mit Modeler Administration Console

Der Administratorzugriff auf SPSS Modeler Server über Modeler Administration Console 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 auskommentiert, Sie müssen diese Zeile also bearbeiten, um den Zugriff für bestimmte Personen zuzulassen, oder Sie verwenden `*`, um allen Benutzern den Zugriff zu gewähren, wie in den folgenden Beispielen gezeigt wird:

```
administrators, "*"
administrators, "jschmidt,mjonas,abauer"
```

- 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 Domännennamen.
- Verwenden Sie das Sternchen (`*`) umsichtig. Mit dieser Einstellung kann sich jede Person mit einem gültigen Benutzerkonto für IBM® SPSS® Modeler Server (meistens jede beliebige Person im Netzwerk) anmelden und die Konfigurationsoptionen ändern.

SPSS Modeler Server-Verbindungen

Sie müssen für jeden SPSS Modeler Server in Ihrem Netzwerk, den Sie verwalten möchten, eine Verbindung einrichten. Anschließend müssen Sie sich bei jedem Server anmelden. Die Serververbindung wird für Modeler Administration Console-Sitzungen gespeichert, nicht jedoch die Anmeldeinformationen. Sie müssen sich bei jedem Start von Modeler Administration Console erneut anmelden.

So richten Sie eine Serververbindung ein

- ▶ Stellen Sie sicher, dass der IBM® SPSS® Modeler Server-Service gestartet wurde.
- ▶ Wählen Sie im Menü “Datei” die Option Neu und dann Verwaltete Serververbindung.
- ▶ Geben Sie in 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.
- ▶ 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 Fertig stellen. Die neue Serververbindung wird im Server Administrator-Bereich angezeigt.

Um Verwaltungsaufgaben durchzuführen, müssen Sie sich anmelden.

So melden Sie sich beim Server an

- ▶ Doppelklicken Sie im Server Administrator-Bereich auf den Server, bei dem Sie sich anmelden möchten.
- ▶ Geben Sie im Dialogfeld “Anmelden” Ihre Anmeldeinformationen ein. (Verwenden Sie Ihr Benutzerkonto für den Serverhost.) Klicken Sie auf OK.

Wenn die Anmeldung fehlschlägt und die Meldung Administratorrechte auf dem Server nicht verfügbar erscheint, ist die wahrscheinlichste Ursache, dass der Administratorzugriff nicht korrekt konfiguriert wurde. [Für weitere Informationen siehe Thema Konfigurieren von Zugriffsrechten mit Modeler Administration Console auf S. 25.](#)

Wenn die Anmeldung fehlschlägt und die Fehlermeldung Verbindung zum Server fehlgeschlagen ‘<server>’ angezeigt wird, überprüfen Sie, dass die Benutzer-ID und das Passwort korrekt sind und dass der SPSS Modeler Server-Service ausgeführt wird. Gehen Sie z. B. unter Windows auf Systemsteuerung > Verwaltung > Dienste und überprüfen Sie die Eingabe für IBM SPSS Modeler Server. Wenn die Statusspalte 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 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 eine Änderung der Portnummer erst nach dem Neustart des Servers wirksam wird.)

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.

Verbindungen/Sitzungen

Abbildung 4-2
Verbindungs- und Sitzungseinstellungen

Connections/Sessions	
Maximum Number of Connections	-1
Port Number*	28047

Maximale Anzahl an Verbindungen. (max_sessions) Maximale Anzahl an Benutzern, die gleichzeitig verbunden sein können. 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 (28047). Endbenutzer müssen die Portnummer kennen, um mit SPSS Modeler Server arbeiten zu können.

Protokollierung

Abbildung 4-3
Protokolleinstellungen

Logging	
Directory for Log File*	
Log Database Access Information	No
Log Database Errors	No
Log File Access Information	No
Log File Access Errors	No

Verzeichnis für Protokolldatei. (log_directory) Speicherort der Datei *messages.log*. Ein leerer Wert bedeutet *[serverinstallationspfad]/log*. Beachten Sie, dass normale Schrägstriche (/) zu verwenden sind. Informationen zum Inhalt dieser Protokolldatei finden Sie unter [Server-Protokoll auf S. 43](#).

Datenbankzugriffsinformationen protokollieren. (log_database_access) Der Protokolldatei werden Zugriffsinformationen für die Datenbank hinzugefügt.

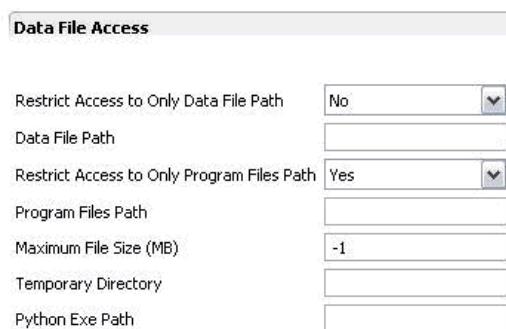
Datenbankfehler protokollieren. (log_database_errors) Der Protokolldatei werden Datenbankfehler hinzugefügt.

Zugriffsinformationen für Dateien aufzeichnen. (log_file_access) Der Protokolldatei werden Zugriffsinformationen für Dateien hinzugefügt.

Dateizugriffsfehler protokollieren. (log_file_errors) Der Protokolldatei werden Zugriffsinformationen für Dateien hinzugefügt.

Datendateizugriff

Abbildung 4-4
Datendateizugriff-Einstellungen



Data File Access	
Restrict Access to Only Data File Path	No
Data File Path	
Restrict Access to Only Program Files Path	Yes
Program Files Path	
Maximum File Size (MB)	-1
Temporary Directory	
Python Exe Path	

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 mehrere Verzeichnisse durch Strichpunkte (z. B. *[serverinstallationspfad]/data;c:/data;c:/temp*).

Zugriff ausschließlich auf Programmdateipfad beschränken. (program_files_restricted) Ist dieser Wert auf *Ja* eingestellt, wird der Zugriff auf Programmdateien im Standardverzeichnis *bin* und den im Programmdateipfad aufgelisteten Dateien eingeschränkt.

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 Strichpunkte.

Maximale Dateigröße (MB). (max_file_size) Maximale Größe von temporären und exportierten Datendateien, die während der Stream-Ausführung 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 (Cache-Dateien) 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 Data Preparation-Vorgänge (beispielsweise die Sortierung) bei der Ausführung Parallelität verwenden. Dadurch, dass die einzelnen Ausführungs-Threads gesonderte Datenträger für die temporäre Speicherung verwenden können, ist eine Leistungssteigerung möglich. Verwenden Sie bei allen Pfadangaben normale Schrägstriche (/). *Hinweis:* Beim Starten von SPSS Modeler Server werden in diesem Verzeichnis temporäre Dateien erzeugt. Stellen Sie sicher, dass Sie über die notwendigen Zugriffsrechte für dieses Verzeichnis verfügen (beispielsweise wenn es sich beim temporären Verzeichnis um einen freigegebenen Netzwerkordner handelt), anderenfalls kann SPSS Modeler Server nicht gestartet werden.

Pfad zur ausführbaren Python-Datei. (python_exe_path) Vollständiger Pfad zur ausführbaren Python-Datei einschließlich Name der exe-Datei. Je nachdem, wo Python installiert ist, kann es erforderlich sein, dass für Zugriff ausschließlich auf Programmdateipfad beschränken die Option *Nein* angegeben sein muss.

Leistung und Optimierung

Abbildung 4-5
Leistungs- und Optimierungseinstellungen

Performance/Optimization	
Stream Rewriting	Yes
Parallelism	-1
Buffer Size	32768
Cache Compression	1
Memory Usage Multiplier	100
Modelling Memory Limit Percentage	25
Allow Modelling Memory Override	Yes
Minimum Server Port	0
Maximum Server Port	0
Array Fetch Optimization	1

Stream-Neuerstellung. (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 Daten-Sets so früh wie möglich zu verringern. Die Deaktivierung 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.

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 deaktivieren, stellen Sie diesen Wert auf 1 ein. Um eine eingeschränkte parallele Verarbeitung zu gestatten,

setzen Sie den Wert auf eine Zahl, die geringer ist als die auf Ihrem Computer verfügbare Anzahl an Prozessoren. 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.

Cache-Komprimierung. (`cache_compression`) Ein ganzzahliger Wert von 0 bis 9, der die Komprimierung von Cache und anderen Dateien im temporären Verzeichnis des Servers steuert. Durch Komprimierung reduziert sich die verwendete Speichermenge, was bei begrenzter Speicherkapazität von Bedeutung ist. Außerdem wird die Leistungsfähigkeit erhöht, da weniger Datenträgeraktivität zum Erstellen und Lesen des Caches erforderlich 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 Netzwerktrainingsalgorithmen verwendet werden. Mit dem Wert 0 wird die Komprimierung vollständig deaktiviert. Die Werte von 1 aufwärts bieten einen steigenden Komprimierungsgrad, jedoch jeweils mit entsprechend erhöhter Zugriffszeit. Der Standardwert ist 1; höhere Werte sind selten erforderlich, es sei denn, der Datenträgerspeicher ist besonders knapp.

Multiplikator für Arbeitsspeicherverwendung. (`memory_usage`) Legt fest, welcher Anteil an physikalischem Speicher für die Sortierung und andere arbeitsspeicherinterne Cache-Vorgänge zugewiesen wird. Die Standardeinstellung ist 100. Dies entspricht etwa 10 % des physikalischen Speichers. Wenn freier Arbeitsspeicher verfügbar 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`) Legt fest, welcher Anteil an physikalischem Speicher für das Training von Netzwerk-, Kohonen- und *K*-Means-Modellen zugewiesen wird. Die Standardeinstellung liegt bei 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. deaktiviert die Option Auf Geschwindigkeit optimieren 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 deaktiviert 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 Stream-Ausführung 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: "keine Einschränkung".

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-Serverport 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 Stream-Ausführung.

Standardmäßig verwendet 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 Fehlermeldung angezeigt. Zur Konfiguration des Portbereichs benötigt SPSS Modeler einen verfügbaren offenen Port (zusätzliche zum Hauptport des Servers) pro gleichzeitigem Stream plus einen zusätzlichen Port für jeden angeschlossenen Client, der über eine ODBC-Verbindung verfügt. Um mehr Streams gleichzeitig auszuführen, können Sie den Bereich der Portnummern erweitern und entsprechend mehr Ports in der Firewall öffnen.

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

Array-Abufr Optimierung. (`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

Abbildung 4-6
SQL-Einstellungen

The image shows a dialog box titled "SQL". It contains two configuration options:

- "Maximum SQL string length" is set to "2048" in a text input field.
- "Automatic SQL Generation" is set to "Yes" in a dropdown menu.

Maximale SQL-Zeichenkettenlänge. (`max_sql_string_length`) Maximale Länge für Zeichenketten, die aus einer Datenbank mit SQL importiert werden. Zeichenkettenwerte, die länger sind als dieser Wert, werden ohne Warnmeldung rechts gekürzt. Der zulässige Bereich liegt zwischen 1 und 65.535 Zeichen.

Automatische SQL-Erzeugung. (`sql_generation_enabled`) Ermöglicht automatische SQL-Erzeugung für Streams, wodurch die Leistung erheblich verbessert werden kann. Diese Option ist standardmäßig aktiviert. Die Deaktivierung 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 für Scoring-Zwecke die SQL-Erzeugung unabhängig von dieser Einstellung für jeden Modellierungsknoten gesondert aktiviert werden muss.

SSL

Abbildung 4-7
SSL-Einstellungen

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

Zertifikatdatei. (`ssl_certificate_file`) Die Datei mit dem öffentlichen SSL-Zertifikat, das beim Start des Servers geladen werden soll.

Datei für privaten Schlüssel. (`ssl_private_key_file`) Die Datei mit dem privaten SSL-Schlüssel, der beim Start des Servers geladen werden soll.

Passwort für privaten Schlüssel. (`ssl_private_key_password`) Passwortschlüssel für das spezifizierte Zertifikat. Dieser Schlüssel muss von mindestens einer der in der konfigurierten Zertifikatdatei aufgeführten Zertifizierungsstellen signiert worden sein.

Konfiguration von Coordinator of Processes

Abbildung 4-8
Konfigurationseinstellungen für Coordinator of Processes

Items marked (*) require a server restart to take effect.

Host. (`cop_host`) Der Hostname bzw. die IP-Adresse des Coordinator of Processes-Diensts. Der Standardname “spsscop” 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-Diensts. Der Standardport, 8080, ist der IBM SPSS Collaboration and Deployment Services-Standardport.

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

Passwort. (cop_password) Das verschlüsselte Passwort für die Authentifizierung beim Coordinator of Processes-Dienst.

Zum Erstellen der Passwortdatei benötigen Sie das Dienstprogramm für Passwörter (*pwutil*), das sich im Verzeichnis *bin* der SPSS Modeler Server-Installation befindet.

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

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-Meldungen; der Standardwert ist 2.

Gewicht. (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.

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

Optionen nur in options.cfg sichtbar

default_sql_string_length. Gibt die Standardbreite der Zeichenkettenspalten an, die in den Datenbank-Cache-Tabellen erstellt werden. Zeichenkettenfelder in Datenbank-Cache-Tabellen werden mit einer Standardbreite von 255 erstellt, wenn keine Informationen von einem weiter oben liegenden Typ zur Verfügung stehen. Wenn Sie in Ihren Daten breitere Werte haben, instanziiieren Sie entweder einen weiter oben liegenden Typknoten mit diesen Werten, oder setzen Sie diesen Parameter auf einen Wert, der groß für diese Zeichenkettenwerte ist.

max_transfer_size. Nur für systeminterne Verwendung, nicht bearbeiten.

start_process_as_login_user. Legen Sie dies auf `true` fest, wenn Sie eine Einzelanmeldung verwenden und mit einem IBM® SPSS® Collaboration and Deployment Services-Server verbunden sind, der nicht die gleiche Authentifizierungsmethode wie der IBM® SPSS® Modeler Server-Host verwendet. Durch Festlegen der Option auf `true` können die untergeordneten Prozesse die Anmeldedaten von SPSS Modeler Server verwenden.

administrators. Geben Sie die Benutzernamen der Benutzer an, den Sie Administratorzugriff gewähren möchten. [Für weitere Informationen siehe Thema Konfigurieren von Zugriffsrechten mit Modeler Administration Console auf S. 25.](#)

`supported_sql_type_logging`. Nur zu Debugging-Zwecken; nicht bearbeiten, außer Sie werden dazu von einem Mitarbeiter des technischen Supports angewiesen.

`allow_config_custom_overrides`. Nicht bearbeiten, außer Sie werden dazu von einem Mitarbeiter des technischen Supports angewiesen.

SPSS Modeler Server-Überwachung

Im Überwachungsbereich von Modeler Administration Console wird ein Snapshot aller aktuell auf dem SPSS Modeler Server-Computer laufenden Prozesse angezeigt, ähnlich wie im Windows Task-Manager. Sie öffnen den Überwachungsbereich, indem Sie im Server Administrator-Bereich auf den Überwachungsknoten unter dem gewünschten Server doppelklicken. Der Bereich wird dann mit einem aktuellen Snapshot 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.

Abbildung 4-9
Überwachungsbereich von IBM SPSS Modeler Administration Console

Executable	Memory	CPU	User Name
FrameworkService	13,036 K	00	SYSTEM
LocalSch	1,640 K	00	SYSTEM
Mcshield	23,364 K	00	SYSTEM
SoftMon	1,632 K	00	SYSTEM
System	176 K	00	SYSTEM
UpdaterUI	2,156 K	00	cstrahs
VsTaskMgr	464 K	00	SYSTEM
clemserv	10,452 K	00	SYSTEM
clemserv	22,312 K	00	SYSTEM

54 Processes 3M/4M

Verwenden der Datei `options.cfg`

Die Datei `options.cfg` befindet sich im Verzeichnis `[Serverinstallationspfad]/config`. Die einzelnen Einstellungen werden als durch Kommata getrenntes Namen-Werte-Paar angegeben, wobei der **Name** den Namen der Option bezeichnet und der **Wert** den Wert für die Option. Rauten (#) zeigen Kommentare an.

Hinweis: Mit Ausnahme der folgenden Optionen können alle Konfigurationsoptionen anstatt mit dieser Konfigurationsdatei mit IBM® SPSS® Collaboration and Deployment Services Deployment Manager geändert werden.

- `administrators`
- `allow_custom_config_overrides`
- `max_transfer_size`
- `start_process_as_login_user`
- `supported_sql_type_logging`

Durch die Verwendung von Deployment Manager können Sie Serverneustarts für alle Optionen mit Ausnahme des Serverports vermeiden. [Für weitere Informationen siehe Thema Verwendung von IBM SPSS Modeler Administration Console auf S. 24.](#)

Konfigurationsoptionen, die zur Standarddatei hinzugefügt werden können

In-Database-Caching wird standardmäßig mit IBM® SPSS® Modeler Server aktiviert. Sie können diese Funktion deaktivieren, indem Sie folgende Zeile in die Datei *options.cfg* einfügen:

```
enable_database_caching, N
```

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

So können Sie die Konfigurationsoptionen für IBM SPSS Modeler Server anzeigen bzw. ändern:

- ▶ Öffnen Sie die Datei *options.cfg* mit einem Texteditor.
- ▶ Suchen Sie die relevanten Optionen. Eine vollständige Optionsliste finden Sie unter [SPSS Modeler Server-Konfiguration auf S. 27.](#)
- ▶ Bearbeiten Sie die Werte nach Bedarf. Beachten Sie, dass bei allen Pfadnamenwerten ein normaler Schrägstrich (/) und kein umgekehrter Schrägstrich (\) als Trennzeichen für den Pfadnamen verwendet werden muss.
- ▶ Speichern Sie die Datei.
- ▶ Stoppen Sie SPSS Modeler Server und starten Sie das Programm erneut, um die Änderungen zu übernehmen. [Für weitere Informationen siehe Thema Starten und Stoppen von IBM SPSS Modeler Server auf S. 21.](#)

Nicht verwendete Datenbankverbindungen schließen

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

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

So schließen Sie nicht verwendete Datenbankverbindungen:

- ▶ Suchen Sie das Verzeichnis *[Serverinstallationspfad]/config* .
- ▶ Fügen Sie folgende Datei hinzu (oder öffnen Sie sie, wenn sie bereits existiert):
odbc-custom-properties.cfg
- ▶ Fügen Sie folgende Zeile in die Datei ein:
`cache_connection, N`
- ▶ Speichern und schließen Sie die Datei.
- ▶ Starten Sie IBM® SPSS® Modeler Server neu, um die Änderungen zu übernehmen.

IBM SPSS Statistics-Lizenzspeicherort

Damit IBM® SPSS® Modeler den Statistics-Transformations-, Statistics-Modell- und den Statistics-Ausgabeknoten verwenden kann, muss auf dem Computer, auf dem der Stream ausgeführt wird, eine Kopie von IBM® SPSS® Statistics installiert und lizenziert sein. Bei der Ausführung mithilfe eines Servers muss die Lizenzkonfiguration auf dem betreffenden Server erfolgen.

Windows

- ▶ Wechseln Sie an der Eingabeaufforderung in das Verzeichnis “bin” von IBM® SPSS® Modeler Server und führen Sie folgenden Befehl aus:

```
statisticsutility -location =<Pfad zur SPSS Server-Lizenzdatei>IBM SPSS Statistics
```

Dabei ist <path to SPSS Statistics Server license file> das Installationsverzeichnis eines lizenzierten SPSS Statistics-Servers.

UNIX

- ▶ Wechseln Sie an der Eingabeaufforderung in das Verzeichnis “bin” von SPSS Modeler Server und führen Sie folgenden Befehl aus:

```
statisticsutility -location =<Pfad zur IBM SPSS Statistics Server-Lizenzdatei>
```

Dabei ist <path to SPSS Statistics Server license file> das Installationsverzeichnis eines lizenzierten SPSS Statistics-Servers.

Verwenden von SSL zur sicheren Datenübertragung

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

Wie SSL funktioniert

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.

- ▶ Wenn ein Client eine Verbindung zu einem Server aufbaut, authentifiziert der Client den Server mit dem Zertifikat für den öffentlichen Schlüssel.
- ▶ Der Client generiert dann eine Zufallszahl, verschlüsselt die Zahl mit dem öffentlichen Schlüssel des Servers und sendet die verschlüsselte Meldung zurück an den Server.
- ▶ Der Server entschlüsselt die Zufallszahl mit seinem privaten Schlüssel.
- ▶ 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 selbstsignierte Zertifikate generiert werden.

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

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

- ▶ Installieren Sie OpenSSL auf dem Server-Computer.
- ▶ Beziehen und installieren Sie das SSL-Zertifikat und die Schlüssel.
- ▶ Aktivieren und konfigurieren Sie SSL in der Server-Verwaltungsanwendung (IBM® SPSS® Collaboration and Deployment Services Deployment Manager).
- ▶ Installieren Sie nach Wunsch Verschlüsselung mit unbegrenzter Stärke auf den Client-Computern.
- ▶ Wenn Sie ein selbstsigniertes Zertifikat verwenden, kopieren Sie es auf den Client-Computer.
- ▶ Fügen Sie das Zertifikat zum Client-Keystore hinzu.
- ▶ Weisen Sie die Endbenutzer an, bei der Verbindung zum Server SSL zu aktivieren.

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

OpenSSL installieren

Wenn OpenSSL noch nicht auf dem Server verfügbar ist, müssen Sie es installieren.

- Laden Sie OpenSSL von <http://www.openssl.org/> herunter. Verwenden Sie unbedingt die geeignete OpenSSL-Version für die Server-Version:

Tabelle 4-1
OpenSSL-Versionen

Server-Produkt	Kompatible OpenSSL-Version
IBM® SPSS® Statistics 17-19 (nicht Linux® auf System z®) IBM® SPSS® Modeler 14.2 (nicht Linux® auf System z®)	0.9.8 und Unterversionen (0.9.8a, 0.9.8b usw.)
SPSS Statistics 19 (Linux® auf System z®) SPSS Modeler 14.2 (Linux® auf System z®)	1.0.0

Anmerkung: Wenn IBM® SPSS® Modeler Server für UNIX für SSL konfiguriert ist, versucht das Programm eine Verbindung mit zwei Bibliotheken mit Bezeichnungen wie den folgenden herzustellen (.so kann von der Plattform abhängen):

- *libssl.so.0.9.8*
- *libcrypto.so.0.9.8*

Bei der Verwendung einer Unterversion von SSL (0.9.8a, 0.9.8b usw.) müssen Sie für diese Bibliotheken symbolische Verbindungen mit den Versionenxxx.0.9.8 erstellen, nach denen SPSS Modeler Server sucht.

- Installieren und konfigurieren Sie die Software gemäß den Anweisungen. Wenn Sie OpenSSL selbst erstellen, beachten Sie die folgenden Richtlinien:

Windows. OpenSSL muss mit DLLs (die standardmäßig mehrere Threads benutzen) erstellt werden.

UNIX. OpenSSL muss mehrere Threads (was nicht immer standardmäßig der Fall ist) und gemeinsam verwendete Bibliotheken unterstützen.

- Stellen Sie sicher, dass das System die SSL-Bibliotheken finden kann:

Windows. Stellen Sie sicher, dass sich die OpenSSL-Module im Systempfad befinden.

UNIX. Hängen Sie die SSL-Bibliotheken an das Ende des in *modelersrv.sh* definierten Pfads zum Laden der Bibliothek an. Fügen Sie unter Solaris beispielsweise die folgende Zeile zum Server-Startskript hinzu:

```
LD_LIBRARY_PATH_64=$LD_LIBRARY_PATH_64:/usr/local/openssl; export LD_LIBRARY_PATH_64
```

Beziehen und Installieren von SSL-Zertifikat und -Schlüssel

- Beziehen Sie ein SSL-Zertifikat und eine Schlüsseldatei. Dazu haben Sie zwei Möglichkeiten.

- Erwerben Sie sie von einer öffentlichen Zertifizierungsstelle (z. B. Verisign oder Thawte). Die öffentliche Zertifizierungsstelle signiert das Zertifikat, um den Server zu verifizieren, der es benutzt.
 - Generieren Sie Schlüssel und Zertifikatdateien mithilfe einer internen Zertifizierungsstelle für selbstsignierte Zertifikate. OpenSSL bietet ein Zertifikatverwaltungstool für diesen Zweck. Alternativ können Sie im Internet nach Anweisungen zur Erstellung von selbstsignierten SSL-Zertifikaten suchen.
- ▶ Kopieren Sie Zertifikat und Schlüsseldateien in ein lokales Verzeichnis auf dem Server. Der öffentliche und der private Schlüssel können zusammen in einer Datei gespeichert werden. Sorgen Sie dafür, dass Zugriffsberechtigungen ein zufälliges Anzeigen dieses Verzeichnisses verhindern.
 - ▶ Installieren Sie SSL-Zertifikat und Schlüssel auf dem Anwendungsserver. Weitere Informationen dazu, wie die Schlüssel und das Zertifikat mit einem bestimmten Anwendungsserver zusammenarbeiten, finden Sie in der entsprechenden Herstellerdokumentation. Beachten Sie, dass Sie das Zertifikat und die Schlüssel möglicherweise dem Java Keystore hinzufügen müssen.

Aktivieren und Konfigurieren von SSL in der IBM SPSS Collaboration and Deployment Services Deployment Manager

- ▶ Starten Sie die Server-Verwaltungsanwendung (IBM® SPSS® Collaboration and Deployment Services Deployment Manager) und bauen Sie eine Verbindung zum Server auf.
- ▶ Stellen Sie auf der Konfigurationsseite die Option Secure Sockets Layer auf Ja ein .
- ▶ Geben Sie in SSL-Datei - öffentlicher Schlüssel den vollständigen Pfad zur Datei mit dem öffentlichen Schlüssel an.
- ▶ 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.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:
Datei > Speichern
- ▶ Starten Sie Server-Dienst oder -Daemon neu. Beim Neustart werden Sie zur Eingabe des SSL-Passworts aufgefordert. Unter Windows können Sie Passwort speichern wählen, um das Passwort sicher zu speichern. Damit ist es nicht mehr nötig, bei jedem Serverstart das Passwort einzugeben.

Einrichten von SSL in IBM SPSS Modeler Server (nur UNIX-Systeme)

Wenn zum Zugriff auf die private SSL-Schlüsseldatei ein Passwort erforderlich ist, können Sie an UNIX-Systemen das Passwort auf eine der folgenden Arten angeben:

- Geben Sie es in der Datei *options.cfg* mithilfe der Option `ssl_private_key_password` an. [Für weitere Informationen siehe Thema SSL auf S. 32.](#)
- Fügen Sie der Datei *modelersrv.sh* einen Schalter hinzu, um das Passwort beim Start anzufordern.

Wenn Sie die zweite Option verwenden, suchen Sie in der Datei *modelersrv.sh* die folgende Zeile:

```
if "$INSTALLEDPATH/$SCLEMDNAME" -server $ARGS; then
```

Fügen Sie den Schalter `-request_ssl_password` wie folgt hinzu:

```
if "$INSTALLEDPATH/$SCLEMDNAME" -request_ssl_password -server $ARGS; then
```

Installieren von 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 empfehlen wir, die Verschlüsselungsstufe auf unbegrenzte Stärke zu erhöhen.

- ▶ Laden Sie dazu die “Java Cryptography Extension (JCE) Unlimited Strength Jurisdiction Policy Files 5.0” von <http://java.sun.com/javase/downloads/index.jsp> herunter.
- ▶ Entpacken Sie die heruntergeladene Datei.
- ▶ Kopieren Sie die beiden *.jar*-Dateien *local_policy.jar* und *US_export_policy.jar* in das Verzeichnis `<Installationsordner>/jre/lib/security`, wobei `<Installationsordner>` der Ordner ist, in dem Sie das Produkt installiert haben.

Hinzufügen des Zertifikats zum Client-Keystore (für Verbindungen zum Repository oder IBM SPSS Modeler Server)

Hinweis: Überspringen Sie diesen Schritt, wenn Sie ein Zertifikat verwenden, das von einer Zertifizierungsstelle signiert wurde.

Wenn Sie SSL für die Verbindung zu einem IBM® SPSS® Collaboration and Deployment Services-Repository oder IBM® SPSS® Modeler Server verwenden und außerdem selbstsignierte Zertifikate verwenden, müssen Sie das Zertifikat dem Java Keystore des Clients hinzufügen. Die folgenden Schritte werden am *Client*-Computer ausgeführt.

Beachten Sie, dass auch ein Server-Computer als Client fungieren kann. Ein Beispiel ist ein IBM® SPSS® Statistics-Server, der eine Verbindung zum IBM® SPSS® Collaboration and Deployment Services Repository aufbaut. In diesem Fall ist SPSS Statistics-Server der *client*. Daher müssen Sie das Zertifikat für den IBM SPSS Collaboration and Deployment Services Repository-Server auf den SPSS Statistics-Server kopieren.

- ▶ Kopieren Sie die Zertifikatdatei (.cer) vom SSL-Server in das Verzeichnis `<install_dir>\jre\lib\security` auf dem Client-Computer.
- ▶ Klicken Sie im selben Verzeichnis mit der rechten Maustaste auf die Datei `cacerts` und wählen Sie die Option `Eigenschaften`.
- ▶ Deaktivieren Sie das Kontrollkästchen `Schreibgeschützt`.
- ▶ Öffnen Sie eine Befehlszeilen-Eingabeaufforderung und wechseln Sie in folgendes Verzeichnis, wobei `<Produkt-Installationsverzeichnis>` das Verzeichnis ist, in dem Sie das Produkt installiert haben:

```
<Produkt-Installationsverzeichnis>/jre/bin
```

- ▶ Geben Sie folgenden Befehl ein:

```
keytool -import -alias <Aliasname> -file <Pfad zum Zertifikat> -keystore <Pfad zum Keystore>
```

Während `<Aliasname>` ein beliebiger Name für das Zertifikat ist, ist `<Pfad zum Zertifikat>` der vollständige Pfad zum Zertifikat und `<Pfad zum Keystore>` der vollständige Pfad zum Java Keystore, der `<Produkt-Installationsverzeichnis>/lib/security/jssecacerts` oder `<Produkt-Installationsverzeichnis>/lib/security/cacerts` sein kann.

- ▶ Wenn Sie dazu aufgefordert werden, geben Sie das Keystore-Passwort ein (standardmäßig `changeit`).
- ▶ Wenn Sie gefragt werden, ob dem Zertifikat vertraut werden soll, geben Sie `yes` (Ja) ein.

Anweisung an Endbenutzer, SSL zu aktivieren

Wenn Endbenutzer durch ein Client-Produkt eine Verbindung zum Server aufbauen, müssen Sie SSL im Dialogfeld für die Verbindung zum Server aktivieren. Fordern Sie Ihre Benutzer unbedingt auf, das korrekte Kontrollkästchen zu markieren. Beispielsweise sollten die Benutzer in IBM® SPSS® Modeler das Kontrollkästchen `Sichere Verbindung gewährleisten` im Dialogfeld `Repository: Server` aktivieren, das über die Option `Repository` im Menü `Extras` aufgerufen werden kann.

Konfiguration des URL-Präfix

Wenn IBM® SPSS® Collaboration and Deployment Services Repository für SSL-Zugriff eingerichtet wird, muss die Einstellung der URL-Präfix-Konfiguration wie folgt geändert werden:

1. Melden Sie sich beim Repository mit der browserbasierten Konsole an.
2. Öffnen Sie die Konfigurationsoption `URL-Präfix`.
Konfiguration > Setup > URL-Präfix
3. Stellen Sie den Wert des Präfix auf `https` anstelle von `http` ein und setzen Sie den Portwert auf die SSL-Portnummer. Beispiel:

```
[default]
http://<hostname>:<port>
```

[SSL-enabled]
https://<hostname>:<SSLport>

Schützen von LDAP mit SSL

Lightweight Directory Access Protocol (LDAP) ist ein Internet Engineering Task Force- (IETF-)Standard für den Informationsaustausch zwischen Netzwerkverzeichnissen und Datenbanken mit jedem Informationsgehalt. Für Systeme, die zusätzliche Sicherheit benötigen, können LDAP-Anbieter wie Microsofts Active Directory über Secure Socket Layer (SSL) betrieben werden, vorausgesetzt der Web- oder Anwendungsserver unterstützt LDAP über SSL. Die Verwendung von SSL in Kombination mit LDAP kann sicherstellen, dass Anmeldepasswörter, Anwendungsdaten und andere vertrauliche Daten nicht gefährdet, abgefangen oder gestohlen werden.

Das folgende Beispiel illustriert, wie LDAPS mithilfe von Microsofts Active Directory als Sicherheitsanbieter aktiviert wird. Genauere Informationen zu jedem dieser Schritte oder Details zu einem bestimmten Release des Sicherheitsanbieters finden Sie in der originalen Herstellerdokumentation.

1. Stellen Sie sicher, dass Active Directory und die Enterprise-Zertifizierungsstelle installiert sind und funktionieren.
2. Generieren Sie mithilfe der Zertifizierungsstelle ein Zertifikat und importieren Sie dieses Zertifikat in den Zertifikatspeicher der IBM® SPSS® Collaboration and Deployment Services Deployment Manager-Installation. Dies ermöglicht, dass eine LDAPS-Verbindung zwischen IBM® SPSS® Collaboration and Deployment Services Repository und einem Active Directory-Server aufgebaut wird.

Um Deployment Manager für sichere Active Directory-Verbindungen zu konfigurieren, stellen Sie sicher, dass eine Verbindung zum Repository besteht.

3. Starten Sie IBM® SPSS® Collaboration and Deployment Services Deployment Manager.
4. Wählen Sie im Menü “Extras” die Option Server-Verwaltung.
5. Melden Sie sich beim zuvor definierten verwalteten Server an.
6. Doppelklicken Sie auf das Symbol Konfiguration für den Server, um die Hierarchie zu erweitern.
7. Doppelklicken Sie auf das Symbol Sicherheitsanbieter, um die Hierarchie zu erweitern.
8. Doppelklicken Sie auf den Active Directory-Sicherheitsanbieter.
9. Geben Sie Konfigurationswerte für die Active Directory-Instanz mit installierten Sicherheitszertifikaten ein.
10. Markieren Sie das Kontrollkästchen SSL verwenden.
11. Geben Sie den Namen im Feld “Domänenbenutzer” an. Nachfolgende Anmeldungen mit Active Directory werden mit SSL authentifiziert.

Weitere Informationen zum Installieren, Konfigurieren und Implementieren von LDAPS auf einem bestimmten Anwendungsserver finden Sie in der entsprechenden Herstellerdokumentation.

Server-Protokoll

IBM® SPSS® Modeler Server speichert Aufzeichnungen über die wichtigen Aktionen in der Protokolldatei *messages.log*. Standardmäßig befindet sich diese Datei im Verzeichnis *[Serverinstallationspfad]/log*, der Speicherort kann jedoch konfiguriert werden. Beim Start von SPSS Modeler Server wird die Protokolldatei aktualisiert, um den Startzeitpunkt und die Nummer des überwachten Ports anzugeben. Ab diesem Zeitpunkt wird die Protokolldatei jedes Mal aktualisiert, wenn ein Client versucht, eine Verbindung zum Server herzustellen.

Der Server schreibt standardmäßig folgende Ereignisse in die Protokolldatei:

- Server gestartet
- Server gestoppt
- Verbindung akzeptiert
- Verbindung verweigert
- Sitzung gestartet
- Sitzung gestoppt
- Systemfehler
- Anmeldung erfolgreich
- Anmeldung fehlgeschlagen

Wenn Sie Zugriffsinformationen für die Datenbank aufzeichnen möchten, legen Sie in der Konfiguration von SPSS Modeler Server für Datenbankzugriffsinformationen protokollieren und Datenbankfehler protokollieren die Option *Ja* fest. Die folgenden Zugriffsinformationen für die Datenbank werden aufgezeichnet:

- Anmeldung bei Datenbank erfolgreich
- Anmeldung bei Datenbank fehlgeschlagen
- Abmeldung bei Datenbank

Wenn Sie Zugriffsinformationen für Dateien aufzeichnen möchten, legen Sie in der Konfiguration von SPSS Modeler Server für Dateizugriffsinformationen protokollieren und Dateizugriffsfehler protokollieren die Option *Ja* fest. Die folgenden Zugriffsinformationen für Dateien werden aufgezeichnet:

- Datei geöffnet/erstellt
- Datei geschlossen

Im Folgenden finden Sie ein Beispiel für eine Protokolldatei (Zeilennummern wurden hinzugefügt):

1. 2006/06/14 08:41:23 [167]: IBM SPSS Modeler Server Version 0.1.0 started
2. 2006/06/14 08:41:23 [167]: Listening on port 28028
3. 2006/06/14 08:49:28 [167]: Accepted connection from localhost:1827
4. 2006/06/14 08:49:28 [167]: Session 105 (localhost:1827) started
5. 2006/06/14 08:49:29 [167-105]: Login failed for user: SPSS\awilson
6. 2006/06/14 09:27:27 [167]: Session 105 ended
7. 2006/06/14 09:53:39 [167]: Accepted connection from localhost:1837

8. 2006/06/14 09:53:39 [167]: Session 63 (localhost:1837) started
9. 2006/06/14 09:53:40 [167-63]: Login succeeded for user: SPSS\awilson
10. 2006/06/14 10:27:31 [167]: Session 63 ended

Alle Einträge in der Protokolldatei enthalten einen Zeitstempel. Die Zeilen 1 und 2 wurden der Protokolldatei beim Start von SPSS Modeler Server hinzugefügt. Zeile 3 wurde hinzugefügt, als ein Benutzer auf dem Computer *localhost* mit der Portnummer *1827* versuchte, eine Verbindung zum Server herzustellen. Zeile 4 gibt an, dass mit diesem Benutzer eine Serversitzung gestartet wurde. Zeile 5 gibt an, dass sich der Benutzer nicht richtig anmelden konnte. Etwas später hat der Benutzer die Sitzung beendet. Eine halbe Stunde später, wurde ein weiterer Verbindungsversuch von *localhost* durchgeführt und eine Sitzung wurde gestartet. Dieses Mal, war die Anmeldung erfolgreich. Eine weitere halbe Stunde später wurde die Sitzung beendet und die Verbindung getrennt.

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 DBMS und verwenden Sie die SQL-Optimierung, wann immer dies möglich ist.
- Verwenden Sie Hardware, die die Empfehlungen unter [Kapitel 2](#) erfüllt oder übertrifft.
- Stellen Sie sicher, dass die Leistungs- und Optimierungseinstellungen für Ihren Server ordnungsgemäß konfiguriert sind. Einige dieser Einstellungen sind auch im Client verfügbar, sie können jedoch nur bei der Nutzung von IBM® SPSS® Modeler im Standalone-Modus verwendet werden. Sobald IBM® SPSS® Modeler mit einem Server verbunden wird, nutzt der Client automatisch die Leistungs- und Optimierungseinstellungen des Servers.
- Optimieren Sie Streams für maximale Leistung. [Für weitere Informationen siehe Thema Erläuterungen zur Leistung für Streams und Knoten in Kapitel 13 in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)

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 Anwendung IBM® SPSS® Collaboration and Deployment Services Deployment Manager anpassen, die in SPSS Modeler Server enthalten ist. [Für weitere Informationen siehe Thema Verwendung von IBM SPSS Modeler Administration Console in Kapitel 4 auf S. 24.](#)

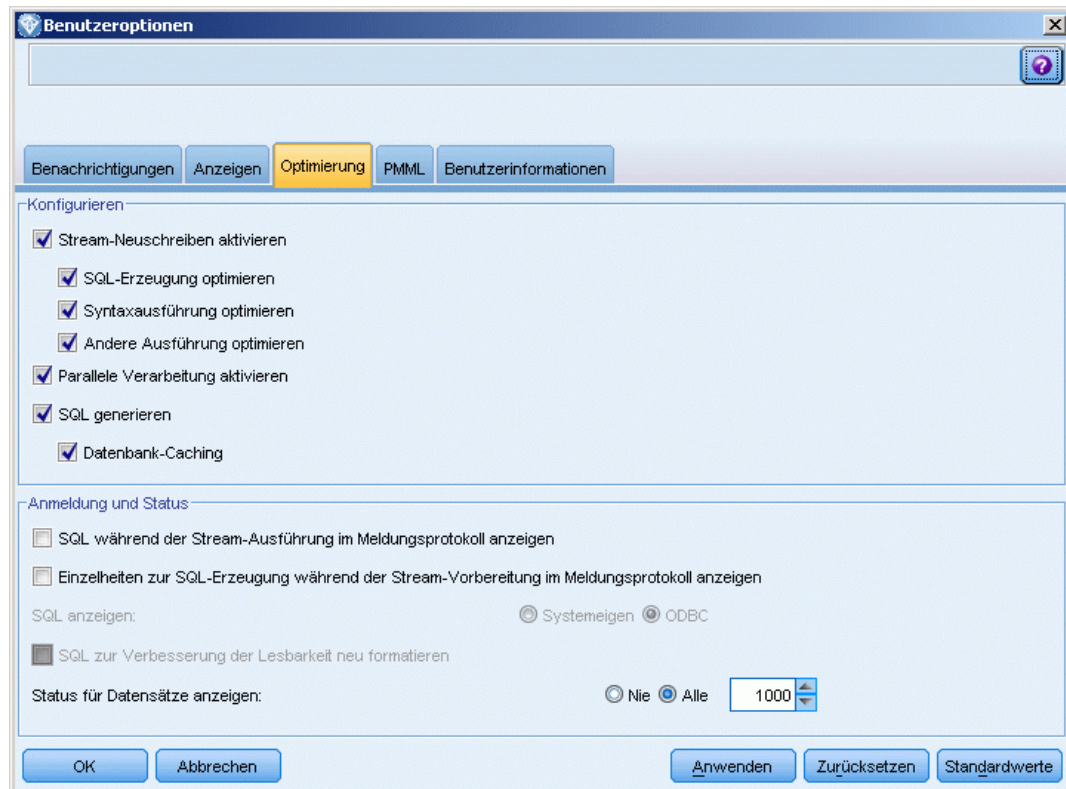
Diese Einstellungen sind unter der Überschrift Leistung und Optimierung im Konfigurationsfenster von Deployment Manager zusammengefasst. Die Einstellungen sind bei den meisten Installationen für die optimale Leistung vorkonfiguriert. Je nach Hardware, Datenmenge und Stream-Inhalt kann es erforderlich sein, diese Einstellungen anzupassen. [Für weitere Informationen siehe Thema Leistung und Optimierung in Kapitel 4 auf S. 29.](#)

Leistungs- und Optimierungseinstellungen für Ihren Client

Auf der Registerkarte “Optimierung” im Dialogfeld “Benutzeroptionen” können Sie die Stream-Leistung optimieren. Beachten Sie, dass die Leistungs- und Optimierungseinstellungen auf IBM® SPSS® Modeler Server (sofern verwendet) alle Einstellungen auf dem Client außer Kraft setzen. *Diese Registerkarte ist nur verfügbar, wenn SPSS Modeler Server-Konnektivität auf dem IBM® SPSS® Modeler-Computer aktiviert wurde.*

Hinweis: Für Datenbankmodellierung und SQL-Optimierung muss auf dem SPSS Modeler-Computer die SPSS Modeler Server-Konnektivität aktiviert sein. Wenn diese Einstellung aktiviert ist, können Sie auf Datenbankalgorithmen zugreifen, SQL direkt aus SPSS Modeler per Pushback übertragen und auf SPSS Modeler Server zugreifen. Um den aktuellen Lizenzstatus zu überprüfen, klicken Sie im Menü “Hilfe” in SPSS Modeler auf Info. [Für weitere Informationen siehe Thema Verbindung mit IBM SPSS Modeler Server in Kapitel 3 in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)

Abbildung 5-1
Dialogfeld “Benutzeroptionen,” Registerkarte “Optimierung”



Anmerkung: Die Unterstützung für SQL-Pushback ist von der verwendeten Datenbanktyp abhängig. Neueste Informationen zu Datenbanken und ODBC-Treibern, die für die Verwendung mit IBM® SPSS® Modeler 14.2 unterstützt werden und getestet wurden, finden Sie in den Produktkompatibilitätsdiagrammen auf der unternehmenseigenen Support-Site unter <http://www.ibm.com/support>.

Stream-Neuschreiben aktivieren. Wählen Sie diese Option, um die Stream-Neuerstellung in SPSS Modeler zu aktivieren. Es stehen zwei Neuerstellungsarten zur Verfügung. Sie können eine oder beide auswählen. Die Stream-Neuerstellung ordnet die Knoten in einem Stream für eine effizientere Operation im Hintergrund neu an, ohne die Stream-Semantik zu ändern.

- **SQL-Erzeugung 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 weiter unten im Stream nach Knoten, die in SQL wiedergegeben und ohne Beeinträchtigung der Stream-Semantik vor den problematischen Knoten verschoben werden können. Nicht nur, dass die Datenbank die Operationen effizienter als SPSS Modeler durchführen kann, solche Pushbacks dienen auch dazu, die Größe des Daten-Sets, das an SPSS Modeler zur Verarbeitung zurückgegeben wird, zu reduzieren. So wird wiederum der Netzverkehr reduziert und die Geschwindigkeit der Stream-Operationen erhöht. Das Kontrollkästchen SQL generieren muss aktiviert sein, damit die SQL-Optimierung funktionieren kann. [Für weitere Informationen siehe Thema SQL-Optimierung in Kapitel 6 auf S. 50.](#)

- **Syntaxausführung optimieren.** Diese Methode der Stream-Neuerstellung 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 Stream-Neuerstellung 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. Bei der Verwendung mehrerer Knoten oder der folgenden Einzelknoten können Sie von der parallelen Verarbeitung profitieren: C5.0, Verbinden (Zusammenführungsknoten, durch Schlüssel), Sortieren, Intervall (Klassierknoten, Rangordnung und Perzentile) sowie Aggregieren (unter Verwendung eines oder mehrerer Schlüsselfelder).

SQL generieren. Wählen Sie diese Option zur Aktivierung der SQL-Erzeugung. Dabei wird zugelassen, dass Stream-Operationen 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-Erzeugung 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.

- **Datenbank-Caching.** Bei Streams, die in der Datenbank auszuführende 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 Datenbank-Caching 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 Knoten weiter unten im Stream erzeugt werden, was zu weiteren Leistungsverbesserungen führt. (Alternativ kann diese Option bei Bedarf deaktiviert werden, beispielsweise wenn es aufgrund von Richtlinien oder Berechtigungen nicht möglich ist, Daten in die Datenbank zu

schreiben. Wenn Datenbank-Caching oder SQL-Optimierung nicht aktiviert ist, wird der Cache stattdessen in das Dateisystem geschrieben. [Für weitere Informationen siehe Thema Cachen von Optionen für Knoten in Kapitel 5 in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)

Hinweis: 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 SPSS Modeler gelieferten Ergebnissen ab. Aus ähnlichen Gründen können diese Unterschiede auch je nach Datenbankanbieter variieren.

SQL wird während der Stream-Ausführung im Meldungsprotokoll angezeigt. Gibt an, ob die während der Ausführung des Streams erzeugte SQL an das Meldungsprotokoll übergeben wird.

Einzelheiten zur SQL-Erzeugung während der Stream-Vorbereitung im Meldungsprotokoll anzeigen. Legt während der Stream-Vorschau fest, ob eine Vorschau der SQL, die generiert werden würde, an das Meldungsprotokoll übergeben wird.

SQL anzeigen. Legt fest, ob eine im Protokoll angezeigte SQL systemeigene SQL-Funktionen oder Standard-ODBC-Funktionen in der Form {fn FUNC(...)}, wie von SPSS Modeler erzeugt, enthalten soll. Ersteres richtet sich nach den ODBC-Treiberfunktionen, die u. U. nicht implementiert werden. Dieses Steuerelement hätte z. B. für SQL Server keinerlei Auswirkungen.

SQL zur Verbesserung der Lesbarkeit neu formatieren. Legt fest, ob im Protokoll angezeigte SQL zur Verbesserung der Lesbarkeit formatiert werden soll.

Status für Datensätze anzeigen. Legt fest, ob Datensätze gemeldet werden sollen, wenn sie an Endknoten ankommen. Geben Sie eine Zahl an, um den Status alle N Datensätze zu aktualisieren.

Klicken Sie auf Standardwerte, um zu den Systemstandardeinstellungen für diese Registerkarte zurückzukehren.

Datenbankverwendung und -optimierung

Datenbankserver. Wenn möglich, erstellen Sie eine dedizierte Datenbankinstanz für das Data-Mining, sodass der Produktionsserver nicht von IBM® SPSS® Modeler-Anfragen beeinflusst wird. Die von 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.

In-Database Mining. Viele Datenbankhersteller 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 In-Database Mining-Funktionen von SPSS Modeler ergänzen und erweitern die Fähigkeiten des Produkts hinsichtlich der SQL-Erzeugung und bieten auf diese Weise eine Möglichkeit, die herstellereigene Datenbankerweiterungen zu steuern. In manchen Fällen kann diese Herangehensweise den manchmal kostenintensiven Aufwand für den Datentransfer zwischen SPSS Modeler und der Datenbank reduzieren. Auch das Datenbank-Caching kann dazu beitragen, diese Vorteile weiter auszubauen. Weitere Informationen finden Sie in der Datei *DatabaseMiningGuide.pdf* im Ordner *\Documentation* auf der IBM® SPSS® Modeler Installations-CD.

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 unter [Kapitel 6](#).

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. [Für weitere Informationen siehe Thema Erweiterte Optionen für den Datenbankexport in Kapitel 7 in IBM SPSS Modeler 14.2- Quellen-, Prozess- und Ausgabeknoten.](#)

IBM® SPSS® Modeler kann mit den externen Ladeprogrammen für viele gängige Datenbanksysteme zusammenarbeiten. In der Software sind mehrere Skripts enthalten. Diese stehen (mit Dokumentation) im Unterverzeichnis `/scripts` unter dem 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 5-1
Leistungsvorteil des Massenladens

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

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 Erzeugung 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 Daten-Sets arbeiten, können diese **Pushbacks** die Leistung auf unterschiedliche Weise erheblich steigern.

- Durch Reduzierung der Größe des Ergebnis-Sets, das aus dem DBMS nach SPSS Modeler übertragen werden soll. Wenn große Ergebnis-Sets mithilfe eines ODBC-Treibers gelesen werden, kann es zu Netzwerk-E/A- oder Treiberfehlern kommen. Aus diesem Grund sind Zeilen- und Spaltenauswahl und -aggregation (Auswahl-, Stichproben-, Aggregatknoten), mit denen üblicherweise die Größe des zu übertragenden Daten-Sets verringert wird, die Operationen, die am meisten von der SQL-Optimierung profitieren. Daten können außerdem an kritischen Punkten im Stream (z. B. nach einem Merge- oder Auswahlknoten) 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 DBMS häufig die Vorteile paralleler Verarbeitung, leistungsstärkerer Hardware, einer ausgereifteren Verwaltung des Speicherplatzes und des Vorhandenseins von Indizes nutzen kann.

Angesichts dieser Vorteile wurde SPSS Modeler so entworfen, dass es die Menge an SQL, die von den einzelnen Streams erzeugt wird, maximiert, sodass nur diejenigen Operationen, die sich nicht in SQL kompilieren lassen, von SPSS Modeler Server ausgeführt werden. Aufgrund der Beschränkungen der Ausdrucksmöglichkeiten in der Standard-SQL (SQL-92) werden bestimmte Operationen nicht unterstützt. [Für weitere Informationen siehe Thema Tipps für die Maximierung der SQL-Erzeugung auf S. 56.](#)

Hinweis: 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 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 Zeichenketten können 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 Datenbankverwalter, wenn Sie Unterstützung bei der Konfiguration Ihrer Datenbank benötigen. Für größtmögliche Kompatibilität mit SPSS Modeler sollte beim Vergleich von Datenbankzeichenketten zwischen Groß- und Kleinschreibung unterschieden werden.

Hinweis: Für Datenbankmodellierung und SQL-Optimierung muss auf dem SPSS Modeler-Computer die SPSS Modeler Server-Konnektivität aktiviert sein. Wenn diese Einstellung aktiviert ist, können Sie auf Datenbankalgorithmen zugreifen, SQL direkt aus SPSS Modeler per Pushback übertragen und auf SPSS Modeler Server zugreifen. Um den aktuellen Lizenzstatus zu überprüfen, klicken Sie im Menü “Hilfe” in SPSS Modeler auf Info. [Für weitere Informationen siehe Thema Verbindung mit IBM SPSS Modeler Server in Kapitel 3 in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)

Datenbankanforderungen

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

Beachten Sie, dass durch Verwendung von Datenbankmodellierung eventuell weitere Leistungsverbesserungen möglich sind. [Für weitere Informationen siehe Thema Übersicht über die Datenbank-Modellierung in Kapitel 2 in IBM SPSS Modeler 14.2 – In-Database Mining-Handbuch.](#)

Funktionsweise der SQL-Erzeugung

Die ersten Fragmente eines Streams, der von den Quellenknoten der Datenbank ausgeht, sind die Hauptziele für die SQL-Erzeugung. 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 Stream-Vorbereitung und vor der Ausführung findet die SQL-Erzeugung wie folgt statt.

- Der Server ordnet Streams neu, um weiter unten liegende Knoten in die “SQL-Zone” zu verschieben, sofern es erwiesenermaßen problemlos möglich ist, dies zu tun. (Diese Funktion kann auf dem Server deaktiviert 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.
- Die SQL wird ausgehend von den Knoten mit den kompliziertesten SQL-Äquivalenten zurück zu den Quellenknoten auf Validität überprüft. Die SQL, die erfolgreich validiert wurde, wird für die Ausführung ausgewählt.
- Knoten, für die alle Operationen SQL erzeugt haben, werden im Stream-Zeichenbereich 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. [Für weitere Informationen siehe Thema Tipps für die Maximierung der SQL-Erzeugung auf S. 56.](#)

Auftreten von Verbesserungen

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

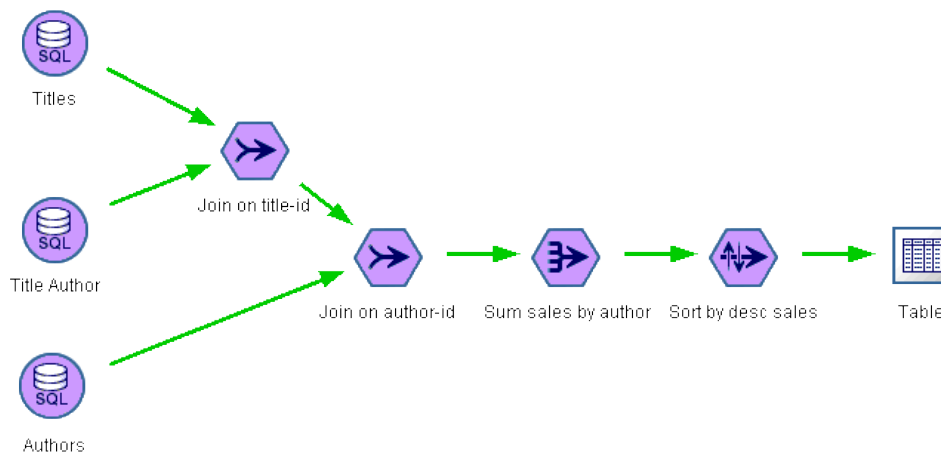
- **Joins (Zusammenführen (Merge) nach Schlüssel)** Join-Operationen können die Optimierung innerhalb von Datenbanken steigern.
- **Aggregation.** Die Aggregat-, Verteilungs- und Webknoten verwenden jeweils Aggregation zum Erstellen der Ergebnisse. Zusammengefasste Daten benötigen erheblich weniger Bandbreite 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.** 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 Textdateien: Der Server muss die überflüssigen Felder zwar lesen, er weist jedoch keinen Speicher für sie zu.
- **Scoring.** SQL lässt sich aus Entscheidungsbäumen, Regelmengen, linearer Regression und faktorgenerierten Modellen erzeugen.

Beispiel für SQL-Erzeugung

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

Abbildung 6-1

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



Erzeugte SQL

Die für diesen Stream erzeugte SQL lautet wie folgt:

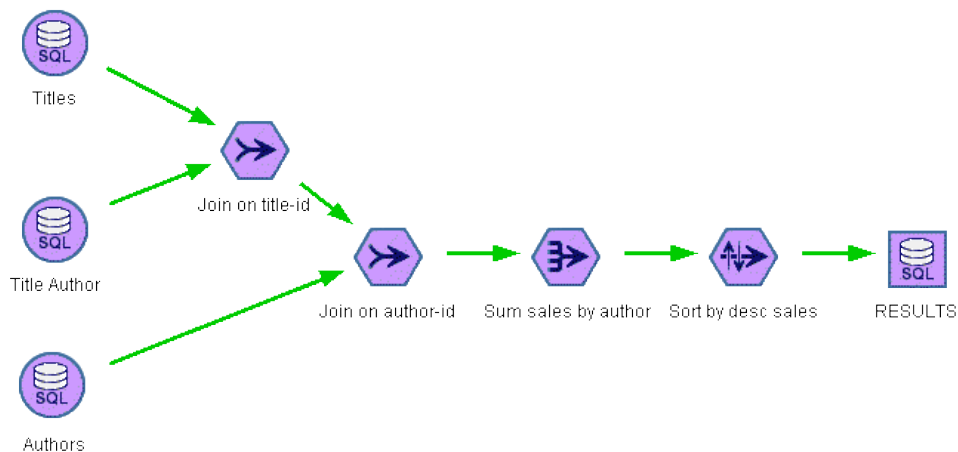
```
SELECT
  T2.au_name AS C0,
  T2.au_fname AS C1,
  SUM({fn CONVERT(T0.ytd_sales,SQL_BIGINT)}) AS C2
FROM
  dbo.titles T0,
  dbo.titleauthor T1,
  dbo.authors T2
WHERE
  (T0.title_id = T1.title_id )
  AND (T1.au_id = T2.au_id )
GROUP BY T2.au_name ,T2.au_fname
ORDER BY 3 DESC
```

Ausführen des Streams

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

Abbildung 6-2

Gesamter, in der Datenbank ausgeführter Stream



Konfigurieren der SQL-Optimierung

- ▶ Installieren Sie einen ODBC-Treiber und konfigurieren Sie eine Datenquelle für die zu verwendende Datenbank. [Für weitere Informationen siehe Thema Datenzugriff in Kapitel 2 auf S. 11.](#)
- ▶ Erstellen Sie einen Stream, der einen Quellenknoten zum Extrahieren von Daten aus dieser Datenbank verwendet.
- ▶ Vergewissern Sie sich gegebenenfalls, dass die SQL-Erzeugung auf Client und Server aktiviert ist. Standardmäßig ist sie für beide aktiviert.

So aktivieren Sie SQL-Optimierung auf dem Client

- ▶ Wählen Sie im Menü “Extras” die Option Benutzeroptionen aus.
- ▶ 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. [Für weitere Informationen siehe Thema Leistungs- und Optimierungseinstellungen für Ihren Client in Kapitel 5 auf S. 45.](#)

So aktivieren Sie SQL-Optimierung auf dem Server

Da die Servereinstellungen alle auf dem Client angegebenen Spezifikationen überschreiben, müssen die Serverkonfigurationseinstellungen Stream-Neuschreiben und Automatische SQL-Erzeugung jeweils aktiviert sein. Weitere Einstellungen zum Ändern der IBM® SPSS® Modeler Server-Einstellungen finden Sie im Abschnitt [Leistung und Optimierung in Kapitel 4.](#)

So aktivieren Sie die Optimierung beim Scoring von Modellen

Zum Zwecke des Scorings muss die SQL-Erzeugung 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 die generierte SQL ausgeführt werden soll, weil die SQL zu groß oder zu komplex ist.

Unter Umständen erfordert es einige Anläufe, um festzustellen, ob die SQL-Erzeugung die Leistung eines bestimmten Modells verbessert. Dies geschieht auf der Registerkarte “Einstellungen”, nachdem ein generiertes Modell zu einem Stream hinzugefügt wurde. [Für weitere Informationen siehe Thema Verwendung von Modell-Nuggets in Streams in Kapitel 3 in IBM SPSS Modeler 14.2-Modellierungsknoten.](#)

Vorschau für erzeugte SQL

Sie können die erzeugte SQL im Meldungsprotokoll in der Vorschau anzeigen, bevor sie in der Datenbank ausgeführt wird. Dies ist u. U. für die Fehlersuche nützlich und ermöglicht außerdem den Export der erzeugten 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.

- ▶ Vergewissern Sie sich, dass die Optionen SQL während der Stream-Ausführung im Meldungsprotokoll anzeigen und Einzelheiten zur SQL-Erzeugung während der Stream-Vorbereitung im Meldungsprotokoll anzeigen im Dialogfeld “Benutzeroptionen” ausgewählt sind. [Für weitere Informationen siehe Thema Leistungs- und Optimierungseinstellungen für Ihren Client in Kapitel 5 auf S. 45.](#)
- ▶ Wählen Sie im Stream-Zeichenbereich den Knoten oder Stream aus, den Sie in der Vorschau anzeigen möchten.

- Klicken Sie in der Symbolleiste auf die Schaltfläche für die SQL-Vorschau.

Alle Knoten, für die SQL erzeugt wird (und die bei der Ausführung des Streams per Pushback in die Datenbank zurückübertragen werden) werden im Stream-Zeichenbereich violett gekennzeichnet.

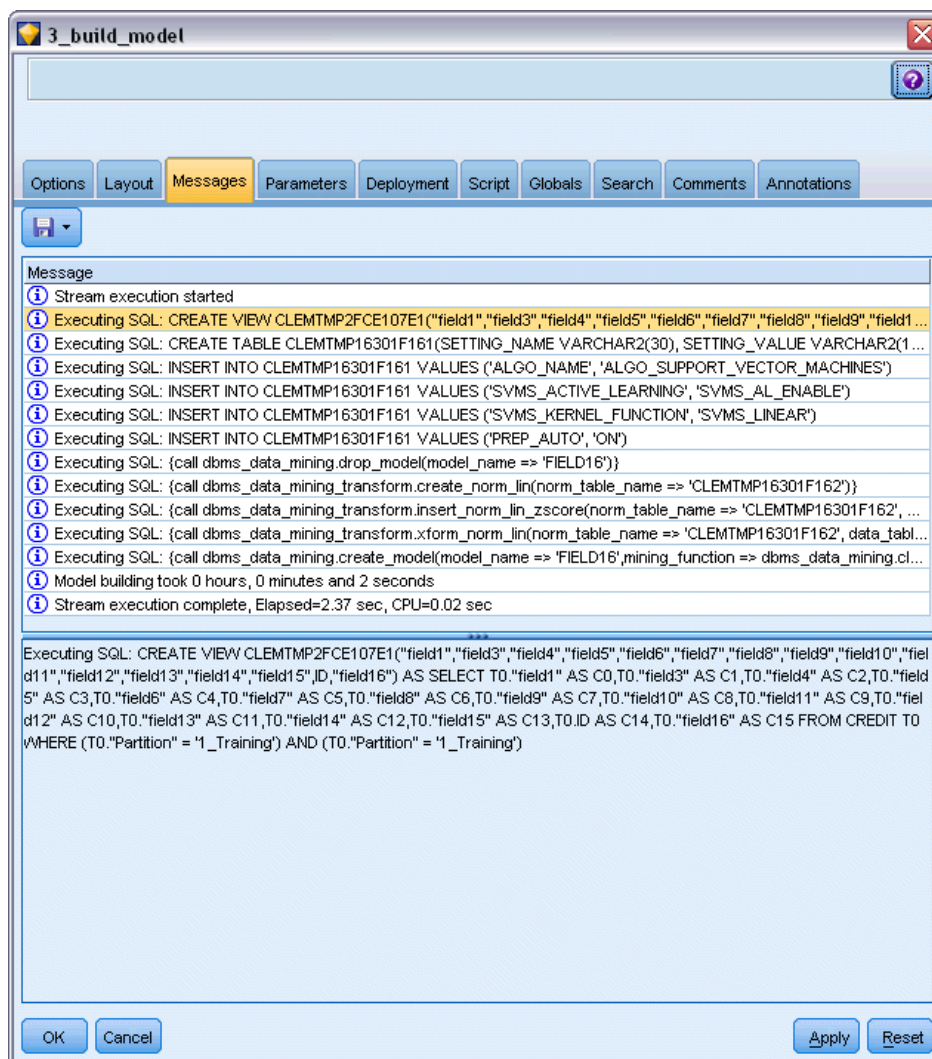
Abbildung 6-3
Schaltfläche für die SQL-Vorschau



- Um die erzeugte SQL in der Vorschau anzuzeigen, wählen Sie die folgende Optionsfolge aus den Menüs aus:

Werkzeuge > Stream-Eigenschaften > Nachrichten...

Abbildung 6-4
Im Meldungsprotokoll angezeigte erzeugte SQL



SQL für Modell-Nuggets anzeigen

Bei einigen Modellen kann SQL für das Modell-Nugget erzeugt werden, wobei die Modell-Scoring-Phase per Pushback an die Datenbank zurückgegeben wird. Diese Funktion dient nicht in erster Linie zur Leistungsverbesserung, sondern dazu, dass die gesamte SQL von Streams, die diese Nuggets enthalten, per Pushback zurückgeführt werden kann. [Für weitere Informationen siehe Thema Knoten zur Unterstützung der SQL-Erzeugung auf S. 58.](#)

Führen Sie folgende Schritte aus, um die SQL für ein Modell-Nugget anzuzeigen, das SQL-Erzeugung unterstützt:

- ▶ Wählen Sie die Registerkarte “Einstellungen” auf dem Modell-Nugget.
- ▶ Wählen Sie nach Bedarf eine der Optionen Mit (Ohne) Unterstützung für fehlende Werte generieren oder SQL für dieses Modell generieren.
- ▶ Wählen Sie im Menü des Modell-Nuggets:
Datei > SQL exportieren
- ▶ Speichern Sie die Datei.
- ▶ Öffnen Sie die Datei, um die SQL anzuzeigen.

Tipps für die Maximierung der SQL-Erzeugung

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

Stream-Reihenfolge. Die SQL-Erzeugung wird 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-Erzeugung auch für alle weiter unten im Stream liegenden 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. Der SQL-Optimizer kann Neuordnungen bis zu einem bestimmten Grad automatisch durchführen (sie müssen jedoch sicherstellen, dass Stream-Neuschreiben 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. [Für weitere Informationen siehe Thema Knoten zur Unterstützung der SQL-Erzeugung auf S. 58.](#)

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-Erzeugung 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-Erzeugung behindern (insbesondere die Sequenzfunktionen [”@-Funktionen”]). Manchmal wird die Generierung angehalten, da die generierte Abfrage so komplex wurde,

sodass die Datenbank nicht mehr damit umgehen kann. [Für weitere Informationen siehe Thema CLEM-Ausdrücke und Operatoren, die die SQL-Erzeugung unterstützen auf S. 61.](#)

Mehrere Quellenknoten. Wenn ein Stream mehrere Datenbankquellenknoten aufweist, wird die SQL-Erzeugung auf jeden Eingabezweig unabhängig angewendet. Wenn die Erzeugung in einem Zweig angehalten wird, kann sie in einem anderen fortgesetzt werden. Wenn zwei Zweige mit einem Merge-Vorgang zusammengeführt werden (und beide Zweige bis zum Merge-Punkt in SQL ausgedrückt werden können), lässt sich die Zusammenführung oft durch einen Datenbank-Join ersetzen und die Erzeugung kann weiter unten im Stream 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. [Für weitere Informationen siehe Thema Übersicht über die Datenbank-Modellierung in Kapitel 2 in IBM SPSS Modeler 14.2 – In-Database Mining-Handbuch.](#)

Scoring von Modellen. 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-Erzeugung separat für die einzelnen Modellknoten aktiviert werden. Wenn sich herausstellt, dass der Modellknoten die SQL-Erzeugung 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-Erzeugung zur Verfügung. Führen Sie Tests aus, um sicherzustellen, dass die Option für Ihre Anwendung nützlich ist. [Für weitere Informationen siehe Thema Knoten zur Unterstützung der SQL-Erzeugung auf S. 58.](#)

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, die (möglicherweise komplexe) generierte SQL zu verarbeiten, und es kann erforderlich sein, SPSS Modeler über den Windows Task-Manager zu schließen.

Datenbank-Caching. Wenn Sie einen Knoten-Cache zum Speichern von Daten an kritischen Punkten im Stream verwenden (beispielsweise nach einem Merge- oder Aggregatknoten), müssen Sie sicherstellen, dass neben der SQL-Optimierung auch Datenbank-Caching aktiviert ist. Dadurch können die Daten in den meisten Fällen eine temporäre Tabelle in der Datenbank (statt des Dateisystems) als Cache fungieren. [Für weitere Informationen siehe Thema Konfigurieren der SQL-Optimierung auf S. 53.](#) Weitere Informationen finden Sie auch hier: [Cachen von Optionen für Knoten.](#)

Herstellerspezifische SQL. Der größte Teil der 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-Erzeugung

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

Sie können die zu erzeugende SQL anzeigen, bevor Sie sie ausführen. [Für weitere Informationen siehe Thema Vorschau für erzeugte SQL auf S. 54.](#)

Quellen

Knoten zur Unterstützung der SQL-Erzeugung	Anmerkungen
Datenbank	Dieser Knoten wird verwendet, um Tabellen und Ansichten für eine weitere Analyse anzugeben. Dieser Knoten ermöglicht die Eingabe von SQL-Abfragen. Vermeiden Sie Ergebnis-Sets mit doppelten Spaltennamen. Für weitere Informationen siehe Thema Schreiben von SQL-Abfragen auf S. 64.

Datensatzoperationen

Knoten zur Unterstützung der SQL-Erzeugung	Anmerkungen
Auswählen	Unterstützt die Erzeugung nur, wenn SQL-Erzeugung 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.
Muster	Bei der einfachen Stichprobenziehung wird die SQL-Erzeugung unterstützt, wie in der folgenden Tabelle gezeigt. Bei der Ziehung komplexer Stichproben wird die SQL-Erzeugung nicht unterstützt.
Aggregieren	
RFM-Aggregat	Unterstützt die Erzeugung, 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(YEAR,MONTH,DAY)</code> ein Pushback-Vorgang durchgeführt wird.
Sortieren	
Zusammenführen	Keine SQL generiert für Zusammenführen (Merge) nach Schlüssel. Ein Zusammenführen nach Schlüssel mit Full oder partiellem Outer Join wird nur unterstützt, falls sie auch von der Datenbank/dem Treiber unterstützt wird. Nicht übereinstimmende Eingabefelder können mittels eines Filterknotens oder über die Registerkarte "Filter" eines Quellenknotens umbenannt werden. Für alle Merge-Typen wird <code>SQL_SP_EXISTS</code> nicht unterstützt, wenn die Eingaben aus verschiedenen Datenbanken stammen.
Anhängen	Unterstützt die Erzeugung bei unsortierten Eingaben.
Duplikat	

Die Unterstützung für SQL-Erzeugung im Stichprobenknoten bei Verwendung einer einfachen Stichprobenziehung wird in der nächsten Tabelle gezeigt.

Modalwert	Muster	Max. Größe	Startwert	DB2 OS/Z	DB2 OS/400	DB2 Windows	Netezza Windows/UNIX	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			J	J	J			J
			auf			J		J			
		max	Aus			J	J	J			J
auf					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			J	J	J			J
			auf			J		J			
		max	Aus			J	J	J			J
			auf			J		J			

Feldfunktionen

Knoten zur Unterstützung der SQL-Erzeugung	Anmerkungen
Typ	Unterstützt die SQL-Erzeugung, wenn der Typknoten instanziiert wird und keine ABORT- oder WARN-Typprüfung angegeben ist.
Filter	
Ableiten	Unterstützt die SQL-Erzeugung, wenn die für den Ableitungsausdruck erzeugte SQL unterstützt wird (siehe Ausdrücke unten).
Ensemble	Unterstützt SQL-Erzeugung für stetige Ziele. Bei anderen Zielen wird die Erzeugung nur unterstützt, wenn die Ensemble-Methode "Höchste Konfidenz hat Vorrang" verwendet wird.
Füller	Unterstützt die SQL-Erzeugung, wenn die für den "derive"-Ausdruck erzeugte SQL unterstützt wird (siehe Ausdrücke unten).
Anonymisieren	Unterstützt die SQL-Erzeugung für stetige Ziele und die partielle SQL-Erzeugung für nominale und Flag-Ziele.
Umkodieren	
Klassierung	Unterstützt SQL-Erzeugung, wenn die Klassiermethode "n-tile (gleiche Anzahl)" verwendet wird und die Option "Von Registerkarte "Klassenwerte" lesen, sofern verfügbar" ausgewählt ist.
RFM-Analyse	Unterstützt SQL-Generierung, wenn die Option "Von Registerkarte "Klassenwerte" lesen, sofern verfügbar" ausgewählt ist, jedoch unterstützen dies weiter abwärts liegende Knoten nicht.
Partitionieren	Unterstützt die SQL-Erzeugung für die Zuweisung von Datensätzen zu Partitionen.
Dichotom	
Umstrukturieren	

Grafiken

Knoten zur Unterstützung der SQL-Erzeugung	Anmerkungen
Diagrammtafel	SQL-Erzeugung wird für folgende Diagrammtypen unterstützt: Fläche, 3D-Fläche, Balken, 3D-Balken, Balken für Häufigkeiten, Hitzekarte, Kreis, 3D-Kreis, Kreis für Häufigkeiten. Bei Histogrammen wird SQL-Erzeugung nur für kategoriale Daten unterstützt.
Verteilung	
Internet	
Evaluation	

Generierte Modellknoten

Bei einigen Modellen kann SQL für das Modell-Nugget erzeugt werden, wobei die Modell-Scoring-Phase per Pushback an die Datenbank zurückgegeben wird. Diese Funktion dient nicht in erster Linie zur Leistungsverbesserung, sondern dazu, dass die gesamte SQL von Streams, die diese Nuggets enthalten, per Pushback zurückgeführt werden kann. [Für weitere Informationen siehe Thema SQL für Modell-Nuggets anzeigen auf S. 56.](#)

SQL-Erzeugung unterstützendes Modell-Nugget	Anmerkungen
C&R-Baum	Unterstützung der SQL-Erzeugung für die Option für einzelne Bäume, nicht jedoch zur Verbesserung, Verstärkung oder für große Daten-Set-Optionen.
QUEST	
CHAID	
C5.0	
Entscheidungsliste	
Linear	Unterstützung der SQL-Erzeugung für die Standardmodelloption, nicht jedoch zur Verbesserung, Verstärkung oder für große Daten-Set-Optionen.
Netzwerk	Unterstützung der SQL-Erzeugung für die Standardmodelloption (nur Mehrschicht-Perzeptron), nicht jedoch zur Verbesserung, Verstärkung oder für große Daten-Set-Optionen.
Faktor/PCA	
Logistisch	Unterstützt SQL-Erzeugung für das multinomiale, nicht jedoch für das binomiale Verfahren. Beim multinomialen Verfahren wird die Erzeugung nicht unterstützt, wenn Konfidenzen ausgewählt wurden, es sei denn, der Zieltyp ist "Flag".
Generierte Regelmengen	

Ausgabe

Knoten zur Unterstützung der SQL-Erzeugung	Anmerkungen
Tabelle	Unterstützt die Erzeugung, wenn die SQL-Erzeugung für den highlight-Ausdruck (siehe Ausdrücke unten) unterstützt wird.
Matrix	Unterstützt die Erzeugung, außer wenn "Alle numerischen Werte" für die Option "Felder" ausgewählt wurde.
Analyse	Unterstützt die Erzeugung je nach den ausgewählten Optionen.

Knoten zur Unterstützung der SQL-Erzeugung	Anmerkungen
Transformieren	
Statistics	Unterstützt die Erzeugung, wenn die Option "Korrelieren" nicht verwendet wird.
Bericht	
Globalwerte	

Exportieren

Knoten zur Unterstützung der SQL-Erzeugung	Anmerkungen
Datenbank	
Publisher	Der veröffentlichte Stream enthält die erzeugte SQL.

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

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

Operatoren

Operation zur Unterstützung der SQL-Erzeugung	Anmerkungen
+	
-	
/	
*	
>>	Wird zur Verkettung von Zeichenfolgen verwendet.

Vergleichsoperatoren

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

Funktionen

Operation zur Unterstützung der SQL-Erzeugung	Anmerkungen
abs	
allbutfirst	
allbutlast	
and	
arccos	
arcsin	
arctan	
arctanh	
cos	
div	
exp	
fracof	
hasstartstring	
hassubstring	
integer	
intof	
isaplhacode	
islowercode	
isnumbercode	
isstartstring	
issubstring	
isuppercode	
last	
length	
locchar	
log	
log10	
lowertoupper	
max	
member	
min	
negate	
not	
number	
or	
pi	
real	
rem	
round	
sign	
sin	
sqrt	

Operation zur Unterstützung der SQL-Erzeugung	Anmerkungen
string	
strmember	
subscrs	
substring	
substring_between	
uppertolower	
to_string	

Sonderfunktionen

Operation zur Unterstützung der SQL-Erzeugung	Anmerkungen
@NULL	
@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	

Aggregate-Funktionen

Operation zur Unterstützung der SQL-Erzeugung	Anmerkungen
Sum	
Mean	
Min	
Max	
Count	
SDev	

Verwenden von SQL-Funktionen in CLEM-Ausdrücken

Die Funktionen @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 proprietäre 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 Corp. entziehen. Ein Deployment ist jedoch in Sonderfällen möglich, normalerweise im Rahmen einer

Dienstleistungsabspache. 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 Daten-Set mit doppelten Spaltennamen führen. Diese doppelten Namen verhindern häufig die SQL-Optimierung für alle abwärts liegenden Knoten.

IBM® SPSS® Modeler verwendet verschachtelte SELECT-Anweisungen, um ein Pushback von SQL für Streams durchzuführen, die eine SQL-Abfrage im Datenbank-Quellenknoten verwenden. Mit anderen Worten: Der Stream verschachtelt die im Datenbank-Quellenknoten angegebene Abfrage in eine oder mehrere SELECT-Anweisungen, die während der Optimierung von abwärts liegenden Knoten erstellt werden. Wenn also das Ergebnis-Set einer Abfrage doppelte Spaltennamen enthält, kann die Anweisung nicht vom RDBMS 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 Daten-Set 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;
```


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 die erzeugte SQL in der Datenbank vollständig optimiert wird.

Richtige Ländereinstellung

Wenn Sie 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 SPSS Modeler Server-Computer mit einer anderen Ländereinstellung ausgeführt wird oder wenn für die Shell, in der SPSS Modeler Server gestartet wurde, die Ländereinstellung nicht vollständig definiert wurde, wird die erzeugte SQL u. U. in Oracle nicht vollständig optimiert. Dafür gibt es folgende Gründe:

- 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 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 SPSS Modeler Server ausgegeben. Dies bedeutet, dass SPSS Modeler keine Streams für Oracle optimieren kann.

So prüfen Sie die Ländereinstellungen und setzen Sie zurück:

- ▶ Führen Sie in einer UNIX-Shell folgenden Befehl aus:

```
#locale
```

Sie erhalten Informationen über die 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
```

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

- ▶ Nachdem Sie den Link erstellt haben, müssen Sie SPSS Modeler Server aus derselben Shell neu starten. ((SPSS Modeler Server erhält seine Ländereinstellungsinformationen aus der Shell, aus der er gestartet wurde.)

Hinweise

Wenn Sie einen UNIX-Rechner 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 In-Database Mining zu optimieren, müssen sowohl **LANG** als auch **LC_ALL** in der zum Starten von SPSS Modeler Server verwendeten Shell definiert sein. **LANG** kann in der Shell genauso definiert werden wie jede andere Umgebungsvariable, bevor SPSS Modeler Server neu gestartet wird. Siehe folgende Definition:

```
#LANG=en_US.ISO8859-15; export LANG
```

- Bei jedem Start von 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.

Konfigurieren von UNIX-Startskripts

Einführung

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

- Konfigurieren von SPSS Modeler Server für einen automatischen Start beim Neustart des Servercomputers.
- Manuelles Stoppen und Neustarten von SPSS Modeler Server.
- Ändern der Berechtigungen für die von SPSS Modeler Server erstellten Dateien.
- Konfigurieren von SPSS Modeler Server, um mit den in SPSS Modeler Server enthaltenen ODBC Connect-Treibern zu arbeiten. [Für weitere Informationen siehe Thema IBM SPSS Modeler Server und das Data Access-Pack auf S. 69.](#)
- Konfigurieren von SPSS Modeler Server für die Verwendung des IBM® SPSS® Collaboration and Deployment Services Enterprise View Driver-Treibers beim Lesen von Daten mit einem Enterprise View-Knoten. [Für weitere Informationen siehe Thema Konfigurieren eines Treibers für den Enterprise-Ansichts-Knoten auf S. 76.](#)

Skripts

IBM® SPSS® Modeler Server verwendet verschiedene Skripts, darunter:

- **modelersrv.sh.** Das Skript für den manuellen Start von SPSS Modeler Server befindet sich im Installationsverzeichnis von SPSS Modeler Server. Es konfiguriert die Serverumgebung, wenn der Serverdaemonprozess *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 Skript wird Ihr System konfiguriert, um den Serverdaemonprozess beim Neustart automatisch zu starten. Führen Sie das Skript ein Mal aus, um Ihr System für den automatischen Start zu konfigurieren. Sie müssen es nicht bearbeiten. Das Skript befindet sich im Installationsverzeichnis von SPSS Modeler Server.
- **rc.modeler.** Beim Ausführen von *auto.sh* wird dieses Skript 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 B-1

Speicherort von "rc.modeler" unter den verschiedenen Betriebssystemen

Betriebssystem	Ort
AIX	/etc/rc.modeler
HP-UX	/sbin/init.d/rc.modeler
Solaris	/etc/init.d/rc.modeler

Automatisches Starten und Stoppen von IBM SPSS Modeler Server

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

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

- ▶ Melden Sie sich als “root” an.
- ▶ Wechseln Sie zum SPSS Modeler Server-Installationsverzeichnis.
- ▶ Führen Sie das Skript aus. Geben Sie an der UNIX-Eingabeaufforderung Folgendes ein:

```
./auto.sh
```

Das automatische Startskript *rc.modeler* wird an dem in der vorangehenden Tabelle angegebenen Speicherort erstellt. Das Betriebssystem verwendet *rc.modeler*, um den SPSS Modeler Server-Daemonprozess zu starten, wenn der Servercomputer neu gestartet wird. Außerdem verwendet das Betriebssystem *rc.modeler* auch, um den Daemon 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 Skript *modelersrv.sh* ausführen.

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

- ▶ Wechseln Sie zum SPSS Modeler Server-Installationsverzeichnis.
- ▶ Um den Server zu starten, geben Sie an der UNIX-Eingabeaufforderung Folgendes ein:

```
./modelersrv.sh start
```

- ▶ Um den Server anzuhalten, geben Sie an der UNIX-Eingabeaufforderung Folgendes ein:

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

Bearbeiten von Skripten

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 Skripts:

- ▶ Stoppen Sie IBM® SPSS® Modeler Server. (Für weitere Informationen siehe Thema [Manuelles Starten und Stoppen von IBM SPSS Modeler Server auf S. 68.](#))

- ▶ Wählen Sie das entsprechende Skript aus. (Für weitere Informationen siehe Thema Skripts auf S. 67.)
- ▶ Öffnen Sie das Skript in einem Texteditor, nehmen Sie die Änderungen vor und speichern Sie die Datei.
- ▶ Starten Sie SPSS Modeler Server, entweder automatisch (indem Sie den Servercomputer 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 Startskript *modelersrv.sh*, *rc.modeler* oder in beiden die Einstellung **UMASK** bearbeiten. (Weitere Informationen finden Sie unter [Bearbeiten von Skripts](#).) Wir empfehlen **077** als Einstellung für **UMASK**. Restriktivere Einstellungen könnten zu Berechtigungsproblemen für SPSS Modeler Server führen.

IBM SPSS Modeler Server und das Data Access-Pack

Wenn Sie die Connect for ODBC-Treiber mit IBM® SPSS® Modeler Server verwenden möchten, muss die Connect for ODBC-Umgebung mit *odbc.sh* konfiguriert werden, wenn der SPSS Modeler Server-Prozess startet. Sie erreichen dies, indem Sie das entsprechende IBM® SPSS® Modeler-Startskript in *modelersrv.sh*, *rc.modeler* oder in beiden konfigurieren. (Für weitere Informationen siehe Thema [Bearbeiten von Skripts auf S. 68](#).)

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

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

- ▶ Laden Sie von der Unternehmens-Website unter <http://www.spss.com/drivers/clientCLEM.htm> das relevante komprimierte tar-Archiv für die Plattform, auf der Sie SPSS Modeler Server installiert haben, herunter. Stellen Sie sicher, dass Sie die Treiber für die korrekte Version von SPSS Modeler Server, die Sie installiert haben, herunterladen. Kopieren Sie die Datei an den Speicherort, an dem Sie die ODBC-Treiber installieren möchten (z. B. */usr/spss/odbc*).
- ▶ Extrahieren Sie das Archiv.
- ▶ Führen Sie das *setodbcpath.sh*-Skript aus, das aus dem Archiv extrahiert wird.

- Bearbeiten Sie das Skript *odbc.sh*, um die Definition von ODBCINI unten zu diesem Skript hinzuzufügen und sie zu exportieren, z. B.:

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

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

- Speichern Sie *odbc.sh*.
- (Nur 64-Bit-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 [ "$LD_LIBRARY_PATH_64" = "" ]; then
    LD_LIBRARY_PATH_64=<library_path>
else
    LD_LIBRARY_PATH_64=<library_path>:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
```

wobei *library_path* der gleiche ist wie für die LD_LIBRARY_PATH-Definition, die bereits in dem Skript 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"-Zeichenketten in den neu angehängten `if`- und `export`-Anweisungen durch "LD_LIBRARY_PATH_64" zu ersetzen.

Ihre finale *odbc.sh*-Datei auf einer 64-Bit-SPSS Modeler Server-Installationen könnte also 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.

- 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).
- Speichern Sie die *odbc.ini*-Datei.

- ▶ Konfigurieren Sie 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>
```

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

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

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

- ▶ Speichern Sie *modelersrv.sh*.

So testen Sie die Verbindung:

- ▶ Starten Sie SPSS Modeler Server neu.
- ▶ Verbinden Sie sich von einem Client mit SPSS Modeler Server.
- ▶ Fügen Sie auf dem Client einen Datenbankquellenknoten zum Zeichenbereich hinzu.
- ▶ 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 Sie Fehlermeldungen bei dem Versuch, sich mit einer von Ihnen definierten Datenquelle zu verbinden, erhalten, folgen Sie dem Verfahren zur Fehlerbehebung. [Für weitere Informationen siehe Thema Fehlerbehebung für die Connect for ODBC-Konfiguration auf S. 72.](#)

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

Wenn Sie erfolgreich von SPSS Modeler Server eine Verbindung zur Datenbank aufbauen können, können Sie eine IBM® SPSS® Modeler Solution Publisher Runtime-Installation auf dem gleichen Server konfigurieren, indem Sie das gleiche *odbc.sh*-Skript aus dem Startskript von SPSS Modeler Solution Publisher Runtime referenzieren.

- ▶ Bearbeiten Sie das *modelerrun*-Skript in SPSS Modeler Solution Publisher Runtime, um die folgende Zeile direkt unterhalb der letzten Zeile des Skripts einzufügen:

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

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

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

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

- ▶ Speichern Sie die *modelerrun*-Skriptdatei.

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

Für ODBC ist keine Konfiguration des IBM® SPSS® Modeler Batch-Skripts erforderlich. Das liegt daran, dass Sie sich mit SPSS Modeler Server von SPSS Modeler Batch verbinden, um Streams auszuführen. Stellen Sie sicher, dass die 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:

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

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

Fehlerbehebung für die Connect for 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 Startskript liegen.

- ▶ Überprüfen Sie, ob `$ODBCINI` in *modelersrv.sh* definiert ist, entweder ausdrücklich im Skript selbst oder im *odbc.sh*-Skript, das in *modelersrv.sh* referenziert ist.
- ▶ Stellen Sie in letzterem Fall sicher, dass `ODBCINI` auf den vollständigen Pfad zur *odbc.ini*-Datei verweist, die Sie verwendet haben, um Ihre ODBC-Datenquellen zu definieren.
- ▶ Wenn die Pfadangabe in `ODBCINI` richtig ist, überprüfen Sie den `$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 dem Punkt, wo Sie `ODBCINI` definieren:

```
echo $ODBCINI
```

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

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


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

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

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

Wenn der korrekte Wert bei ausgeführter *modelersrv.sh* in *stdout* wiederholt wird, sollten Sie die Datenquellennamen im Datenbankquellenknoten sehen können, wenn Sie 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 SPSS Modeler Serverdaemonumgebung 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 Skript aus. Der Wert des Bibliothekspfads wird im Terminal wiederholt, wenn Sie das Skript das nächste Mal ausführen.

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

```
echo $<library_path_variable>
```

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

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 SPSS Modeler Server ausführen, überschreibt `$LD_LIBRARY_PATH_64` `$LD_LIBRARY_PATH`, wenn es 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 es das *lib*-Unterverzeichnis Ihrer ODBC-Installation enthält, und exportieren Sie die Definition.

Datenquellennamen 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 Datenquellennamen (DSN), wie er im [ODBC Data Sources]-Abschnitt oben in der Datei mit der Zeichenkette 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 Fehlermeldung angezeigt, wenn Sie eine Verbindung unter Verwendung des DSN von IBM® SPSS® Modeler herstellen möchten. Es folgt ein Beispiel einer *falschen* Spezifikation:

```
[ODBC-Datenquellen]
Oracle=Oracle Wire Protocol

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

Sie müssen eine der beiden Zeichenketten fett markieren, 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-Datenquellen]
Oracle=Oracle Wire Protocol

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

- ▶ Überprüfen Sie, dass das vom Treiberparameter spezifizierte Shared Object vorhanden ist.
- ▶ Korrigieren Sie den Pfad zum Shared Object, falls dieser falsch ist.
- ▶ Wenn der Treiberparameter im Format

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

spezifiziert ist, bedeutet dies, dass Sie Ihre ODBC-bezogenen Skripte nicht initialisiert haben. Führen Sie das *setodbcpath.sh*-Skript aus, das mit den Treibern installiert wurde. [Für weitere Informationen siehe Thema IBM SPSS Modeler Server und das Data Access-Pack auf S. 69.](#)

Wenn Sie dieses Skript ausgeführt haben, sollten Sie sehen, dass die Zeichenkette "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 Connect for 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 Startskript festlegen.

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

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

```
sh
cd ODBC DIR
. odbcs.sh
./bin/ivtestlib MFor815
```

Hierbei wird `ODBC DIR` durch den Pfad zum Installationsverzeichnis von Connect ODBC ersetzt.

- Lesen die Sie Nachricht, um festzustellen, ob ein Fehler vorliegt. Beispiel:

```
Load of MFor815.so failed: ld.so.1: bin/ivtestlib: fatal: libclntsh.so: open failed: No such file or directory
```

Diese Nachricht gibt an, dass die Oracle-Client-Bibliothek `libclntsh.so` nicht vorhanden ist oder sich nicht unter dem angegebenen Pfad befindet (z. B. nicht unter `LD_LIBRARY_PATH` unter Solaris).

- 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
```

Hierbei wird `/bigdisk/oracle/product/8.1.6/lib` durch den Pfad `libclntsh.so` ersetzt.

`LD_LIBRARY_PATH` ist die Variable für den Bibliothekspfad für Ihr Betriebssystem.

Beachten Sie: Wenn Sie die 64-Bit-Version von SPSS Modeler unter Linux oder Solaris 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
```

- Lesen Sie die Nachricht, um festzustellen, ob der Treiber nun geladen werden kann. Beispiel:

```
Load of MFor815.so successful, qehandle is 0xFF3A1BE4
```

Diese Nachricht gibt an, dass die Oracle-Client-Bibliothek geladen werden kann.

- Berichtigen Sie den Bibliothekspfad im Startskript von SPSS Modeler.

- ▶ Starten Sie SPSS Modeler Server mit dem Startskript 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 B-2

Bibliothekspfade, nach Betriebssystem geordnet

Betriebssystem	Name der Bibliothekspfadvariablen
AIX	LIBPATH
HP-UX	SHLIB_PATH
Solaris oder Linux 32-Bit	LD_LIBRARY_PATH
Solaris oder Linux 64-Bit	LD_LIBRARY_PATH_64

Konfigurieren eines Treibers für den Enterprise-Ansichts-Knoten

Mit dem Enterprise-Ansichts-Knoten können Sie eine Verbindung zwischen einer IBM® SPSS® Modeler-Sitzung und einer IBM SPSS Collaboration and Deployment Services Enterprise View in einem freigegebenen IBM® SPSS® Collaboration and Deployment Services Repository herstellen und aufrechterhalten. [Für weitere Informationen siehe Thema Speichern und Bereitstellen von IBM SPSS Collaboration and Deployment Services Repository-Objekten in Kapitel 9 in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)

Um den Enterprise-Ansichts-Knoten mit IBM® SPSS® Modeler Server verwenden zu können, muss ein IBM® SPSS® Collaboration and Deployment Services Enterprise View Driver installiert und konfiguriert sein. Details zur Installation des Treibers erhalten Sie bei Ihrem lokalen Administrator. Nach der Installation muss ein Verweis auf *pev.sh* in das Startskript aufgenommen werden – entweder in *modelersrv.sh*, *rc.modeler* oder in beiden –, das folgendes Format aufweist:

```
./usr/odbc/pev.sh
```

Ausführen unter UNIX als Non-Root-Prozess

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 SPSS Modeler Server geht davon aus, dass der Serverdaemonprozess 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. Es ist dennoch möglich, den Daemonprozess mit einem Nicht-Root-Konto auszuführen, indem Sie die nachfolgenden Anweisungen befolgen.

Ausführen als Nicht-Root. Ohne Root-Berechtigung authentifiziert SPSS Modeler Benutzer mit einer privaten Passwortdatenbank, die sich von der UNIX-Passwortdatenbank unterscheidet, und startet alle Benutzersitzungen mit demselben UNIX-Benutzerkonto. Dies bedeutet, dass alle Dateidaten, die von SPSS Modeler gelesen oder geschrieben werden, allen SPSS Modeler-Benutzern zugänglich sind. Dies gilt nicht für den Zugriff auf Datenbankdaten, da sich Benutzer unabhängig hiervon bei allen den von ihnen verwendeten Datenbank-Datenquellen authentifizieren müssen.

Konfigurieren von IBM SPSS Modeler Server als Nicht-Root-Prozess

Um IBM® SPSS® Modeler Server als Nicht-Root-Prozess auszuführen, werden die echten und effektiven Benutzer-IDs des Serverdaemonprozesses auf ein Konto Ihrer Wahl gesetzt. Sie müssen sich dennoch als Root anmelden oder als “su to root”, um den Server zu starten, zu stoppen und allgemeine Verwaltungsaufgaben auszuführen.

So konfigurieren Sie SPSS Modeler Server für die 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 verwenden wir *modelerusers*.
2. Erstellen Sie das Benutzerkonto, auf dem SPSS Modeler Server ausgeführt werden soll. Dieses Konto dient zur ausschließlichen Verwendung für den SPSS Modeler Server-Daemonprozess. Sie können dieses Konto also nicht verwenden, um den Server zu starten, zu stoppen oder allgemeine IBM® SPSS® Modeler-Verwaltungsaufgaben durchzuführen. Für dieses Beispiel verwenden wir den Namen *modelerserv*.

Beachten Sie beim Erstellen des Kontos Folgendes:

- Das Konto benötigt keine Anmeldefunktionen.

- Die Hauptgruppe sollte die soeben erstellte Gruppe `<modelerusers>` sein.
 - Das Basisverzeichnis kann das Installationsverzeichnis von 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 Startskripts, um SPSS Modeler Server mit dem neu erstellten Konto zu starten. Öffnen Sie das entsprechende Startskript in einem Texteditor. [Für weitere Informationen siehe Thema Skripts in Anhang B auf S. 67.](#)
 - ▶ Ändern Sie im passenden Skript die Zeile mit der Variablen `MODELERUSER`, sodass diese auf den Namen des zum Ausführen von SPSS Modeler Server erstellten Benutzerkontos verweist. Beispiel:


```
MODELERUSER=modelerserv
```
 - ▶ 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 Passwortdatenbank einzustellen, indem Sie die folgende Zeile anhängen:


```
authentication_methods, "pasw_modeler"
```
 5. Anschließend erstellen Sie eine private Passwortdatenbank, die in der Datei `config/passwords.cfg` gespeichert wird. Die Passwortdatei definiert die Kombinationen aus Benutzername/Passwort, die für die Anmeldung bei SPSS Modeler zulässig sind. *Anmerkung:* Diese Daten sind nur für SPSS Modeler bestimmt und stehen in keinerlei Zusammenhang mit den Benutzernamen und Passwörtern, die für die Anmeldung bei UNIX verwendet werden. Sie können dieselben Benutzernamen verwenden, wenn dies praktikabel ist, die Verwendung derselben Passwörter ist jedoch nicht zulässig.

Zum Erstellen der Passwortdatei benötigen Sie das Dienstprogramm für Passwörter (`pwutil`), das sich im Verzeichnis `bin` der SPSS Modeler Server-Installation befindet. Die Synopse dieses Programms lautet:

```
pwutil [ benutzername [ passwort ] ]
```

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

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

Das Definieren eines einzelnen Benutzernamens ist in den meisten Fällen ausreichend, wenn sich alle Benutzer mit demselben Namen und Passwort anmelden. 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
```

Anmerkung: Mit einem einzelnen > wird der Inhalt von *passwords.cfg* überschrieben, wobei auch die zuvor festgelegten Benutzer überschrieben werden. Beachten Sie, dass unabhängig hiervon alle Benutzer dasselbe UNIX-Konto verwenden.

- Ändern Sie die Besitzerrechte für das 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 .
```

- Erstellen Sie Unterverzeichnisse im Datenverzeichnis für Ihre 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 IBM SPSS Modeler Server unter AIX als Non-Root-Benutzer

Beim Start von 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 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
```

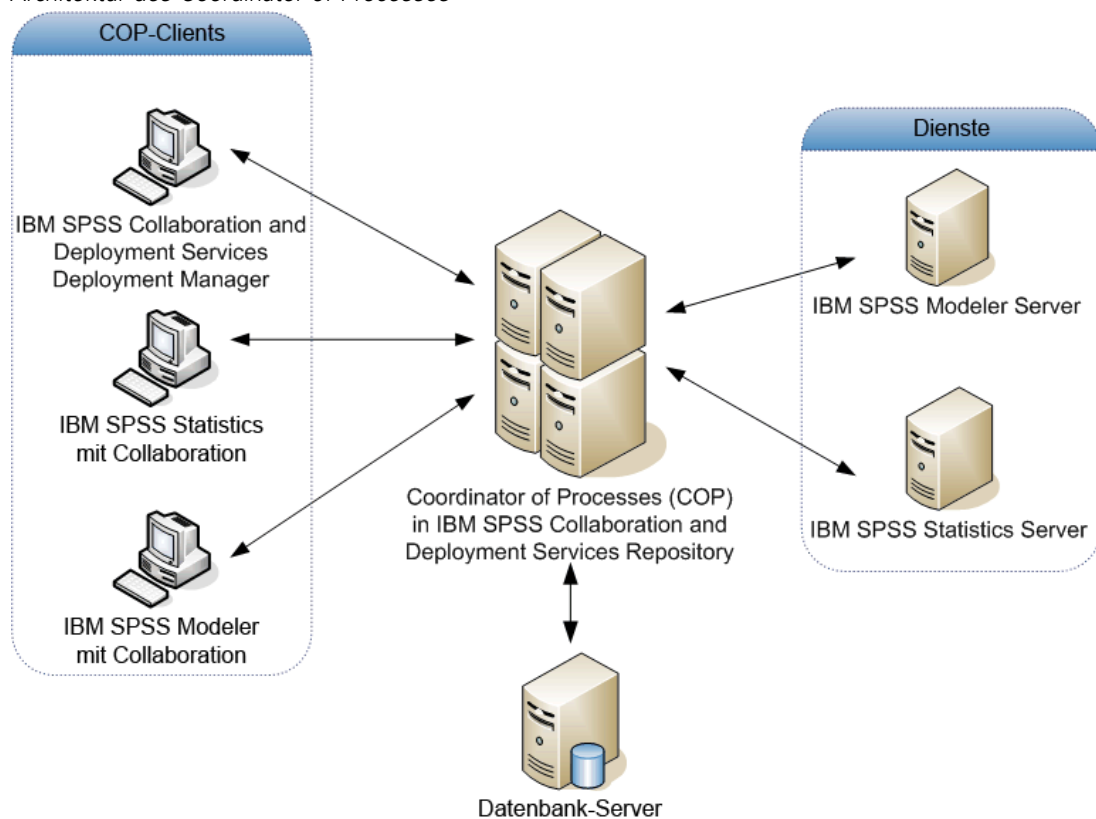
Starten Sie SPSS Modeler Server nach dem Anlegen der Verknüpfung neu.

Lastenausgleich mit Server-Clustern

Mit IBM® SPSS® Collaboration and Deployment Services kann das Plugin Coordinator of Processes verwendet werden, um Dienste im Netzwerk 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 Dienste, beispielsweise IBM® SPSS® Statistics Server oder IBM® SPSS® Modeler Server registrieren sich beim Start beim Coordinator of Processes und senden in regelmäßigen Abständen aktualisierte Statusmeldungen. Die Dienste können außerdem alle notwendigen Konfigurationsdateien in IBM® SPSS® Collaboration and Deployment Services Repository speichern und bei der Initialisierung abrufen.

Abbildung D-1
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 Netzwerk 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's Guide*, das auf Anfrage bei IBM Corp. zur Verfügung gestellt wird.

Wenn Sie in 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. [Für weitere Informationen siehe Thema Verbindung mit IBM SPSS Modeler Server in Kapitel 3 in IBM SPSS Modeler 14.2- Benutzerhandbuch.](#)

Hinweise

This information was developed for products and services offered worldwide.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

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

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

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

Der folgende Absatz gilt nicht für Großbritannien oder andere Länder, in denen derartige Bestimmungen nicht mit dem dort geltenden Recht vereinbar sind. SPSS INC., AN IBM COMPANY, ÜBERNIMMT FÜR DIE VORLIEGENDE DOKUMENTATION KEINERLEI GEWÄHRLEISTUNG IRGENDWELCHER ART, WEDER AUSDRÜCKLICH NOCH STILLSCHWEIGEND, EINSCHLIESSLICH (JEDOCH NICHT DARAUF BEGRENZT) DER STILLSCHWEIGENDEN GEWÄHRLEISTUNGEN IN BEZUG AUF DIE NICHTVERLETZUNG VON RECHTEN DRITTER, AUF HANDELSÜBLICHKEIT ODER DIE EIGNUNG FÜR EINEN BESTIMMTEN ZWECK. Einige Staaten lassen bei bestimmten Transaktionen keine Ausschlussklauseln ausdrücklicher oder stillschweigender Gewährleistungen zu, sodass diese Erklärung möglicherweise nicht auf Sie zutrifft.

Diese Informationen können technische Ungenauigkeiten oder typografische Fehler enthalten. An den hierin enthaltenen Informationen werden in regelmäßigen Abständen Änderungen vorgenommen, die in spätere Ausgaben der Publikation eingearbeitet werden. SPSS Inc. kann jederzeit ohne Vorankündigung Verbesserungen und/oder Veränderungen an den in dieser Publikation beschriebenen Produkten und/oder Programmen vornehmen.

Alle in diesen Ausführungen enthaltenen Verweise auf Websites, die nicht zu SPSS bzw. IBM gehören, dienen lediglich der Information. Die Nennung bedeutet nicht, dass SPSS bzw. IBM den Inhalt dieser Websites unterstützen. Das Material auf diesen Websites ist kein Bestandteil des Materials für dieses SPSS Inc.-Produkt. Sie verwenden diese Websites auf eigenes Risiko.

Wenn Sie Informationen an IBM bzw. SPSS senden, räumen Sie IBM und SPSS das nicht ausschließliche Recht ein, die Informationen in jeglicher Form zu verwenden bzw. weiterzugeben, die dem Unternehmen geeignet erscheint, ohne dass ihm daraus Verbindlichkeiten Ihnen gegenüber entstehen.

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

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

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

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Diese Informationen enthalten Beispiele für Daten und Berichte, die in alltäglichen Betriebsabläufen verwendet werden. Um sie möglichst umfassend darzulegen, enthalten die Beispiele Namen von Einzelpersonen, Unternehmen, Marken und Produkten. Alle diese Namen sind frei erfunden und jegliche Ähnlichkeit mit Namen und Adressen, die von einem tatsächlichen Handelsunternehmen verwendet werden, ist rein zufällig.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Marken

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

SPSS ist eine Marke von SPSS Inc., an IBM Company, die in vielen Ländern weltweit eingetragen sind.

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

IT Infrastructure Library ist eine eingetragene Marke der Central Computer and Telecommunications Agency, die nun zum Office of Government Commerce gehört.

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 und 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.

ITIL ist eine eingetragene Marke und eine eingetragene Gemeinschaftsmarke des Office of Government Commerce und ist beim U.S. Patent and Trademark Office eingetragen.

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

Cell Broadband Engine ist eine Marke von Sony Computer Entertainment, Inc. in den USA und/oder anderen Ländern und wird im Rahmen einer Lizenz dieses Unternehmens verwendet.

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

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

Weitere Produkt- oder Servicenamen können Marken von IBM, SPSS oder anderen Unternehmen sein.

- 64-Bit-Betriebssysteme, 9
- Administratorzugriff
 - für IBM SPSS Modeler Server, 25
- allow_modelling_memory_override
 - options.cfg*, Datei, 29
- Anmeldung bei IBM SPSS Modeler Server, 14
- Anwendungsbeispiele, 3
- Arbeitsspeicher, 11
- Architektur
 - Komponenten, 6
- Authentifizierung, 19
- auto.sh* (UNIX)
 - Speicherort von, 67
- Automatischer Serverstart
 - Konfiguration unter UNIX, 68
- Beispiele
 - Anwendungshandbuch, 3
 - Übersicht, 4
- Benutzer-ID
 - IBM SPSS Modeler Server, 14
- Benutzerauthentifizierung, 19
- Benutzerkonten
 - IBM SPSS Modeler Server, 19
- Betriebssysteme
 - 64-Bit, 9
- cache_compression
 - options.cfg*, Datei, 29
- Cache-Komprimierung, 29
- Caching, In-Database, 35
- chemsrv.sh* (UNIX)
 - Speicherort von, 67
- CLEM-Ausdrücke
 - SQL-Erzeugung, 61
- Connect for ODBC
 - Konfiguration unter UNIX, 69
- Coordinator of Processes, 17
 - Lastenausgleich, 80
 - Server-Cluster, 80
- Coordinator of Processes, Konfiguration
 - für IBM SPSS Modeler Server, 32
- COP, 17
 - Lastenausgleich, 80
 - Server-Cluster, 80
- cop_enabled
 - options.cfg*, Datei, 32
- cop_host
 - options.cfg*, Datei, 32
- COP-Konfiguration
 - für IBM SPSS Modeler Server, 32
- cop_password
 - options.cfg*, Datei, 32
- cop_port_number
 - options.cfg*, Datei, 32
- cop_service_description
 - options.cfg*, Datei, 32
- cop_service_host
 - options.cfg*, Datei, 32
- cop_service_name
 - options.cfg*, Datei, 32
- cop_service_weight
 - options.cfg*, Datei, 32
- cop_update_interval
 - options.cfg*, Datei, 32
- cop_user_name
 - options.cfg*, Datei, 32
- Data Access-Pack
 - Connect for ODBC, Konfiguration unter UNIX, 69
 - Fehlerbehebung für Connect for ODBC unter UNIX, 72
 - Konfigurieren von UNIX, 69
 - und UNIX-Bibliothekspfade, 76
- data_file_path
 - options.cfg*, Datei, 28
- data_files_restricted
 - options.cfg*, Datei, 28
- Datei *messages.log*, 43
- Dateiberechtigungen
 - Konfiguration unter UNIX, 69
 - unter IBM SPSS Modeler Server, 19
- Dateinamen
 - UNIX, 13
 - Windows, 13
- Datenbank-Caching
 - controlling from *options.cfg*, 35
 - SQL-Erzeugung, 57
- Datenbanken
 - Zugriff, 11
- Datenbankserver, 48
- Datenbankverbindungen
 - Schließen, 35
- Datendateien
 - IBM SPSS Statistics, 13
 - Importieren und Exportieren, 13
- Datenzugriff, 11
- DB2
 - SQL-Optimierung, 50–51
- Dokumentation, 3
- Domänenname (Windows)
 - IBM SPSS Modeler Server, 14
- Einzelanmeldung, 16
- Ergebnisse
 - Datensatzreihenfolge, 20
 - Runden von, 20

- Unterschiede zwischen Client und Server, 20
- Fehler bei der Stream-Ausführung, 31
- Festplatten, 11
- Firewall-Einstellungen
 - options.cfg*, Datei, 31
- Hardwareempfehlungen
 - für IBM SPSS Modeler Server, 8
- Hinzufügen von IBM SPSS Modeler Server-Verbindungen, 16–17
- Hostname
 - IBM SPSS Modeler Server, 14, 16
- IBM SPSS Modeler, 1
 - Dokumentation, 3
- IBM SPSS Modeler Administration Console, 24
 - Administratorzugriff, 25
- IBM SPSS Modeler Server
 - Administratorzugriff, 25
 - Andere Ergebnisse als der Client, 20
 - Benutzer-ID, 14
 - Benutzerauthentifizierung, 19
 - Benutzerkonten, 19
 - Coordinator of Processes, Konfiguration, 32
 - COP-Konfiguration, 32
 - Dateierstellung, 19
 - Domänenname (Windows), 14
 - Hostname, 14, 16
 - Informationen für Endbenutzer, 18
 - Konfigurationsoptionen, 27
 - nicht reagierende Prozesse, 22
 - password, 14
 - Portnummer, 14, 16, 27
 - Protokolldateien, 27
 - Serverprozesse, 34
 - Temp-Verzeichnis, 28
 - Überwachung, 34
 - Verwaltung, 24
 - Verwaltungsoptionen, 23
- IBM SPSS Statistics Data Access-Technologie, 11
- IBM SPSS Statistics Datendateien
 - Importieren und Exportieren, 13
- IBM SPSS Statistics-Lizenzspeicherort, 36
- IBM SPSS Text Analytics, 2
- In-Database Mining, 48
- In-Database-Caching, 35
- io_buffer_size*
 - options.cfg*, Datei, 29
- Kernel-Limits in UNIX, 22
- Knoten
 - Unterstützung von SQL-Erzeugung, 58
- Knoten-Caching
 - Schreiben in Datenbank, 57
 - SQL-Erzeugung, 57
- Konfigurationsoptionen
 - Anmeldeversuche, 27
 - Automatische SQL-Erzeugung, 31
 - Coordinator of Processes, 32
 - COP, 32
 - Leistung und Optimierung, 29
 - Parallele Verarbeitung, 29
 - Portnummer, 27
 - Protokollierung, 27
 - Speicherverwaltung, 29
 - SQL-Zeichenkettenlänge, 31
 - SSL-Datenverschlüsselung, 32
 - Stream-Neuerstellung, 29
 - Temp-Verzeichnis, 28
 - Übersicht, 27
 - Verbindungen und Sitzungen, 27
 - von IBM SPSS Modeler Server, 23
 - Zugriff auf Datendateien, 28
- LDAP, 42
 - Schützen, 42
- log_database_access*
 - options.cfg*, Datei, 27
- log_database_errors*
 - options.cfg*, Datei, 27
- log_directory*
 - options.cfg*, Datei, 27
- log_file_access*
 - options.cfg*, Datei, 27
- log_file_errors*
 - options.cfg*, Datei, 27
- Marken, 84
- max_file_size*
 - options.cfg*, Datei, 28
- max_login_attempts*
 - options.cfg*, Datei, 27
- max_parallelism*
 - options.cfg*, Datei, 29
- max_sessions*
 - options.cfg*, Datei, 27
- max_sql_string_length*
 - options.cfg*, Datei, 31
- Mehrere Streams, Ausführung, 31
- Meldungen
 - Anzeigen von erzeugter SQL, 48, 54
- memory_usage*
 - options.cfg*, Datei, 29
- Microsoft SQL Server
 - SQL-Optimierung, 50–51
- Modell-Nuggets
 - SQL anzeigen für, 56
- Modellierung
 - Speicherverwaltung, 29
- modelling_memory_limit_percentage*
 - options.cfg*, Datei, 29

-
- ODBC-Datenquellen
 - Connect for ODBC- und UNIX-Skripts, 69
 - und UNIX, 69
 - Operatoren
 - SQL-Erzeugung, 61
 - Optimierung, 45
 - SQL-Erzeugung, 50–51, 53
 - Option “cache_connection”, 35
 - options.cfg*, Datei, 34
 - Oracle
 - SQL-Optimierung, 50–51, 65

 - Parallele Verarbeitung
 - Aktivieren, 45
 - Steuerung, 29
 - password
 - IBM SPSS Modeler Server, 14
 - performance
 - von IBM SPSS Modeler Server, 45
 - Pfade, 13
 - port_number
 - options.cfg*, Datei, 27
 - Porteinstellungen
 - options.cfg*, Datei, 29
 - Portnummer
 - IBM SPSS Modeler Server, 14, 16, 27
 - program_file_path
 - options.cfg*, Datei, 28
 - program_files_restricted
 - options.cfg*, Datei, 28
 - Protokolldateien
 - Anzeigen von erzeugter SQL, 48, 54
 - für IBM SPSS Modeler Server, 27, 43
 - Prozesse, nicht reagierend, 22
 - Prozessoren, 8
 - mehrfach, 29
 - Pushbacks, 45, 50–51, 53, 58
 - CLEM-Ausdrücke, 61

 - RAM, 11
 - rc.modeler* (UNIX)
 - Speicherort von, 67
 - Rechtliche Hinweise, 82

 - Schützen
 - LDAP, 42
 - Secure Sockets Layer, 36
 - Server
 - Anmeldung, 14
 - Hinzufügen von Verbindungen, 16
 - Suchen nach COP für Server, 17
 - Serverport, Einstellungen
 - options.cfg*, Datei, 29
 - Sicherheit
 - Dateierstellung, 19
 - Konfigurieren der Dateierstellung unter UNIX, 69
 - SSL, 36
 - Speicherplatz, 10
 - berechnen, 10
 - Speicherverwaltung
 - Verwaltungsoptionen, 29
 - SPSS Modeler Server, 1
 - SQL
 - Abfragen, 64
 - Anzeige für Modell-Nuggets, 56
 - Doppelte Spaltennamen, 64
 - Optimierung von Oracle, 65
 - Vorschau für erzeugte, 54
 - SQL Server
 - SQL-Optimierung, 50–51
 - SQL-Erzeugung, 45, 50–51, 58
 - Aktivieren, 53
 - Aktivieren für IBM SPSS Modeler Server, 31
 - Anzeige für Modell-Nuggets, 56
 - CLEM-Ausdrücke, 56, 61
 - Protokollierung, 48, 54
 - Stream-Neuerstellung, 56
 - Tipps, 56
 - Vorschau, 48, 54
 - sql_generation_enabled
 - options.cfg*, Datei, 31
 - SQL-Pushback. *Siehe auch* SQL-Erzeugung, 50
 - @SQLFN-Funktion, 63
 - SSL, 36
 - Kommunikation schützen, 37
 - Übersicht, 37
 - ssl_certificate_file
 - options.cfg*, Datei, 32
 - SSL-Datenverschlüsselung
 - Aktivieren für IBM SPSS Modeler Server, 32
 - ssl_enabled
 - options.cfg*, Datei, 32
 - ssl_private_key_file
 - options.cfg*, Datei, 32
 - ssl_private_key_password
 - options.cfg*, Datei, 32
 - Starten von IBM SPSS Modeler Server
 - unter UNIX, 21
 - unter Windows, 21
 - Status
 - von IBM SPSS Modeler Server unter UNIX, 21
 - von IBM SPSS Modeler Server unter Windows, 21
 - Stoppen von IBM SPSS Modeler Server
 - unter UNIX, 21
 - unter Windows, 21
 - Stream-Neuerstellung, 56
 - Aktivieren, 45
 - stream_rewriting_enabled
 - options.cfg*, Datei, 29
 - Stufen der
 - Datenbankunterstützung, 11
 - Stufen, Datenbankunterstützung, 11
 - Suchen nach COP für Verbindungen, 17

temp_directory

options.cfg, Datei, 28

Temp-Verzeichnis

für IBM SPSS Modeler Server, 28

Temporäre Dateien, 10

Berechtigungen (IBM SPSS Modeler Server), 19

UNC-Dateinamen, 13

UNIX

Benutzerauthentifizierung, 19

Bibliothekspfade, 76

Konfigurieren der Dateiberechtigungen, 69

UNIX-Kernel-Limits, 22

UNIX-Skripts

auto.sh, 67

bearbeiten, 68

modelersrv.sh, 67

rc.modeler, 67

URL-Präfix, 41

Verbindungen

mit IBM SPSS Modeler Server, 14, 16–17

Server-Cluster, 17

Verschlüsselung

SSL, 36

Verwaltung

des Speicherorts der IBM SPSS Statistics-Lizenz, 36

von IBM SPSS Modeler Server, 23

Violette Knoten, 45

SQL-Optimierung, 52

Vorschau

SQL-Erzeugung, 54

Zombie-Prozesse, IBM SPSS Modeler Server, 22