

Administration et guide des
performances de IBM SPSS Modeler
Server 14.2



Remarque : avant d'utiliser ces informations et le produit dont elles traitent, consultez les informations générales dans la rubrique Avis sur p. .

Ce document contient des informations sur la propriété de SPSS Inc., an IBM Company. Il est fourni sous accord de licence et est protégé par la loi sur le copyright. Les informations contenues dans cette publication n'incluent pas les garanties de produit et aucune des provisions de ce manuel ne devra être interprétée comme tel.

Lorsque vous envoyez des informations à IBM ou SPSS, vous octroyez à IBM et SPSS le droit non exclusif d'utiliser ou de distribuer ces informations de la façon dont IBM ou SPSS le juge approprié sans aucune obligation envers vous.

© **Copyright IBM Corporation 1994, 2011..**

Préface

IBM® SPSS® Modeler est le puissant utilitaire de Data mining de IBM Corp.. SPSS Modeler aide les entreprises et les organismes à améliorer leurs relations avec les clients et les citoyens grâce à une compréhension approfondie des données. A l'aide des connaissances plus précises obtenues par le biais de SPSS Modeler, les entreprises et les organismes peuvent conserver les clients rentables, identifier les opportunités de vente croisée, attirer de nouveaux clients, détecter les éventuelles fraudes, réduire les risques et améliorer les services gouvernementaux.

L'interface visuelle de SPSS Modeler met à contribution les compétences professionnelles de l'utilisateur, ce qui permet d'obtenir des modèles prédictifs plus efficaces et de trouver des solutions plus rapidement. SPSS Modeler dispose de nombreuses techniques de modélisation, telles que les algorithmes de prévision, de classification, de segmentation et de détection d'association. Une fois les modèles créés, l'utilisateur peut utiliser IBM® SPSS® Modeler Solution Publisher pour les remettre aux responsables, où qu'ils se trouvent dans l'entreprise, ou pour les transférer vers une base de données.

A propos de IBM Business Analytics

Le logiciel IBM Business Analytics fournit des informations complètes, cohérentes et précises que les preneurs de décision utilisent avec confiance pour améliorer la performance du marché. Un portefeuille étendu d'outils de [business intelligence](#), d'[analyses prédictives](#), de [performance financière et de gestion de stratégie](#), et des [applications analytiques](#) offre des connaissances claires, immédiates et applicables pour améliorer l'efficacité actuelle ainsi que la capacité de prévoir les résultats futurs. Combinées avec de riches solutions industrielles, des pratiques éprouvées et des services professionnels, les organisations de toutes tailles peuvent atteindre la productivité la plus élevée, automatiser des décisions en toute tranquillité et fournir de meilleurs résultats.

Dans le cadre de ce portefeuille, le logiciel IBM SPSS Predictive Analytics aide les organisations à prévoir des événements futurs et à agir en conséquence pour mener à de meilleurs résultats. Des clients dans le domaine commercial, gouvernemental et académique à travers le monde font confiance à la technologie IBM SPSS et considèrent qu'elle représente un avantage compétitif pour attirer, retenir et ajouter des clients, tout en réduisant la fraude et en atténuant les risques. En incorporant le logiciel IBM SPSS dans leur opérations quotidiennes, les organisations deviennent des entreprises prédictives – capables de diriger et d'automatiser les décisions pour atteindre les buts qu'ils se sont fixés et obtenir des avantages compétitifs sensibles. Pour informations supplémentaires ou pour joindre un revendeur, visitez le site <http://www.ibm.com/spss>.

Assistance technique

L'assistance technique est à la disposition des clients pour la maintenance des produits. Les clients peuvent contacter l'assistance technique pour obtenir de l'aide concernant l'utilisation des produits IBM Corp. ou l'installation dans l'un des environnements matériels pris en charge. Pour joindre l'assistance technique, consultez le site Web de IBM Corp. à l'adresse <http://www.ibm.com/support>. Lorsque vous contactez l'assistance technique, n'oubliez pas de préparer vos identifiants, le nom de votre société et votre contrat d'assistance.

Contenu

1	<i>A propos de IBM SPSS Modeler</i>	1
	IBM SPSS Modeler Server	1
	IBM SPSS Modeler Options	2
	IBM SPSS Text Analytics	2
	Documentation de IBM SPSS Modeler	2
	Exemples d'application	4
	Dossier Demos	4
2	<i>Recommandations relatives à l'architecture et au matériel</i>	6
	IBM SPSS Modeler Architecture	6
	Description de l'architecture	6
	Recommandations matérielles	8
	Exigences en termes d'espace disque temporaire et de mémoire RAM	9
	Accès aux données	11
	Référencement des fichiers de données	13
	Importation des fichiers de données IBM SPSS Statistics	13
	Instructions d'installation	13
3	<i>Support de IBM SPSS Modeler</i>	14
	Connexion au IBM SPSS Modeler Server	14
	Ajout et modification d'une connexion à IBM SPSS Modeler Server	16
	Recherche de serveurs dans IBM SPSS Collaboration and Deployment Services.	17
	Systèmes de données et de fichiers	18
	Authentification de l'utilisateur	19
	Création de fichiers	19
	Différences au niveau des résultats	19
4	<i>Administration de IBM SPSS Modeler Server</i>	21
	Démarrage et arrêt de IBM SPSS Modeler Server	21
	Pour démarrer ou arrêter le serveur sous Windows, ou vérifier son état	21
	Pour démarrer ou arrêter le serveur sous UNIX, ou vérifier son état.	21
	Traitement des processus de serveurs qui ne répondent pas (systèmes UNIX)	23

Administration	23
Utilisation de IBM SPSS Modeler Administration Console	24
Démarrage de Modeler Administration Console	24
Configuration de l'accès avec Modeler Administration Console	25
SPSS Modeler Server Connexions	26
Configuration de SPSS Modeler Server	27
SPSS Modeler Server Surveillance	34
Utilisation du fichier options.cfg	34
Fermeture de connexions inutilisées à une base de données	35
Emplacement de la licence IBM SPSS Statistics	36
Utilisation du protocole SSL pour sécuriser le transfert de données	36
Fonctionnement de SSL	37
Sécurisation des communications Client-Serveur et Serveur-Serveur avec le protocole SSL	37
Configuration du préfixe d'URL	41
Sécurisation de LDAP avec SSL	41
Journal du serveur	42

5 Considérations relatives aux performances 44

Paramètres de performances et d'optimisation du serveur	44
Paramètres de performances et d'optimisation du client	45
Utilisation et optimisation de la base de données	47
Optimisation SQL	48

6 Optimisation SQL 49

Fonctionnement de la génération SQL	50
Exemple de génération SQL	51
Configuration de l'optimisation SQL	52
Prévisualisation du code SQL généré	53
Affichage SQL pour des nuggets de modèle	55
Astuces relatives à l'optimisation de la génération SQL	55
Noeuds prenant en charge la génération SQL	57
Opérateurs et expressions CLEM prenant en charge la génération SQL	60
Utilisation de fonctions SQL dans les expressions CLEM	62
Rédaction de requêtes SQL	63

Annexes

A	Configuration d'Oracle pour les plates-formes UNIX	64
	Configuration d'Oracle pour l'optimisation SQL	64
B	Configuration des scripts de démarrage UNIX	66
	Introduction	66
	Scripts	66
	Démarrage et arrêt automatiques de IBM SPSS Modeler Server	67
	Démarrage et arrêt manuels de IBM SPSS Modeler Server	67
	Édition de scripts	68
	Contrôle des droits d'accès relatifs à la création de fichier	68
	IBM SPSS Modeler Server et Data Access Pack	68
	Dépannage de la configuration Connect for ODBC	71
	Chemins de bibliothèque	75
	Configuration d'un pilote pour le noeud Enterprise View	75
C	Exécution en tant que processus non-racine sous UNIX	76
	Introduction	76
	Configuration de IBM SPSS Modeler Server en tant que processus non-racine	76
D	Équilibrage de charge avec classes de serveur	79
E	Avis	81
	Index	84

A propos de IBM SPSS Modeler

IBM® SPSS® Modeler est un ensemble d'outils de data mining qui vous permet de développer rapidement, grâce à vos compétences professionnelles, des modèles prédictifs et de les déployer dans des applications professionnelles afin de faciliter la prise de décision. Conçu autour d'un modèle confirmé, le modèle CRISP-DM, SPSS Modeler prend en charge l'intégralité du processus de Data mining, des données à l'obtention de meilleurs résultats commerciaux.

SPSS Modeler propose différentes méthodes de modélisation issues des domaines de l'apprentissage automatique, de l'intelligence artificielle et des statistiques. Les méthodes disponibles dans la palette Modélisation vous permettent d'extraire de nouvelles informations de vos données et de développer des modèles prédictifs. Chaque méthode possède ses propres avantages et est donc plus adaptée à certains types de problème spécifiques.

Il est possible d'acquérir SPSS Modeler comme produit autonome ou de l'utiliser en combinaison avec SPSS Modeler Server. Plusieurs autres options sont également disponibles, telles que décrites dans les sections suivantes. Pour plus d'informations, consultez <http://www.ibm.com/software/analytics/spss/products/modeler/>.

IBM SPSS Modeler Server

Grâce à une architecture client/serveur, SPSS Modeler adresse les demandes d'opérations très consommatrices de ressources à un logiciel serveur puissant. Il offre ainsi des performances accrues sur des ensembles de données plus volumineux. D'autres mises à jour ou produits que ceux mentionnés ici peuvent également être disponibles. Pour plus d'informations, consultez <http://www.ibm.com/software/analytics/spss/products/modeler/>.

SPSS Modeler. La version SPSS Modeler intègre toutes les fonctions du produit. Elle est installée et exécutée sur l'ordinateur de bureau de l'utilisateur. Pour obtenir de meilleures performances lors du traitement d'ensembles de données volumineux, vous pouvez l'exécuter en mode local, comme produit autonome, ou en mode réparti, en association avec IBM® SPSS® Modeler Server.

SPSS Modeler Server. SPSS Modeler Server s'exécute en continu en mode d'analyse réparti, conjointement avec une ou plusieurs installations IBM® SPSS® Modeler, obtenant ainsi de meilleures performances lors du traitement des ensembles de données volumineux. En effet, les opérations consommatrices de mémoire s'effectuent sur le serveur : inutile de télécharger les données sur l'ordinateur client. SPSS Modeler Server offre également une prise en charge de l'optimisation SQL et des capacités de modélisation au sein de la base de données, ce qui présente de nouveaux avantages en matière de performances et d'automatisation. Pour lancer les analyses, vous devez disposer d'au moins une installation SPSS Modeler.

IBM SPSS Modeler Options

Vous pouvez acheter et obtenir une licence pour les fonctions et composants suivants à utiliser avec SPSS Modeler. Notez que des produits ou des mises à jour supplémentaires deviennent parfois disponibles. Pour plus d'informations, consultez <http://www.ibm.com/software/analytics/spss/products/modeler/>.

- Accès à SPSS Modeler Server, assurant une meilleure extensibilité et des performances accrues pour les ensembles de données volumineux, ainsi qu'une prise en charge de l'optimisation SQL et des capacités de modélisation au sein de la base de données.
- SPSS Modeler Solution Publisher, pour le scoring en temps réel ou automatique en dehors de l'environnement SPSS Modeler. [Pour plus d'informations, reportez-vous à la section IBM SPSS Modeler Solution Publisher dans le chapitre 2 dans *IBM SPSS Modeler 14.2 Solution Publisher*.](#)
- Adaptateurs permettant le déploiement dans IBM SPSS Collaboration and Deployment Services ou dans l'application client léger IBM SPSS Modeler Advantage. [Pour plus d'informations, reportez-vous à la section Stockage et et déploiement des objets IBM SPSS Collaboration and Deployment Services Repository dans le chapitre 9 dans *Guide de l'utilisateur de IBM SPSS Modeler 14.2*.](#)

IBM SPSS Text Analytics

IBM® SPSS® Text Analytics est un complément totalement intégré pour SPSS Modeler, qui utilise des technologies linguistiques avancées et le traitement du langage naturel pour traiter rapidement un grand nombre de données texte non structurées, extraire et organiser les concepts-clés et grouper ces concepts en catégories. Les concepts extraits et les catégories peuvent ensuite être combinés aux données structurées existantes, telles que les données démographiques, et appliqués à la modélisation grâce à la gamme complète d'outils de Data mining de IBM® SPSS® Modeler, afin de favoriser une prise de décision précise et efficace.

- Le Text Mining offre une modélisation des concepts et des catégories, ainsi qu'un utilitaire interactif où vous pouvez exécuter une exploration avancée des liens textuels et des classes, créer vos propres catégories et affiner les modèles de ressources linguistiques.
- De nombreux formats d'importation sont pris en charge, y compris les blogs et sources Web.
- Enfin, des modèles personnalisés, des bibliothèques et des dictionnaires spécialisés dans des domaines précis, tels que la gestion de la relation client et la génomique, sont fournis.

Remarque : Une licence distincte est requise pour accéder à ce composant. Pour plus d'informations, consultez <http://www.ibm.com/software/analytics/spss/products/modeler/>.

Documentation de IBM SPSS Modeler

Une documentation complète au format d'aide en ligne est disponible dans le menu Aide de SPSS Modeler. Vous y trouverez la documentation de SPSS Modeler, SPSS Modeler Server et de SPSS Modeler Solution Publisher, ainsi que le Guide des applications et d'autres documentations utiles.

La documentation complète de chaque produit au format PDF est disponible dans le dossier *\Documentation* de chaque DVD de produit.

- **Guide de l'utilisateur IBM SPSS Modeler.** Introduction générale à SPSS Modeler : création de flux de données, traitement des valeurs manquantes, création d'expressions CLEM, utilisation des projets et des rapports et regroupement des flux pour le déploiement dans IBM SPSS Collaboration and Deployment Services, des applications prédictives ou IBM SPSS Modeler Advantage.
- **Noeuds de Source, d'exécution et de sortie IBM SPSS Modeler.** Descriptions de tous les noeuds utilisés pour lire, traiter et renvoyer les données de sortie dans différents formats. En pratique, cela signifie tous les noeuds autres que les noeuds de modélisation.
- **IBM SPSS Modeler Noeuds de modélisation.** Description de tous les noeuds utilisés pour créer des modèles de Data mining. IBM® SPSS® Modeler propose différentes méthodes de modélisation issues des domaines de l'apprentissage automatique, de l'intelligence artificielle et des statistiques. [Pour plus d'informations, reportez-vous à la section Description des noeuds de modélisation dans le chapitre 3 dans Noeuds de modélisation IBM SPSS Modeler 14.2.](#)
- **Guide des Algorithmes IBM SPSS Modeler.** Descriptions des fondements mathématiques des méthodes de modélisation utilisées dans SPSS Modeler.
- **Guide des applications IBM SPSS Modeler.** Les exemples de ce guide fournissent des introductions brèves et ciblées aux méthodes et techniques de modélisation. Une version en ligne de ce guide est également disponible dans le menu Aide. [Pour plus d'informations, reportez-vous à la section Exemples d'application dans Guide de l'utilisateur de IBM SPSS Modeler 14.2.](#)
- **Génération de scripts et automatisation IBM SPSS Modeler.** Informations sur l'automatisation du système via la génération de scripts, y compris les propriétés permettant de manipuler les noeuds et les flux.
- **IBM SPSS Modeler Guide de déploiement.** Informations sur l'exécution des scénarios et des flux SPSS Modeler comme étapes des tâches d'exécution sous IBM® SPSS® Collaboration and Deployment Services Deployment Manager.
- **IBM SPSS Modeler CLEF Guide du développeur.** CLEF permet d'intégrer des programmes tiers tels que des programmes de traitement de données ou des algorithmes de modélisation en tant que noeuds dans SPSS Modeler.
- **Guide d'exploration de base de données IBM SPSS Modeler.** Informations sur la manière de tirer parti de la puissance de votre base de données pour améliorer les performances et étendre la gamme des fonctions analytiques via des algorithmes tiers.
- **IBM SPSS Modeler Server et Guide des performances.** Informations sur le mode de configuration et d'administration de IBM® SPSS® Modeler Server.
- **Guide de l'utilisateur de IBM SPSS Modeler Administration Console.** Informations concernant l'installation et l'utilisation de l'interface utilisateur de la console permettant de surveiller et de configurer SPSS Modeler Server. La console est implémentée en tant que plug-in à l'application Deployment Manager.

- **Guide IBM SPSS Modeler Solution Publisher.** SPSS Modeler Solution Publisher est un module complémentaire qui permet aux entreprises de publier des flux destinés à être utilisés en dehors de l'environnement SPSS Modeler.
- **Guide CRISP-DM IBM SPSS Modeler** Guide détaillé sur l'utilisation de la méthodologie CRISP-DM pour le Data mining avec SPSS Modeler

Exemples d'application

Tandis que les outils de Data mining de SPSS Modeler peuvent vous aider à résoudre une grande variété de problèmes commerciaux et organisationnels, les exemples d'application fournissent des introductions brèves et ciblées aux méthodes et aux techniques de modélisation. Les ensembles de données utilisés ici sont beaucoup plus petits que les énormes entrepôts de données gérés par certains Data miners, mais les concepts et les méthodes impliqués doivent pouvoir être adaptés à des applications réelles.

Vous pouvez accéder aux exemples en cliquant Exemples d'application dans le menu Aide de SPSS Modeler. Les fichiers de données et les flux d'échantillons sont installés dans le dossier *Demos*, sous le répertoire d'installation du produit. [Pour plus d'informations, reportez-vous à la section Dossier Demos dans *Guide de l'utilisateur de IBM SPSS Modeler 14.2*.](#)

Exemples de modélisation de bases de données. Consultez les exemples dans le *IBM SPSS ModelerGuide d'exploration de base de données*.

Exemples de génération de scripts. Consultez les exemples dans le *IBM SPSS ModelerGuide de génération de scripts et d'automatisation*.

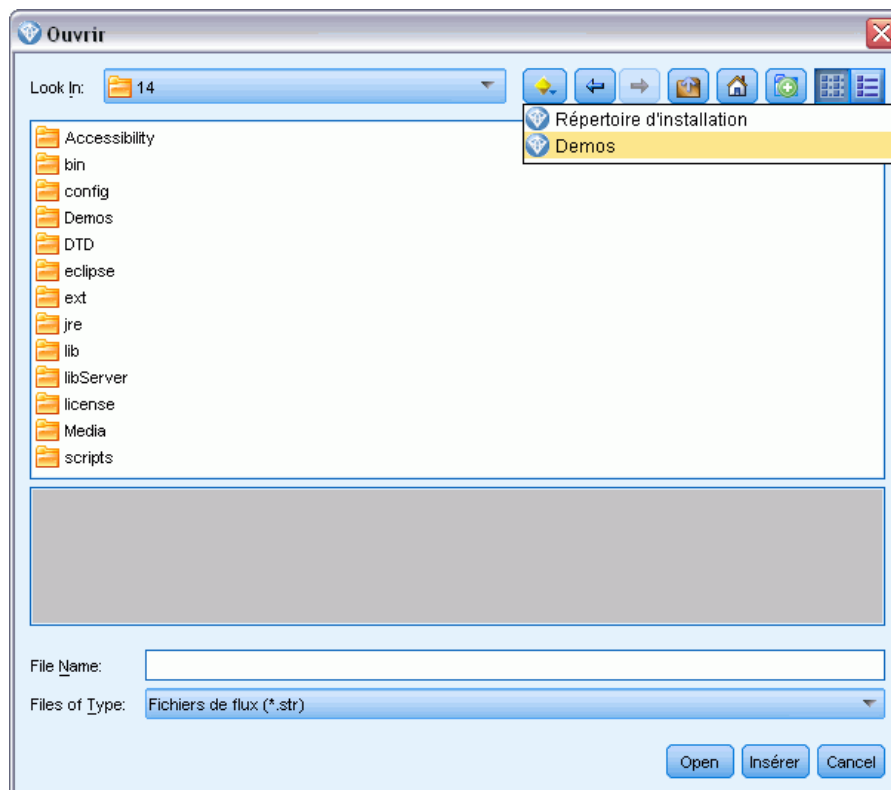
Dossier Demos

Les fichiers de données et les flux d'échantillons utilisés avec les exemples d'application sont installés dans le dossier *Demos*, sous le répertoire d'installation du produit. Ce dossier est également accessible à partir du groupe de programmes sous IBM SPSS Modeler 14.2 dans le

menu Démarrer de Windows, ou en cliquant sur *Demos* dans la liste des répertoires récents de la boîte de dialogue Ouverture de fichier.

Figure 1-1

Sélection du dossier *Demos* dans la liste des répertoires récemment consultés



Recommandations relatives à l'architecture et au matériel

IBM SPSS Modeler Architecture

Cette section décrit l'architecture de IBM® SPSS® Modeler Server, et plus particulièrement le logiciel serveur, le logiciel client et la base de données. Il explique ce qui fait de SPSS Modeler Server une solution capable d'offrir des performances optimales et fournit des recommandations quant au choix d'un matériel adapté pour profiter au mieux de ces performances. Il conclut enfin sur une section consacrée à l'accès aux données, qui décrit l'emplacement d'installation des pilotes ODBC nécessaires.

Description de l'architecture

IBM® SPSS® Modeler Server utilise une architecture répartie à trois niveaux. Les opérations logicielles sont réparties entre les ordinateurs client et serveur. L'installation et l'utilisation de SPSS Modeler Server (plutôt que de IBM® SPSS® Modeler en mode autonome) présentent de nombreux avantages, tout particulièrement lorsque vous traitez des ensembles de données volumineux :

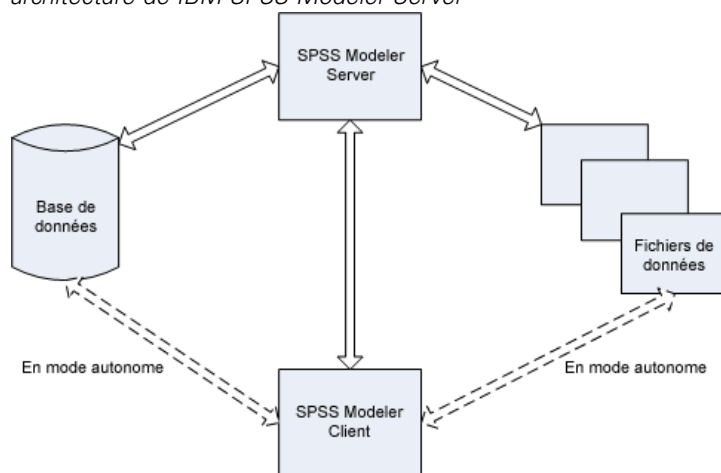
- SPSS Modeler Server peut être exécuté sous UNIX, en plus de Windows, ce qui vous offre davantage de souplesse dans le choix de son emplacement d'installation. Quelle que soit la plate-forme employée, vous pouvez utiliser un ordinateur serveur plus rapide et plus puissant dédié aux seuls traitements de Data mining.
- SPSS Modeler Server est optimisé pour assurer des performances élevées. Lorsqu'il est impossible de répercuter les opérations dans la base de données, SPSS Modeler Server stocke les résultats intermédiaires en tant que fichiers temporaires sur le disque et non dans la mémoire RAM. Etant donné que les serveurs disposent généralement d'une grande quantité d'espace disque disponible, SPSS Modeler Server peut effectuer des opérations de tri, de fusion et d'agrégation sur des ensembles de données très volumineux.
- Grâce à l'architecture client-serveur, vous pouvez centraliser les processus de Data mining de votre entreprise. Cette fonction de centralisation vous permet de formaliser le rôle du Data mining dans vos processus métier.
- Grâce à des outils d'administrateur comme le IBM® SPSS® Collaboration and Deployment Services Deployment Manager (fourni avec SPSS Modeler Server) et le IBM® SPSS® Collaboration and Deployment Services et IBM® SPSS® Collaboration and Deployment Services Repository (vendus séparément), vous pouvez contrôler les processus de Data mining et vérifier que des ressources de calcul appropriées sont disponibles. Vous pouvez automatiser certaines tâches de Data mining, gérer l'accès aux modèles de données et partager les résultats avec les autres utilisateurs de votre entreprise.

Les différents composants de l'architecture répartie de IBM® SPSS® Modeler sont présentés dans le « architecture de IBM SPSS Modeler Server » graphique sur p. 7.

- **SPSS Modeler** Le logiciel client est installé sur l'ordinateur de l'utilisateur final. Il fournit une interface utilisateur et affiche les résultats du Data mining. La version client consiste en une installation complète de SPSS Modeler, mais se distingue par le fait que lorsqu'elle est connectée à SPSS Modeler Server pour une analyse répartie, son moteur d'exécution est inactif. SPSS Modeler s'exécute sur les systèmes d'exploitation Windows uniquement.
- **SPSS Modeler Server.** Le logiciel serveur est installé sur un serveur disposant d'une connexion réseau à la fois avec les SPSS Modeler(s) et la base de données. SPSS Modeler Server s'exécute comme un service (sous Windows) ou un processus démon (sous UNIX), et attend que les clients se connectent. Il gère l'exécution des flux et des scripts créés à l'aide de SPSS Modeler.
- **Serveur de base de données.** Le serveur de base de données peut consister en un entrepôt de données en direct (par exemple, Oracle installé sur un puissant serveur UNIX) ou, pour réduire l'impact sur les autres systèmes opérationnels, un mini-entrepôt sur un serveur local/départemental (SQL Serveur sous Windows, par exemple).

Architecture de IBM SPSS Modeler Server

Figure 2-1
architecture de IBM SPSS Modeler Server



Dans le cas d'une architecture répartie, la plupart des processus s'effectuent sur l'ordinateur serveur. Lorsque l'utilisateur final exécute un flux, SPSS Modeler envoie une description du flux au serveur. Le serveur détermine les opérations qui peuvent être exécutées en langage SQL et crée les requêtes appropriées. Ces requêtes sont exécutées dans la base de données et, pour tout traitement impossible à exprimer en langage SQL, les données obtenues comme résultat sont transmises au serveur. Une fois le traitement terminé, seuls les résultats pertinents sont renvoyés au client.

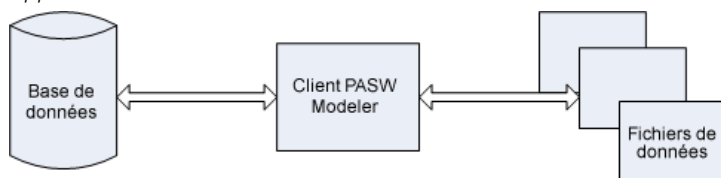
Si nécessaire, SPSS Modeler Server peut exécuter toutes les opérations SPSS Modeler en dehors de la base de données. Il équilibre automatiquement la quantité de mémoire RAM et de mémoire disque utilisée pour stocker les données en vue de leur manipulation. Grâce à ce processus, SPSS Modeler Server est entièrement compatible avec les fichiers plats.

L'équilibrage de la charge est aussi disponible grâce à un groupe de serveurs pour le traitement. La classification est disponible dans IBM SPSS Collaboration and Deployment Services 3.5 par l'intermédiaire du plug-in Coordinator of Processes. [Pour plus d'informations, reportez-vous à la section Equilibrage de charge avec classes de serveur dans l'annexe D sur p. 79.](#) Vous pouvez vous connecter à un serveur ou à un groupe de serveurs gérés dans Coordinator of Processes directement grâce à la boîte de dialogue Connexion au serveur de SPSS Modeler. [Pour plus d'informations, reportez-vous à la section Connexion au IBM SPSS Modeler Server dans le chapitre 3 dans *Guide de l'utilisateur de IBM SPSS Modeler 14.2.*](#)

Client en mode autonome

Vous pouvez également configurer SPSS Modeler tel qu'il s'exécute en tant qu'application de bureau indépendante, comme l'illustre le « Application IBM SPSS Modeler autonome » graphique ci-dessous. [Pour plus d'informations, reportez-vous à la section Support de IBM SPSS Modeler dans le chapitre 3 sur p. 14.](#)

Figure 2-2
Application IBM SPSS Modeler autonome



Recommandations matérielles

Lors de la phase de préparation de votre installation IBM® SPSS® Modeler Server, pensez au matériel que vous utiliserez. SPSS Modeler Server est certes conçu pour être rapide, mais pour profiter au maximum de son efficacité, utilisez du matériel adapté à vos tâches de Data mining. En général, le moyen le plus simple et le moins onéreux pour améliorer les performances consiste à mettre à niveau le matériel existant.

Serveur dédié. Installez SPSS Modeler Server sur un serveur dédié, dont il ne disputera pas les ressources avec d'autres applications, notamment les bases de données SPSS Modeler Server auxquelles il peut se connecter. Les opérations de création de modèles, en particulier, sont très consommatrices de ressources et bénéficient d'une exécution plus efficace lorsqu'elles n'entrent pas en concurrence avec d'autres applications.

Remarque : L'installation de SPSS Modeler Server sur le même ordinateur que la base de données peut permettre de réduire le temps de transfert entre la base de données et le serveur puisqu'elle supprime toute durée de traitement réseau. Toutefois, dans la plupart des cas, il s'avère plus judicieux de placer le serveur et la base de données sur des ordinateurs distincts afin d'éviter toute interférence en termes de ressources. Établissez une connexion rapide entre eux deux pour réduire au maximum le temps de transfert des données.

Processeurs. Le nombre de processeurs installés sur l'ordinateur ne doit pas être inférieur au nombre de tâches simultanées (flux exécutés en même temps) que vous pensez effectuer régulièrement. En général, il est recommandé d'utiliser le plus grand nombre de processeurs possible.

- Une instance unique de SPSS Modeler Server accepte les connexions en provenance de plusieurs clients (utilisateurs) et chaque connexion client peut lancer plusieurs exécutions de flux. A tout instant, plusieurs tâches d'exécution peuvent ainsi être en cours en même temps sur un même serveur.
- En règle générale, il est conseillé d'utiliser un processeur pour un ou deux utilisateurs, deux processeurs pour un maximum de quatre utilisateurs et quatre processeurs pour un maximum de huit utilisateurs. Ajoutez ensuite un processeur par tranche supplémentaire de deux à quatre utilisateurs, en fonction de la diversité des tâches à effectuer.
- Dans la mesure où certains processus peuvent être répercutés dans la base de données via la fonction d'optimisation SQL, il est possible de partager une UC entre plusieurs utilisateurs avec un minimum de pertes de performances.
- La fonction multi-thread permet à une tâche unique d'utiliser plusieurs processeurs. Par conséquent, grâce à l'ajout de processeurs, vous pouvez améliorer les performances du système, et ce même lorsqu'une seule tâche est exécutée à la fois. La fonction multi-thread est généralement utilisée pour la création de modèles C5.0, ainsi que pour certaines opérations de préparation des données (tri, agrégation et fusion).

Plates-formes 64 bits. Si vous prévoyez de traiter ou de créer des modèles sur de grands volumes de données, utilisez Solaris, Windows ou Linux 64 bits comme plate-forme SPSS Modeler Server, et augmentez au maximum la quantité de mémoire RAM de l'ordinateur. En effet, dans le cas d'ensembles de données volumineux, le serveur peut rapidement épuiser la quantité maximale de mémoire allouée par processus par les plates-formes 32 bits ; cela entraîne alors un débordement des données sur le disque et une augmentation sensible du temps d'exécution. L'ajout de mémoire RAM peut être bénéfique pour les implémentations de serveur 64 bits ; un minimum de 8 gigaoctets (Go) est recommandé. La prise en charge 64 bits est assurée pour les plates-formes Solaris, Windows et Linux.

Prévision des besoins futurs. Assurez-vous, dans la mesure du possible, que le matériel serveur est évolutif en termes de mémoire et d'UC ; vous pourrez ainsi l'adapter à une utilisation plus intensive (par exemple, un nombre accru d'utilisateurs simultanés ou des exigences utilisateur supérieures à celles existantes en termes de traitement), ainsi qu'aux futures fonctions multi-thread améliorées de SPSS Modeler Server.

Exigences en termes d'espace disque temporaire et de mémoire RAM

IBM® SPSS® Modeler Server utilise l'espace disque temporaire pour le traitement des grands volumes de données. La quantité d'espace temporaire nécessaire dépend du volume et du type des données traitées, ainsi que du type d'opérations effectuées. Le volume de données est proportionnel au nombre de lignes *et* de colonnes. Plus le nombre de lignes et de colonnes traitées est important, plus vous avez besoin d'espace disque.

Cette section décrit les conditions qui requièrent l'utilisation d'espace disque temporaire et de mémoire RAM supplémentaire ; elle explique également comment évaluer la quantité nécessaire. Toutefois, cette section ne traite pas des exigences en termes d'espace disque temporaire des

processus qui se déroulent dans une base de données, puisque ces exigences sont propres à chaque base de données.

Conditions nécessitant l'utilisation d'espace disque temporaire

La puissante fonction d'optimisation SQL de IBM® SPSS® Modeler Server permet d'effectuer les traitements dans la base de données (et non sur le serveur) lorsque cela s'avère possible. Toutefois, il est impossible d'utiliser l'optimisation SQL si l'une des conditions suivantes se vérifie :

- Les données à traiter sont contenues dans un fichier plat et non dans une base de données.
- La fonction d'optimisation SQL est désactivée.
- L'opération de traitement ne peut pas être optimisée via le code SQL.

Lorsqu'il est impossible d'utiliser l'optimisation SQL, les fonctions CLEM et les noeuds de manipulation des données suivants créent dans l'espace disque temporaire une copie de tout ou partie des données. Si les flux utilisés sur votre site contiennent les fonctions ou commandes de traitement ci-dessous, il se peut que vous deviez réserver de l'espace disque supplémentaire sur le serveur.

- noeud Agréger
- noeud Distinguer
- noeud Discrétiser
- noeud Fusionner lors de l'utilisation de l'option de fusion par clé
- tout noeud de modélisation
- noeud Trier
- Noeud de sortie Table
- fonctions @OFFSET dont la condition de recherche utilise @THIS.
- Toute fonction @ , telle que @MIN, @MAX, et @AVE, dans lequel le paramètre de décalage est calculé.

Calcul de la quantité d'espace disque temporaire nécessaire

En général, IBM® SPSS® Modeler Server doit être capable d'écrire un fichier temporaire qui est au moins *trois fois plus grand* que l'ensemble de données d'origine. Par exemple, si le fichier de données fait 2 Go et que la fonction de génération SQL n'est pas utilisée, SPSS Modeler Server requiert 6 Go d'espace disque pour traiter les données. Etant donné que chaque compte d'utilisateur simultané crée ses propres fichiers temporaires, vous devrez augmenter l'espace disque en conséquence, pour chacun de ces utilisateurs.

Si vous constatez que votre entreprise utilise souvent des fichiers temporaires volumineux, envisagez de recourir à un système de fichiers distinct, créé sur un disque à part, pour les fichiers temporaires de IBM® SPSS® Modeler. Pour des résultats optimaux, vous pouvez utiliser la technologie RAID 0 (ou un ensemble de données partitionnées) qui regroupe plusieurs disques physiques afin d'accélérer les opérations sur disque, l'idéal étant de placer chaque disque du système de fichiers partitionné sur un contrôleur de disque distinct.

Exigences en termes de mémoire RAM

Pour la plupart des traitements qui ne peuvent pas être exécutés dans la base de données, IBM® SPSS® Modeler Server stocke les résultats intermédiaires en tant que fichiers temporaires sur le disque et non dans la mémoire (RAM). Toutefois, dans le cas des noeuds de modélisation, la mémoire RAM est utilisée autant que possible. Les noeuds R. neurones, Kohonen et K-means requièrent une grande quantité de mémoire RAM. Si ces noeuds sont fréquemment utilisés sur votre site, envisagez d'installer davantage de mémoire RAM sur le serveur.

Le nombre d'octets de mémoire RAM nécessaire est généralement calculé comme suit :

$$(\text{number_of_records} * \text{number_of_cells_per_record}) * \text{number_of_bytes_per_cell}$$

où le `number_of_cells_per_record` peut devenir très important en présence de champs nominaux.

Consultez la section concernant la configuration système requise dans le guide d'installation du serveur pour les recommandations en mémoire RAM. Pour quatre utilisateurs simultanés ou plus, il est conseillé d'utiliser encore plus de mémoire RAM. La mémoire doit être partagée entre les tâches simultanées, et doit donc être augmentée en conséquence. En règle générale, l'ajout de mémoire constitue souvent l'un des moyens les plus économiques d'améliorer les performances du système.

Accès aux données

Pour lire ou écrire sur une base de données, vous devez installer et configurer une source de données ODBC pour la base de données appropriée, avec, le cas échéant, des autorisations en lecture et en écriture. Le IBM® SPSS® Data Access Pack comprend un ensemble de pilotes ODBC qui peuvent être utilisés à cette fin. Ces pilotes sont disponibles sur le disque d'installation IBM SPSS Data Access Pack fourni avec cette version. Si vous avez des questions sur la création ou la définition d'autorisations pour les sources de données ODBC, contactez l'administrateur de votre base de données.

Dans IBM® SPSS® Modeler, la prise en charge de la base de données est classée en trois niveaux différents de prise en charge pour l'optimisation et le pushback SQL, en fonction du fournisseur de la base de données. Les différents niveaux de prise en charge sont implémentés au moyen d'un certain nombre de paramètres système qui peuvent être personnalisés pour faire partie du contrat de services SPSS.

Les trois niveaux de prise en charge de la base de données sont :

Table 2-1
niveaux de prise en charge de la base de données

Niveau de prise en charge	Description
Niveau 1	Tout pushback SQL possible est disponible, avec l'optimisation SQL spécifique à la base de données.
Niveau 2	La plupart des pushback SQL possibles sont disponibles, sans optimisation SQL spécifique à la base de données.
Niveau 3	Aucune répercussion SQL ou optimisation : uniquement la lecture des données depuis la base de données et l'écriture des données dans la base de données sont disponibles.

Pilotes ODBC pris en charge

Pour obtenir les informations les plus récentes sur les bases de données et pilotes ODBC pris en charge et testés pour une utilisation avec SPSS Modeler 14.2, consultez les matrices de compatibilité des produits sur le site Web de support technique de l'entreprise (<http://www.ibm.com/support>).

Où installer les pilotes

Vous devez installer et configurer les pilotes ODBC sur chaque ordinateur où le traitement a lieu.

- Si vous exécutez IBM® SPSS® Modeler en mode local (autonome), vous devez installer les pilotes sur l'ordinateur local.
- Si vous exécutez SPSS Modeler en mode distribué sur IBM® SPSS® Modeler Server en mode distant, les pilotes ODBC doivent être installés sur le même ordinateur d'installation que SPSS Modeler Server.
- Si vous devez accéder aux mêmes sources de données provenant de SPSS Modeler et de SPSS Modeler Server, les pilotes ODBC doivent être installés sur les deux ordinateurs.
- Si vous exécutez SPSS Modeler sur Terminal Services, vous devez installer les pilotes ODBC sur le serveur Terminal Services sur lequel vous disposez de SPSS Modeler.
- Si vous utilisez IBM® SPSS® Modeler Solution Publisher Runtime pour exécuter des flux publiés sur un ordinateur distinct, vous devez aussi installer et configurer les pilotes ODBC sur cet ordinateur.

Remarque : Si vous utilisez SPSS Modeler Server sous UNIX pour accéder à une base de données Teradata, vous devez utiliser le gestionnaire de pilote ODBC installé avec le pilote ODBC Teradata. Afin de procéder à ces modifications dans SPSS Modeler Server, veuillez spécifier une valeur pour ODBC_DRIVER_MANAGER_PATH en haut du script modelersrv.sh, à l'endroit indiqué par les commentaires. Cette variable d'environnement doit être définie sur l'emplacement du gestionnaire de pilote ODBC fourni avec le pilote ODBC Teradata (/usr/odbc/lib dans une installation du pilote ODBC Teradata par défaut). Vous devez redémarrer SPSS Modeler Server pour que la modification prenne effet. Pour obtenir plus de détails sur les plateformes de SPSS Modeler Server qui prennent en charge l'accès à Teradata, et sur la version du pilote ODBC Teradata prise en charge, consultez le site Web de support technique de l'entreprise à l'adresse <http://www.ibm.com/support>.

Remarque : Les règles précisées ci-dessus s'appliquent tout particulièrement à l'accès aux données d'une base de données. Les autres types d'opération sur les fichiers, tels que l'ouverture et l'enregistrement de flux, de projets, de modèles, de noeuds, de modèles PMML, de sortie et de fichiers de script, sont toujours effectués sur le client et définis par rapport au système de fichiers de l'ordinateur client. En outre, la commande Définir le répertoire de SPSS Modeler définit le répertoire de travail pour les objets client *locaux* (les flux, par exemple) mais n'a aucune incidence sur le répertoire de travail du serveur.

UNIX. Pour obtenir des informations sur le mode de configuration de SPSS Modeler Server sous UNIX afin d'utiliser la technologie IBM® SPSS® Statistics Data Access, reportez-vous à [Configuration des scripts de démarrage UNIX](#) sur p. 66.

Référencement des fichiers de données

Windows. Si vous stockez les données sur l'ordinateur sur lequel IBM® SPSS® Modeler Server est installé, il est recommandé de fournir le chemin d'accès aux données en vous plaçant côté ordinateur serveur (par exemple, *C:\ServerData\Sales 1998.csv*). Les performances sont plus élevées lorsque vous n'utilisez pas le réseau pour localiser le fichier.

Si les données sont stockées sur un hôte différent, il est recommandé d'utiliser des références de fichier UNC (par exemple, *\\mydataserver\ServerData\Sales 1998.csv*). Notez que les noms UNC fonctionnent uniquement lorsque le chemin contient le nom d'une ressource réseau partagée. L'ordinateur de référence doit disposer d'un accès en lecture pour le fichier indiqué. Si vous basculez fréquemment entre le mode d'analyse réparti et le mode d'analyse local, utilisez des références de fichier UNC puisque celles-ci fonctionnent dans n'importe quel mode.

UNIX. Pour faire référence à des fichiers de données figurant sur un serveur UNIX, indiquez le chemin d'accès complet des fichiers en utilisant des barres obliques normales (par exemple, */public/data/ServerData/Sales 1998.csv*). Évitez d'utiliser des barres obliques inverses dans le répertoire UNIX et dans les noms de fichier des données utilisées avec SPSS Modeler Server. Un fichier texte peut utiliser aussi bien le format UNIX que DOS, puisque tous deux sont gérés automatiquement.

Importation des fichiers de données IBM SPSS Statistics

Si vous exécutez également IBM® SPSS® Statistics Server sur votre site, il se peut que les utilisateurs souhaitent importer ou exporter des données SPSS Statistics lorsqu'ils se trouvent en mode réparti. Gardez à l'esprit que lorsque IBM® SPSS® Modeler est exécuté en mode réparti, il offre une vue du système de fichiers du serveur. Le client SPSS Statistics fonctionne de la même manière. Pour pouvoir effectuer des opérations d'importation et d'exportation entre ces deux applications, les deux clients doivent présenter le même mode d'exécution. Si tel n'est pas le cas, ils offrent une vue différente des systèmes de fichiers et ne peuvent pas partager de fichiers. Les noeuds SPSS Statistics de IBM® SPSS® Modeler peuvent démarrer automatiquement le client SPSS Statistics, mais les utilisateurs doivent d'abord s'assurer que le client SPSS Statistics présente le même mode d'exécution que SPSS Modeler.

Instructions d'installation

Pour obtenir des informations sur l'installation de IBM® SPSS® Modeler Server, reportez-vous aux instructions contenues dans le dossier *\documentation\installation* du CD-ROM correspondant. Des documents différents sont disponibles pour Windows et UNIX.

Pour des informations complètes sur l'installation et l'utilisation de IBM® SPSS® Modeler, reportez-vous au CD-ROM correspondant.

Support de IBM SPSS Modeler

Cette section est destinée aux administrateurs et au personnel du service d'assistance qui apportent leur aide aux utilisateurs de IBM® SPSS® Modeler. Il couvre les sujets suivants :

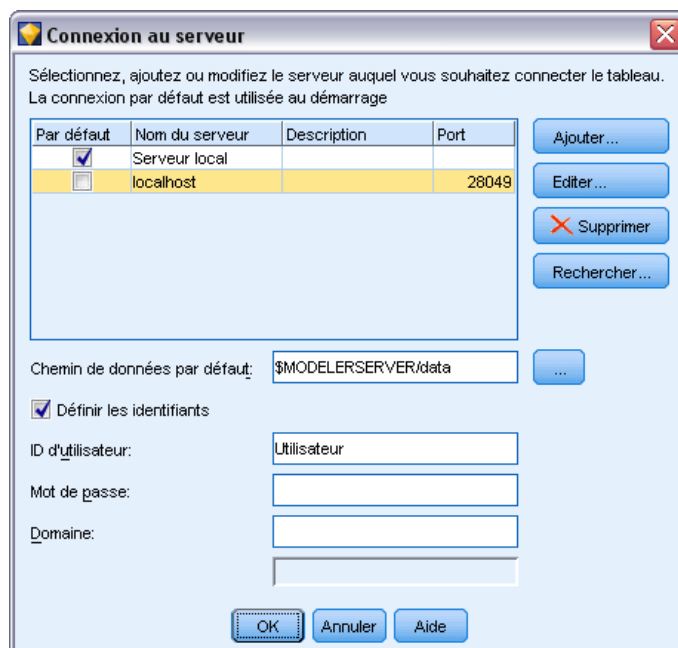
- Comment vous connecter à IBM® SPSS® Modeler Server (ou l'exécuter en tant qu'application autonome en vous déconnectant d'un serveur).
- Systèmes de fichiers et de données nécessaires aux utilisateurs
- Comptes utilisateur et droits d'accès aux fichiers relatifs au serveur SPSS Modeler Server
- Différences au niveau des résultats obtenus avec SPSS Modeler Server et IBM® SPSS® Modeler

Connexion au IBM SPSS Modeler Server

Il est possible d'exécuter IBM® SPSS® Modeler comme une application autonome ou un client connecté directement à IBM® SPSS® Modeler Server ou à SPSS Modeler Server ou à un groupe de serveurs par le biais du Coordinateur des processus connecté à partir de IBM® SPSS® Collaboration and Deployment Services. L'état de la connexion apparaît en bas à gauche de la fenêtre SPSS Modeler.

Lorsque vous souhaitez vous connecter à un serveur, vous pouvez saisir manuellement son nom ou sélectionner un nom que vous aurez préalablement défini. En revanche, si vous avez IBM SPSS Collaboration and Deployment Services, vous avez la possibilité de chercher dans une liste de serveurs ou de groupes de serveurs à partir de la boîte de dialogue Connexion au serveur. Vous pouvez naviguer via les services Statistics s'exécutant sur un réseau grâce au Coordinateur des processus. [Pour plus d'informations, reportez-vous à la section Equilibrage de charge avec classes de serveur dans l'annexe D sur p. 79.](#)

Figure 3-1
Boîte de dialogue Connexion au serveur



Pour vous connecter à un serveur

- ▶ Dans le menu Outils, cliquez sur Connexion au serveur. La boîte de dialogue Connexion au serveur s'affiche. Vous pouvez également double-cliquer sur la zone d'état de la connexion dans la fenêtre SPSS Modeler.
- ▶ Dans la boîte de dialogue, indiquez les options de connexion à l'ordinateur du serveur local ou sélectionnez une connexion dans le tableau.
 - Cliquez sur Ajouter ou Modifier pour ajouter ou modifier une connexion. [Pour plus d'informations, reportez-vous à la section Ajout et modification d'une connexion à IBM SPSS Modeler Server dans Guide de l'utilisateur de IBM SPSS Modeler 14.2.](#)
 - Cliquez sur Rechercher pour accéder au serveur ou à un groupe de serveurs dans le Coordinateur de processus. [Pour plus d'informations, reportez-vous à la section Recherche de serveurs dans IBM SPSS Collaboration and Deployment Services dans Guide de l'utilisateur de IBM SPSS Modeler 14.2.](#)

Tableau de serveur. Ce tableau comprend un ensemble de connexions au serveur définies. Il affiche la connexion par défaut, le nom du serveur, sa description et le numéro du port. Vous pouvez ajouter manuellement une nouvelle connexion ainsi que sélectionner ou rechercher une connexion existante. Pour définir un serveur particulier comme connexion par défaut, cochez la case dans la colonne Par défaut du tableau de la connexion.

Chemin de données par défaut. Indiquez le chemin d'accès aux données situées sur l'ordinateur serveur. Cliquez sur le bouton ... pour accéder à l'emplacement requis.

Définir les informations d'identification. Laissez cette case décochée pour permettre à la fonction de **connexion unique** de se connecter au serveur à l'aide de vos informations de nom d'utilisateur et de mot de passe locaux. Si la connexion unique n'est pas disponible, ou si vous cochez la case pour désactiver la connexion unique (par exemple pour vous connecter à un compte administrateur), les champs suivants sont activés et vous permettent d'entrer vos informations d'identification.

Nom d'utilisateur. Entrez le nom d'utilisateur avec lequel effectuer la connexion au serveur.

Mot de passe : Entrez le mot de passe associé au nom d'utilisateur défini.

Domaine. Indiquez le domaine utilisé pour la connexion au serveur. Le nom de domaine n'est requis que si l'ordinateur serveur se trouve dans un autre domaine Windows que l'ordinateur client.

- Cliquez sur OK pour terminer la connexion.

Pour se déconnecter d'un serveur

- Dans le menu Outils, cliquez sur Connexion au serveur. La boîte de dialogue Connexion au serveur s'affiche. Vous pouvez également double-cliquez sur la zone d'état de la connexion dans la fenêtre SPSS Modeler.
- Dans la boîte de dialogue, sélectionnez le serveur local puis cliquez sur OK.

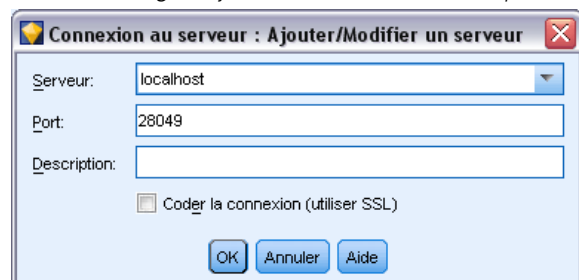
Ajout et modification d'une connexion à IBM SPSS Modeler Server

Dans la boîte de dialogue Connexion au serveur, vous pouvez modifier ou ajouter une connexion au serveur. Cliquez sur Ajouter pour accéder à une boîte de dialogue Ajouter/Modifier un serveur non renseignée, dans laquelle vous pourrez entrer les données de la connexion au serveur. Si vous sélectionnez une connexion existante et cliquez sur Modifier, dans la boîte de dialogue Connexion au serveur, la boîte de dialogue Ajouter/Modifier un serveur s'ouvre, affichant les données de cette connexion, vous permettant ainsi d'apporter toutes les modifications que vous souhaitez.

Remarque : Vous ne pouvez pas modifier une connexion au serveur qui a été ajoutée à partir de IBM® SPSS® Collaboration and Deployment Services, car le nom, le port et d'autres détails sont définis dans IBM SPSS Collaboration and Deployment Services.

Figure 3-2

Boîte de dialogue Ajouter/Modifier un serveur pour la connexion au serveur



Pour ajouter des connexions au serveur

- Dans le menu Outils, cliquez sur Connexion au serveur. La boîte de dialogue Connexion au serveur s'affiche.

- ▶ Dans la boîte de dialogue, cliquez sur Ajouter. La boîte de dialogue Ajouter/Modifier un serveur pour la connexion au serveur s'ouvre.
- ▶ Saisissez les données de connexion au serveur puis cliquez sur OK pour enregistrer la connexion et retourner à la boîte de dialogue Connexion au serveur.
 - **Serveur.** Indiquez un serveur disponible ou sélectionnez-en un dans la liste. L'ordinateur serveur peut être identifié par un nom alphanumérique (par exemple, *monserveur*) ou une adresse IP qui lui est affectée (par exemple, 202.123.456.78).
 - **Port.** Indiquez le numéro de port d'écoute du serveur. Si ce port par défaut ne fonctionne pas, demandez à l'administrateur système le numéro de port correct.
 - **Description.** Saisissez une description optionnelle pour la connexion à ce serveur.
 - **Coder la connexion (utiliser SSL).** Indiquez si une connexion SSL (**Secure Sockets Layer**) doit être utilisée. Le protocole SSL est fréquemment utilisé pour la sécurisation des données sur un réseau. Pour pouvoir utiliser cette fonction, vous devez activer le protocole SSL sur le serveur hébergeant le IBM® SPSS® Modeler Server. Si nécessaire, contactez votre administrateur local pour plus d'informations.

Pour modifier des connexions au serveur

- ▶ Dans le menu Outils, cliquez sur Connexion au serveur. La boîte de dialogue Connexion au serveur s'affiche.
- ▶ Dans la boîte de dialogue, sélectionnez la connexion que vous souhaitez modifier puis cliquez sur Modifier. La boîte de dialogue Ajouter/Modifier un serveur pour la connexion au serveur s'ouvre.
- ▶ Modifiez les données de connexion au serveur puis cliquez sur OK pour enregistrer les changements et retourner à la boîte de dialogue Connexion au serveur.

Recherche de serveurs dans IBM SPSS Collaboration and Deployment Services

Au lieu d'entrer manuellement une connexion au serveur, vous pouvez sélectionner un serveur ou un groupe de serveurs disponible sur le réseau par le biais du Coordinateur de processus, disponible dans IBM® SPSS® Collaboration and Deployment Services. Un groupe de serveurs contient plusieurs serveurs, et permet au Coordinateur de processus de déterminer le serveur qui répond le mieux à la demande de traitement. [Pour plus d'informations, reportez-vous à la section *Équilibrage de charge avec classes de serveur dans l'annexe D sur p. 79.*](#)

Bien que vous ne puissiez pas ajouter manuellement de serveurs dans la boîte de dialogue Connexion au serveur, la recherche de serveurs disponibles vous permet de vous connecter aux serveurs sans que vous ayez besoin de connaître le nom du serveur et le numéro du port. Ces informations sont fournies automatiquement. Il vous faut néanmoins corriger les données de connexion telles que le nom de l'utilisateur, le domaine et le mot de passe.

Remarque : Si vous n'avez pas accès au Coordinateur de processus, vous pouvez tout de même saisir manuellement le nom du serveur auquel vous souhaitez vous connecter ou sélectionner un nom que vous aurez défini au préalable. [Pour plus d'informations, reportez-vous à la section *Ajout et modification d'une connexion à IBM SPSS Modeler Server dans Guide de l'utilisateur de IBM SPSS Modeler 14.2.*](#)

Figure 3-3
Boîte de dialogue Recherche de serveurs



Pour rechercher des serveurs et des groupes de serveurs

- ▶ Dans le menu Outils, cliquez sur Connexion au serveur. La boîte de dialogue Connexion au serveur s'affiche.
- ▶ Cliquez sur Rechercher pour ouvrir la boîte de dialogue Recherche de serveurs. Si vous n'êtes plus connecté à IBM SPSS Collaboration and Deployment Services, lors de votre tentative d'accès au Coordinateur de processus, il vous sera demandé de vous reconnecter. [Pour plus d'informations, reportez-vous à la section Connexion au IBM SPSS Collaboration and Deployment Services Repository dans le chapitre 9 dans Guide de l'utilisateur de IBM SPSS Modeler 14.2.](#)
- ▶ Sélectionnez le serveur ou le groupe de serveurs dans la liste.
- ▶ Cliquez sur OK pour fermer la boîte de dialogue et ajouter cette connexion au tableau de la boîte de dialogue Connexion au serveur.

Systemes de données et de fichiers

Les utilisateurs utilisant IBM® SPSS® Modeler Server devront probablement accéder à des fichiers de données ainsi qu'à d'autres sources de données sur le réseau, ainsi qu'enregistrer des fichiers sur le réseau. Ils peuvent avoir besoin des informations suivantes :

- **Informations des sources de données ODBC.** Si les utilisateurs doivent accéder aux sources de données ODBC définies sur l'ordinateur serveur, ils ont besoin des noms, descriptions et informations de connexion (y compris les ID et mots de passe de connexion de base de données) correspondantes.
- **Accès aux fichiers de données.** Si les utilisateurs doivent accéder aux fichiers de données disponibles sur l'ordinateur serveur ou à un autre emplacement du réseau, ils ont besoin des noms et emplacements de ces fichiers.
- **Emplacement des fichiers enregistrés.** Lorsque les utilisateurs enregistrent des données tout en étant connectés à SPSS Modeler Server, ils tentent parfois d'enregistrer les fichiers sur l'ordinateur serveur. Toutefois, l'emplacement visé est souvent protégé en écriture. Si c'est le cas, indiquez aux utilisateurs où enregistrer les fichiers de données. (Généralement, il s'agit du répertoire personnel de l'utilisateur.)

Authentification de l'utilisateur

IBM® SPSS® Modeler Server utilise le système d'exploitation de l'ordinateur serveur pour authentifier les utilisateurs qui se connectent au serveur. Lorsqu'un utilisateur se connecte à SPSS Modeler Server, toutes les opérations effectuées sous son nom sont exécutées dans le contexte de sécurité de cet utilisateur. L'accès aux tables de base de données est soumis à des privilèges d'utilisateur et/ou de mot de passe, au niveau de la base de données proprement dite.

Windows. Sous Windows, tout utilisateur possédant un compte valide sur le réseau hôte peut se connecter. Avec l'identification par défaut, les utilisateurs doivent avoir des droits d'accès au répertoire `<modeler_server_install>\Tmp` modifiés. Sans ces droits, les utilisateurs ne peuvent pas se connecter à SPSS Modeler Server en utilisant l'authentification par défaut sous Windows.

UNIX. Par défaut, SPSS Modeler Server est considéré comme s'exécutant en tant que processus racine sous UNIX. Cela permet à tous les utilisateurs disposant d'un compte valide sur le réseau hôte de se connecter et limite l'accès des utilisateurs à leurs propres fichiers et répertoires. Toutefois, vous pouvez configurer SPSS Modeler Server en vue d'une exécution sans privilège racine. Dans ce cas, vous devez créer une base de données de mots de passe privés à utiliser pour l'authentification. Tous les utilisateurs IBM® SPSS® Modeler partagent alors un seul compte utilisateur UNIX (par conséquent, ils partagent l'accès aux fichiers de données). [Pour plus d'informations, reportez-vous à la section Configuration de IBM SPSS Modeler Server en tant que processus non-racine dans l'annexe C sur p. 76.](#)

Sous les plates-formes Solaris, HP-UX, Linux et AIX SPSS Modeler Server utilise PAM pour l'authentification. Vous pouvez utiliser le nom de service `modelerserver` pour configurer les modules PAM pour SPSS Modeler Server, si nécessaire.

Création de fichiers

Lorsque IBM® SPSS® Modeler Server accède à des données et les traite, il effectue souvent une copie temporaire de ces données sur le disque. La quantité d'espace disque qui sera utilisée pour les fichiers temporaires dépend de la taille du fichier de données que l'utilisateur final analyse et du type d'analyse effectué. [Pour plus d'informations, reportez-vous à la section Exigences en termes d'espace disque temporaire et de mémoire RAM dans le chapitre 2 sur p. 9.](#)

UNIX. Les versions UNIX de SPSS Modeler Server utilisent la commande UNIX `umask` pour définir les droits d'accès aux fichiers temporaires. Vous pouvez remplacer les droits d'accès par défaut du serveur. [Pour plus d'informations, reportez-vous à la section Contrôle des droits d'accès relatifs à la création de fichier dans l'annexe B sur p. 68.](#)

Différences au niveau des résultats

Les utilisateurs qui effectuent des analyses dans les deux modes peuvent constater de légères différences dans les résultats obtenus avec IBM® SPSS® Modeler et IBM® SPSS® Modeler Server. Ces dernières sont généralement dues à des divergences au niveau des arrondis ou de l'ordre des enregistrements.

Ordre des enregistrements. Sauf si un flux ordonne explicitement les enregistrements en les triant, l'ordre dans lequel les enregistrements sont présentés peut varier entre les flux exécutés localement et ceux exécutés sur le serveur. Des différences d'ordre peuvent également survenir entre les opérations effectuées dans une base de données et celles exécutées dans SPSS Modeler Server. Ces différences sont dues aux divers algorithmes utilisés par chaque système pour mettre en oeuvre des fonctions susceptibles de réorganiser les enregistrements, tels que l'agrégation. Notez également que SQL ne spécifie pas l'ordre dans lequel les enregistrements sont renvoyés à partir d'une base de données si aucune opération de tri explicite n'a été définie.

Différences au niveau des arrondis. En mode local, IBM® SPSS® Modeler utilise un format interne différent de celui de SPSS Modeler Server pour stocker les valeurs à virgule flottante. En raison de divergences au niveau des arrondis, les résultats peuvent varier légèrement entre chaque version.

Administration de IBM SPSS Modeler Server

Ce chapitre contient des informations sur le démarrage et l'arrêt de IBM® SPSS® Modeler Server, sur la configuration de diverses options du serveur, ainsi que sur l'interprétation du fichier journal. Il décrit l'utilisation de l'application IBM® SPSS® Collaboration and Deployment Services Deployment Manager, qui facilite la configuration et la surveillance du serveur. Pour obtenir des instructions d'installation concernant ce composant, reportez-vous à celles de SPSS Modeler Server, disponibles sur le CD du produit.

Démarrage et arrêt de IBM SPSS Modeler Server

IBM® SPSS® Modeler Server est exécuté en tant que service sous Windows, ou en tant que processus démon sous UNIX .

Remarque concernant le redémarrage : L'arrêt de SPSS Modeler Server déconnecte les utilisateurs finaux et met fin à leurs sessions. Par conséquent, redémarrez le serveur pendant les périodes de faible utilisation. Si ce n'est pas possible, avertissez les utilisateurs de l'arrêt du serveur.

Pour démarrer ou arrêter le serveur sous Windows, ou vérifier son état

Sous Windows, vous contrôlez IBM® SPSS® Modeler Server via la boîte de dialogue Services du Panneau de configuration.

- ▶ **Windows XP.** Ouvrez le menu Démarrer de Windows. Sélectionnez Paramètres, puis Panneau de configuration. Double-cliquez sur Outils d'administration, puis sur Services.

Windows 2003 ou 2008. Ouvrez le menu Démarrer de Windows. Sélectionnez Panneau de configuration, puis Outils d'administration et Services.

- ▶ Sélectionnez le service IBM SPSS Modeler Server <nn.n>. Vous pouvez maintenant vérifier l'état du service, le démarrer ou l'arrêter, et éditer les paramètres de démarrage à votre gré.

Par défaut, le service est configuré pour un démarrage automatique, ce qui signifie que si vous l'arrêtez, il redémarrera automatiquement en même temps que l'ordinateur. Lorsque le service est démarré de cette façon, il fonctionne seul. Vous pouvez donc vous déconnecter de l'ordinateur serveur sans risquer d'interrompre le service.

Pour démarrer ou arrêter le serveur sous UNIX, ou vérifier son état

Sous UNIX, vous démarrez ou arrêtez IBM® SPSS® Modeler Server en exécutant le script `modelersrv.sh` dans le répertoire d'installation de SPSS Modeler Server.

- Remplacez le répertoire actuel par le répertoire d'installation de SPSS Modeler Server. Par exemple, à l'invite de commande UNIX, entrez

```
cd /usr/modelersrv
```

où *modelersrv* est le répertoire d'installation de SPSS Modeler Server.

- À l'invite de commande, pour démarrer le serveur, entrez la ligne suivante :

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

- À l'invite de commande, pour arrêter le serveur, entrez la ligne suivante

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

- Pour vérifier l'état de SPSS Modeler Server, à l'invite de commande UNIX, entrez la ligne suivante :

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

et examinez la sortie, semblable à celle générée par la commande UNIX `ps`. Le premier processus de la liste est le processus démon SPSS Modeler Server, les autres sont les sessions IBM® SPSS® Modeler.

Le programme d'installation de SPSS Modeler Server inclut un script (*auto.sh*) qui configure votre système de façon à démarrer le démon du serveur automatiquement lors du démarrage. Si vous avez exécuté ce script, puis arrêté le serveur, le démon du serveur redémarre automatiquement en même temps que l'ordinateur. [Pour plus d'informations, reportez-vous à la section Démarrage et arrêt automatiques de IBM SPSS Modeler Server dans l'annexe B sur p. 67.](#)

limites du noyau UNIX

Vous devez vous assurer que les limites du noyau sur le système sont suffisantes pour utiliser SPSS Modeler Server. Les limites des données, de la mémoire et les valeurs `ulimit` du fichier sont particulièrement importantes et doivent être définies sur illimitées au sein de l'environnement SPSS Modeler Server. Pour ce faire :

- Ajoutez les commandes suivantes à *modelersrv.sh* :

```
ulimit -d unlimited
```

```
ulimit -m unlimited
```

```
ulimit -f unlimited
```

- Redémarrez SPSS Modeler Server.

Traitement des processus de serveurs qui ne répondent pas (systèmes UNIX)

Les processus IBM® SPSS® Modeler Server peuvent ne pas répondre pour plusieurs raisons, ceci comprend les situations où ils effectuent un appel au système ou un appel au pilote ODBC et que l'appel se bloque (l'appel n'a pas de réponse, ou cela prend très longtemps de recevoir une réponse). Lorsque les processus UNIX sont dans cet état, ils peuvent être arrêtés à l'aide de la commande UNIX kill (les interruptions initiées par le client IBM® SPSS® Modeler ou la fermeture du client SPSS Modeler n'ont aucun effet). Une commande kill est fournie comme alternative à la commande normale stop, et permet à un administrateur d'utiliser *modelersrv.sh* pour émettre facilement la commande kill appropriée.

Sur les systèmes susceptibles d'accumuler des processus de serveurs inutilisables ("zombies"), nous recommandons d'arrêter SPSS Modeler Server et de le redémarrer à intervalles réguliers, à l'aide de la séquence de commandes suivantes :

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

Ces processus SPSS Modeler qui sont arrêtés à l'aide de la commande *modelersrv.sh kill* laisseront des fichiers temporaires (du répertoire temporaire) qui devront être supprimés manuellement. Les fichiers temporaires peuvent aussi ne pas être supprimés automatiquement dans d'autres situations, par exemple des fermetures non souhaitées d'application dues à un manque de ressources du système, des interruptions causées par l'utilisateur, des pannes du système, ou d'autres raisons. Par conséquent nous recommandons de supprimer tous les fichiers existants dans le répertoire temporaire SPSS Modeler, comme partie intégrale du processus de redémarrage de SPSS Modeler Server à intervalles réguliers.

Une fois tous les processus de serveur fermés et les fichiers temporaires supprimés, SPSS Modeler Server peut être redémarré en toute sécurité.

Administration

IBM® SPSS® Modeler Server dispose d'un certain nombre d'options configurables qui contrôlent son comportement. Vous pouvez définir ces options de deux façons :

- Utilisez l'application IBM® SPSS® Collaboration and Deployment Services Deployment Manager, qui est disponible gratuitement pour les clients actuels de IBM® SPSS® Modeler. [Pour plus d'informations, reportez-vous à la section Utilisation de IBM SPSS Modeler Administration Console sur p. 24.](#)
- Utilisez le fichier texte *options.cfg*, qui se trouve dans le répertoire [*dossier d'installation du serveur*]/*config*. [Pour plus d'informations, reportez-vous à la section Utilisation du fichier options.cfg sur p. 34.](#)

Nous vous conseillons d'installer Deployment Manager et de l'utiliser comme outil d'administration, au lieu d'éditer le fichier *options.cfg*. L'édition du fichier requiert l'accès au système de fichiers SPSS Modeler Server, alors que Deployment Manager vous permet d'autoriser

toute personne disposant d'un compte utilisateur à modifier ces options. De plus, Deployment Manager fournit des informations supplémentaires sur les processus serveur. Vous pouvez donc en contrôler l'utilisation et les performances. En outre, la plupart des options de configuration peuvent être modifiées sans redémarrer SPSS Modeler Server, ce qui n'est pas le cas du fichier de configuration.

Pour plus d'informations sur l'utilisation de Deployment Manager et du fichier *options.cfg*, reportez-vous aux sections suivantes.

Utilisation de IBM SPSS Modeler Administration Console

L'application Modeler Administration Console offre une interface utilisateur sous forme de console permettant de surveiller et de configurer les installations SPSS Modeler Server ; elle est disponible gratuitement pour les clients actuels de SPSS Modeler Server. L'application ne peut être installée que sur des ordinateurs Windows ; en revanche, elle peut administrer un serveur installé sur n'importe quelle plate-forme prise en charge.

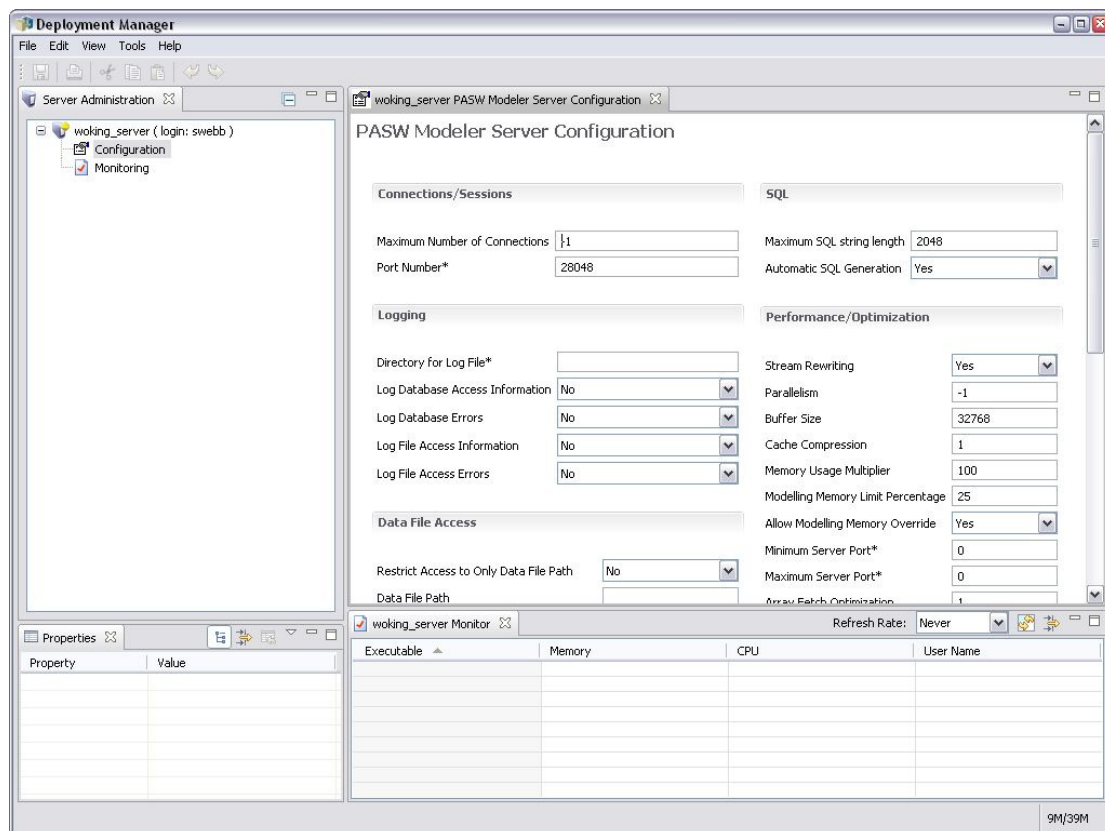
Vous pouvez également spécifier la plupart des options disponibles via Modeler Administration Console dans le fichier *options.cfg*, situé dans le répertoire d'installation de SPSS Modeler Server, sous */config*. Toutefois, l'application Modeler Administration Console fournit une interface graphique partagée qui vous permet de connecter, configurer et surveiller plusieurs serveurs.

Démarrage de Modeler Administration Console

- Dans le menu Démarrer de Windows, choisissez [Tous les] Programmes, puis IBM SPSS Collaboration and Deployment Services, puis Deployment Manager, et IBM SPSS Modeler Administration Console<nn.n>.

Lorsque vous exécutez l'application pour la première fois, les panneaux Administration et propriétés du serveur sont vides (sauf si vous avez installé Deployment Manager avec une connexion au serveur IBM SPSS Collaboration and Deployment Services déjà configurée). Une fois Modeler Administration Console configuré, le panneau Administrateur du serveur situé à gauche affiche un nœud pour chaque SPSS Modeler Server à administrer. Le panneau situé à droite présente les options de configuration relatives au serveur sélectionné. Vous devez d'abord [paramétrer une connexion](#) pour chaque serveur à administrer.

Figure 4-1
interface Modeler Administration Console



Configuration de l'accès avec Modeler Administration Console

L'accès administrateur à SPSS Modeler Server via Modeler Administration Console est contrôlé à l'aide de la ligne `administrators` dans le fichier `options.cfg`, lequel est situé dans le répertoire d'installation de SPSS Modeler Server sous `/config`. Cette ligne est commentée par défaut et vous devez la modifier pour autoriser l'accès à des personnes spécifiques ou utiliser `*` pour autoriser l'accès à tous les utilisateurs, comme dans les exemples suivants :

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

- La ligne doit commencer par `administrators`, et les entrées doivent être placées entre guillemets. Les entrées distinguent les majuscules des minuscules.
- Utilisez une virgule pour séparer les ID utilisateur.
- Pour les comptes Windows, n'utilisez pas de noms de domaine.
- Utilisez l'astérisque avec précaution. En effet, ce signe permet à quiconque disposant d'un compte utilisateur valide pour IBM® SPSS® Modeler Server (dans la plupart des cas, tout utilisateur sur le réseau) de se connecter et de modifier les options de configuration.

SPSS Modeler Server Connexions

Vous devez indiquer une connexion à chaque serveur SPSS Modeler Server du réseau que vous souhaitez administrer. Vous devez ensuite vous connecter à chaque serveur. Bien que la connexion au serveur soit mémorisée d'une session Modeler Administration Console à l'autre, les informations d'identification et de connexion ne le sont pas. Vous devez vous connecter à chaque démarrage de Modeler Administration Console.

Pour configurer une connexion au serveur

- ▶ Vérifiez que le service IBM® SPSS® Modeler Server est démarré.
- ▶ Dans le menu Fichier, choisissez Nouveau, puis Connexion du serveur administré.
- ▶ Dans la première page de l'assistant, saisissez un nom pour la connexion au serveur. Il convient que ce nom, défini pour votre seul usage, soit descriptif ; par exemple, *Serveur de production*. Vérifiez que le type est défini sur IBM SPSS Modeler Server administré , puis cliquez sur Suivant.
- ▶ Dans la seconde page, saisissez le nom d'hôte ou l'adresse IP du serveur. Si vous avez modifié le port par défaut, saisissez le numéro de port. Cliquez sur Terminer. La nouvelle connexion au serveur apparaît dans le panneau Administrateur du serveur.

Pour pouvoir exécuter des tâches d'administration, vous devez alors vous connecter.

Pour vous connecter au serveur

- ▶ Dans le panneau Administrateur du serveur, double-cliquez pour sélectionner le serveur auquel vous souhaitez vous connecter.
- ▶ Dans la boîte de dialogue de connexion, saisissez vos informations d'identification et de connexion. (Utilisez votre compte utilisateur pour l'hôte du serveur.) Cliquez sur OK.

Si la connexion échoue et que le message Impossible d'obtenir des droits d'administration sur le serveur s'affiche, la cause la plus probable est que l'accès de l'administrateur n'a pas été configuré correctement. [Pour plus d'informations, reportez-vous à la section Configuration de l'accès avec Modeler Administration Console sur p. 25.](#)

Si la connexion échoue avec le message Echec de connexion au serveur '<serveur>', vérifiez que l'ID utilisateur et le mot de passe sont corrects puis vérifiez que le service SPSS Modeler Server fonctionne. Par exemple, sous Windows, allez dans le Panneau de configuration, > Outils d'administration > Services et vérifiez l'entrée de IBM SPSS Modeler Server. Si la colonne Etat n'affiche pas Démarré, sélectionnez cette ligne à l'écran et cliquez sur Démarrer, puis essayez de nouveau de vous connecter.

Une fois que vous êtes connecté à SPSS Modeler Server, deux options apparaissent sous le nom du serveur, [Configuration](#) et [Surveillance](#). Double-cliquez sur l'une de ces options.

Configuration de SPSS Modeler Server

Le panneau Configuration présente les options de configuration de SPSS Modeler Server. Utilisez-le pour modifier les options selon vos besoins. Pour enregistrer les modifications, cliquez sur Enregistrer dans la barre d'outils. (Pour que la modification du numéro de port soit effective, un redémarrage du serveur doit avoir lieu.)

Les options sont décrites dans les sections suivantes ; la ligne correspondante du fichier *options.cfg* est fournie entre parenthèses. Les options uniquement visibles dans *options.cfg* sont décrites à la fin de cette section.

Connexions/Sessions

Figure 4-2
Paramètres relatifs aux connexions/sessions

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

Nombre de connexions maximal. (max_sessions) Nombre maximal d'utilisateurs connectés simultanément. La valeur -1 indique l'absence de limite.

Numéro de port. (port_number) Numéro de port sur lequel on écoute SPSS Modeler Server. Modifiez-le si une autre application utilise déjà le numéro par défaut (28047). Pour pouvoir utiliser SPSS Modeler Server, les utilisateurs finaux doivent connaître le numéro de port.

Consignation

Figure 4-3
Paramètres relatifs à la consignation

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

Répertoire du fichier journal. (log_directory) Emplacement du fichier *messages.log*. Une chaîne non renseignée indique [dossier d'installation du serveur]/log. Vous devez utiliser des barres obliques normales. Pour plus d'informations sur le contenu de ce fichier journal, reportez-vous à [Journal du serveur sur p. 42](#).

Consigner les informations sur l'accès à la base de données. (log_database_access) Ajoute, au fichier journal, les informations sur l'accès à la base de données.

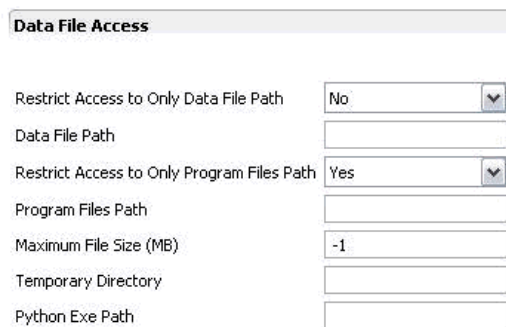
Consigner les erreurs de base de données. (`log_database_errors`) Ajoute, au fichier journal, les erreurs de base de données.

Consigner les informations sur l'accès aux fichiers. (`log_file_access`) Ajoute, au fichier journal, les informations sur l'accès aux fichiers.

Consigner les erreurs d'accès aux fichiers. (`log_file_errors`) Ajoute, au fichier journal, les erreurs sur l'accès aux fichiers.

Accès aux fichiers de données

Figure 4-4
Paramètres relatifs à l'accès aux fichiers de données



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	

Restreindre l'accès au chemin du fichier de données uniquement. (`data_files_restricted`) Lorsque cette option est paramétrée sur *yes* (oui), elle limite l'accès aux fichiers de données du répertoire de données standard et à ceux répertoriés dans la liste Chemin du fichier de données ci-dessous.

Chemin du fichier de données. (`data_file_path`) Liste de répertoires supplémentaires pour lesquels les clients disposent d'un accès en lecture et écriture sur les fichiers de données. A moins que l'option Restreindre l'accès au chemin du fichier de données uniquement ne soit activée, cette option est ignorée. Veillez à utiliser des barres obliques normales (/) dans tous les noms de chemin. Pour indiquer plusieurs répertoires, utilisez des points-virgules (par exemple, *[dossier d'installation du serveur]/données;c:/données;c:/temp*).

Restreindre l'accès au chemin des fichiers de programme uniquement. (`program_files_restricted`) Lorsque cette option est paramétrée sur *yes* (oui), elle limite l'accès aux fichiers de programme du répertoire *bin* standard et à ceux répertoriés dans la liste Chemin des fichiers de programme ci-dessous.

Chemin des fichiers de programme. (`program_file_path`) Liste de répertoires supplémentaires à partir desquels les clients sont autorisés à exécuter des programmes. A moins que l'option Restreindre l'accès au chemin des fichiers de programme uniquement ne soit activée, cette option est ignorée. Veillez à utiliser des barres obliques normales (/) dans tous les noms de chemin. Pour indiquer plusieurs répertoires, utilisez des points-virgules.

Taille de fichier maximale (Mo). (`max_file_size`) Taille maximale des fichiers de données temporaires et exportés, créés durant l'exécution du flux (ne s'applique pas aux fichiers de données SAS et SPSS Statistics). La valeur -1 indique l'absence de limite.

Répertoire temporaire. (temp_directory) Répertoire utilisé pour stocker les fichiers de données temporaires (fichiers de cache). Il est recommandé de placer ce répertoire sur un contrôleur ou un lecteur haut débit distinct car la vitesse d'accès à ce répertoire peut avoir une incidence importante sur les performances. Vous pouvez définir plusieurs répertoires temporaires, en les séparant par une virgule. Ils doivent se trouver sur des disques différents ; le premier répertoire est celui le plus utilisé et les autres permettent le stockage des fichiers de travail temporaires lorsque certaines opérations de préparation des données (telles que le tri) utilisent le parallélisme au cours de leur exécution. Vous pouvez améliorer les performances en permettant à chaque unité d'exécution d'utiliser des disques distincts pour le stockage temporaire. Utilisez des barres obliques (/) dans toutes les spécifications de chemin. *Remarque* : des fichiers temporaires sont générés dans ce répertoire au cours du démarrage de SPSS Modeler Server. Vérifiez que vous possédez les droits d'accès nécessaires pour accéder à ce répertoire (par exemple, si le répertoire temporaire est un dossier réseau partagé), sinon le démarrage de SPSS Modeler Server échouera.

Chemin de l'exécutable Python. (python_exe_path) Chemin complet de l'exécutable python, y compris le nom de l'exécutable. Notez qu'il peut être nécessaire de configurer l'option Restreindre l'accès au chemin des fichiers de programme uniquement sur *Non* en fonction de l'emplacement de l'installation de python.

Performances/Optimisation

Figure 4-5
Paramètres relatifs aux performances/à l'optimisation

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

Réécriture des flux. (stream_rewriting_enabled) Permet au serveur d'optimiser les flux en les réécrivant. Par exemple, le serveur peut rapprocher du noeud source les opérations de réduction des données afin de réduire au maximum la taille des ensembles de données aussi tôt que possible. Il est généralement recommandé de ne désactiver cette option que si l'optimisation entraîne une erreur ou tout autre résultat inattendu. Ce paramètre remplace le paramètre d'optimisation client correspondant.

Parallélisme. (max_parallelism) Décrit le nombre de threads de travail parallèles que SPSS Modeler est autorisé à utiliser lors de l'exécution d'un flux. Si vous paramétrez cette fonction sur 0 ou sur un autre nombre négatif, IBM® SPSS® Modeler fait correspondre le nombre de threads au nombre de processeurs disponibles sur l'ordinateur ; la valeur par défaut pour cette option est -1.

Pour désactiver le traitement parallèle (pour les ordinateurs possédant plusieurs processeurs), paramétrez cette option sur 1. Pour autoriser le traitement parallèle limité, paramétrez-la sur un nombre inférieur au nombre de processeurs installés sur votre ordinateur. Un processeur HyperThread ou à double coeur est traité comme s'il existait deux processeurs.

Taille du tampon (octets). (`io_buffer_size`) Les fichiers de données transférés du serveur au client passent par une mémoire tampon dont le nombre d'octets est indiqué dans cette option.

Compression du cache. (`cache_compression`) Entier compris entre 0 et 9 qui contrôle la compression du cache et d'autres fichiers dans le répertoire temporaire du serveur. La compression réduit la quantité d'espace disque utilisé, ce qui peut s'avérer important lorsque l'espace disponible est limité ; elle permet d'améliorer les performances grâce à la réduction de l'activité de disque nécessaire à la création et à la lecture du cache. La compression augmente également le temps processeur, même si cela s'effectue généralement par réduction du temps d'accès au disque. Notez que seulement certains caches, ceux auxquels vous accédez de manière séquentielle, peuvent être compressés. Cette option ne s'applique pas aux caches à accès aléatoire, tels que ceux utilisés par les algorithmes d'apprentissage du réseau. Une valeur de 0 permet de désactiver intégralement la compression. Les valeurs supérieures ou égales à 1 fournissent des degrés croissants de compression, mais elles induisent une augmentation correspondante du temps d'accès. La valeur par défaut est 1 ; il est généralement inutile de recourir aux valeurs supérieures, sauf lorsque l'espace disque est grandement insuffisant.

Multiplicateur d'utilisation de la mémoire. (`memory_usage`) Contrôle la proportion de mémoire physique allouée au tri et aux autres caches en mémoire. La valeur par défaut, 100, correspond à environ 10 % de la mémoire physique. Augmentez cette valeur pour améliorer les performances de tri lorsque vous disposez d'une certaine quantité de mémoire libre, mais veillez à ne pas indiquer une valeur trop élevée qui entraînerait une pagination excessive.

Limite (en pourcentage) de la mémoire de modélisation. (`modelling_memory_limit_percentage`) Contrôle la proportion de mémoire physique allouée aux modèles d'apprentissage R, neurones, Kohonen et k -means. La valeur par défaut est 25%. Augmentez cette valeur pour améliorer les performances d'apprentissage lorsque vous disposez d'une certaine quantité de mémoire libre, mais veillez à ne pas indiquer une valeur trop élevée qui entraînerait une pagination excessive lors du débordement des données sur le disque.

Autoriser la non-prise en compte de la mémoire de modélisation.

(`allow_modelling_memory_override`) Active ou désactive l'option d'optimisation de la vitesse dans certains noeuds de modélisation. La valeur par défaut est activée. Cette option permet à l'algorithme de modélisation d'obtenir toute la mémoire disponible, passant ainsi outre à l'option de limite exprimée en pourcentage. Vous pouvez désactiver cette option si vous avez besoin de partager les ressources mémoire de l'ordinateur serveur.

Valeurs maximale et minimale de port du serveur. (`max_server_port` et `min_server_port`) Indique l'intervalle de numéros de port pouvant être utilisé pour les connexions supplémentaires entre le client et le serveur qui sont requises pour les modèles interactifs et l'exécution des flux. Ces connexions nécessitent que le serveur écoute sur un autre port ; le fait de ne pas restreindre l'intervalle peut provoquer des problèmes pour les utilisateurs de systèmes dotés de pare-feu. La valeur par défaut pour les deux est -1. Autrement dit, aucune restriction n'est appliquée. Ainsi, par exemple, pour que le serveur écoute sur le port 8000 ou supérieur, vous devrez paramétrer `min_server_port` sur 8000 et `max_server_port` sur -1.

Notez que vous devez ouvrir des ports supplémentaires sur le port du serveur principal pour ouvrir ou exécuter un flux et autant d'autres ports que nécessaire si vous voulez ouvrir ou exécuter des flux simultanés. Cela est nécessaire pour capturer des commentaires issus de l'exécution du flux.

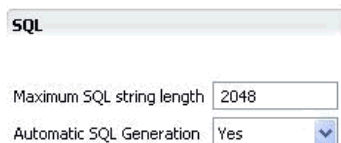
Par défaut, SPSS Modeler utilise tout port disponible ; s'il n'en trouve pas (par exemple, s'ils sont tous fermés par le pare-feu), une erreur s'affiche lorsque vous exécutez le flux. Pour configurer la plage des ports, SPSS Modeler aura besoin de deux ports ouverts (en plus du port du serveur principal) disponibles par flux simultané, ainsi que d'un port supplémentaire pour chaque client connecté ayant une connexion ODBC. Pour exécuter davantage de flux simultanés, vous pouvez élargir la plage des numéros de port et ouvrir plus de ports dans le pare-feu en conséquence.

Remarque : Si vous modifiez ces paramètres, vous devrez redémarrer SPSS Modeler Server pour que la modification prenne effet.

Optimisation de l'extraction du tableau. (`sql_row_array_size`) Contrôle la façon dont SPSS Modeler Server obtient des données de la source de données ODBC. La valeur par défaut est 1, ce qui entraîne l'extraction d'une seule ligne à la fois. L'augmentation de cette valeur incite le serveur à lire les informations par fragments plus importants, ce qui entraîne l'extraction du nombre de lignes spécifié dans un tableau. Avec certaines combinaisons système d'exploitation/base de données, vous pouvez obtenir des améliorations sensibles en matière d'exécution d'instructions SELECT.

SQL

Figure 4-6
Paramètres SQL



SQL

Maximum SQL string length 2048

Automatic SQL Generation Yes

Longueur de chaîne SQL maximale. (`max_sql_string_length`) Longueur maximale d'une chaîne importée depuis la base de données avec SQL. Les valeurs de chaîne supérieures à cette longueur sont tronquées sur la droite sans avertissement. La plage valide est comprise entre 1 et 65 535 caractères.

Génération SQL automatique. (`sql_generation_enabled`) Permet la génération automatique SQL pour les flux, ce qui peut améliorer considérablement les performances. La valeur par défaut est activée. Il n'est conseillé de désactiver cette option que si la base de données ne peut pas prendre en charge les requêtes soumises par SPSS Modeler Server. Remarque : ce paramètre remplace le paramètre d'optimisation client correspondant ; veillez également à activer la génération SQL séparément pour chaque noeud de modélisation à des fins de scoring, quelle que soit la valeur de ce paramètre.

SSL

Figure 4-7
Paramètres SSL

Activer SSL. (ssl_enabled) Active le codage SSL des connexions entre SPSS Modeler et SPSS Modeler Server.

Fichier de certificat. (ssl_certificate_file) Fichier du certificat SSL public à charger lorsque le serveur démarre.

Fichier de clé privée. (ssl_private_key_file) Fichier de clé SSL privé à charger sur le serveur au démarrage.

Mot de passe de la clé privée. (ssl_private_key_password) Mot de passe de la clé pour le certificat spécifié. Notez qu'au moins l'une des autorités de certification répertoriées dans le fichier de certificat configuré doit avoir signé cette clé.

Configuration du Coordinateur de processus

Figure 4-8
Paramètres de configuration du Coordinateur de processus

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

Hôte. (cop_host) Le nom de l'hôte ou l'adresse IP du service du Coordinateur de processus. Le "spsscop" par défaut est un nom personnalisé que les administrateurs peuvent choisir d'ajouter comme un alias pour l'hôte IBM SPSS Collaboration and Deployment Services dans le DNS.

Numéro de port. (cop_port_number) Numéro de port du service du Coordinateur de processus. Le numéro de port par défaut IBM SPSS Collaboration and Deployment Services est 8080.

Nom de connexion. (cop_user_name) Nom d'utilisateur pour l'authentification au service du Coordinateur de processus. Il s'agit d'un nom de connexion IBM SPSS Collaboration and Deployment Services qui peut donc comprendre un préfixe de fournisseur de sécurité (par exemple : ad/jsmith).

Mot de passe : (cop_password) Mot de passe codé pour l'authentification sur le service du Coordinateur de processus.

Pour créer le fichier de mots de passe, utilisez l'utilitaire de création de mots de passe, *pwutil*, qui se trouve dans le répertoire *bin* de l'installation SPSS Modeler Server.

Activé. (cop_enabled) Détermine si le serveur doit tenter de s'enregistrer avec le Coordinateur de processus. La valeur par défaut est de *ne pas* s'enregistrer car l'administrateur doit choisir quels services sont affichés par le du Coordinateur de processus.

Nom de serveur. (cop_service_name) Nom de cette instance SPSS Modeler Server ; par défaut il s'agit du nom de l'hôte.

Description : (cop_service_description) Une description de cette instance.

Intervalle de mise à jour (min). (cop_update_interval) Nombre de minutes entre des messages préservés ; la valeur par défaut est de 2.

Pondération. (cop_service_weight) Pondération de cette instance, spécifiée sous la forme d'un entier compris entre 1 et 10. Plus la pondération est grande, plus elle attire de connexions. La valeur par défaut est 1.

Hôte de service. (cop_service_host) Nom complet de l'hôte de IBM® SPSS® Modeler Server. La valeur par défaut du nom de l'hôte est déduite automatiquement ; l'administrateur peut la remplacer pour des hôtes multiconnectés.

Options visibles uniquement dans options.cfg

`default_sql_string_length`. Spécifie la largeur par défaut des colonnes de la chaîne qui seront créées dans les tables de mise en cache des bases de données. Les champs de type chaîne dans les tables de mise en cache des bases de données seront créés avec une largeur par défaut de 255 si aucune information `Typers` en amont n'existe. Si vous disposez de valeurs plus importantes que celles de vos données, instantiez un noeud `Typers` en amont avec ces valeurs ou définissez ce paramètre sur une valeur assez importante pour accommoder ces valeurs de chaîne.

`max_transfer_size`. Pour système interne uniquement ; ne pas modifier.

`start_process_as_login_user`. Définissez sur `true` si vous utilisez la connexion unique et que vous êtes connecté à un serveur IBM® SPSS® Collaboration and Deployment Services qui n'utilise pas la même méthode d'authentification que l'hôte IBM® SPSS® Modeler Server. Définir cette option sur `true` permet aux processus enfants d'utiliser les informations de connexion de SPSS Modeler Server.

`administrators`. Spécifiez les noms d'utilisateurs des utilisateurs auxquels vous souhaitez octroyer l'accès en tant qu'administrateurs. [Pour plus d'informations, reportez-vous à la section Configuration de l'accès avec Modeler Administration Console sur p. 25.](#)

`supported_sql_type_logging`. Pour le débogage uniquement ; ne pas modifier sauf si le représentant technico-commercial vous a donné des instructions spécifiques.

`allow_config_custom_overrides`. Ne pas modifier sauf si le représentant technico-commercial vous a donné des instructions spécifiques.

SPSS Modeler Server Surveillance

Le panneau de surveillance de l'application Modeler Administration Console présente un instantané de tous les processus en cours d'exécution sur l'ordinateur SPSS Modeler Server (de la même manière que le Gestionnaire des tâches de Windows). Pour activer le panneau de surveillance, double-cliquez sur le noeud de surveillance sous le serveur souhaité dans le panneau Administrateur du serveur. Le panneau affiche alors un instantané des données issues du serveur. Les données sont rafraîchies à la fréquence indiquée (une minute par défaut). Pour rafraîchir manuellement les données, cliquez sur le bouton Rafraîchir. Pour afficher uniquement les processus SPSS Modeler Server dans la liste, cliquez sur le bouton Filtrer les processus non-SPSS Modeler.

Figure 4-9
Panneau de surveillance 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
VsTskMgr	464 K	00	SYSTEM
clemserv	10,452 K	00	SYSTEM
clemserv	22,312 K	00	SYSTEM

54 Processes 3M/4M

Utilisation du fichier `options.cfg`

Le fichier `options.cfg` se trouve dans le répertoire `[dossier d'installation du serveur]/config`. Chaque paramètre est représenté par une paire nom-valeur (séparée par une virgule), le **nom** et la **valeur** étant ceux de l'option. Le signe dièse (#) indique des commentaires.

Remarque : À l'exception de l'option suivante, toutes les options de configuration peuvent être modifiées à l'aide de IBM® SPSS® Collaboration and Deployment Services Deployment Manager, plutôt qu'avec ce fichier de configuration :

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

En utilisant Deployment Manager, vous évitez les redémarrages du serveur pour toutes les options, excepté pour le port du serveur. [Pour plus d'informations, reportez-vous à la section Utilisation de IBM SPSS Modeler Administration Console sur p. 24.](#)

Options de configuration pouvant être ajoutées au fichier par défaut

Par défaut, la mise en mémoire cache dans les bases de données est activée avec IBM® SPSS® Modeler Server. Vous pouvez désactiver cette fonctionnalité en ajoutant la ligne suivante au fichier *options.cfg* :

```
enable_database_caching, N
```

Cette action crée des fichiers temporaires sur le serveur et pas dans la base de données.

Pour afficher ou modifier les options de configuration de IBM SPSS Modeler Server

- ▶ Ouvrez le fichier *options.cfg* dans un éditeur de texte.
- ▶ Recherchez les options qui vous intéressent. Pour obtenir la liste complète des options, reportez-vous à [Configuration de SPSS Modeler Server sur p. 27.](#)
- ▶ Editez les valeurs à votre gré. Notez que toutes les valeurs de nom de chemin doivent utiliser la barre oblique normale (/) plutôt que la barre oblique inverse comme séparateur de nom de chemin.
- ▶ Enregistrez le fichier.
- ▶ Arrêtez et redémarrez SPSS Modeler Server pour que les modifications prennent effet. [Pour plus d'informations, reportez-vous à la section Démarrage et arrêt de IBM SPSS Modeler Server sur p. 21.](#)

Fermeture de connexions inutilisées à une base de données

Par défaut, IBM® SPSS® Modeler garde en cache au moins une connexion à une base de données, une fois que vous avez accédé à celle-ci. La session de la base de donnée reste ouverte même si les flux nécessitant un accès à la base de donnée ne sont pas exécutés.

Les connexions à la base de données mises en cache peuvent améliorer les temps d'exécution en évitant à SPSS Modeler de se reconnecter à la base de données chaque fois qu'un flux est exécuté. Toutefois dans certains environnements, il est important pour les applications de libérer les ressources de base de données aussi vite que possible. Si un trop grand nombre de sessions SPSS Modeler maintiennent des connexions inutiles à la base de donnée, cela peut conduire à un manque de ressources.

Vous pouvez éviter cela en désactivant l'option `SPSS Modelercache_connection` dans un fichier de configuration de base de données personnalisé. En outre cela peut aussi rendre SPSS Modeler plus résistant aux erreurs de connexion à la base de données (par exemple les expirations) qui peuvent se produire lorsque des connexions sont utilisées pendant de longues durées par une session SPSS Modeler.

Pour fermer les connexions inutilisées à une base de données :

- ▶ Recherchez le répertoire [*chemin d'installation du serveur*]/*config*.
- ▶ Ajoutez le fichier suivant (ou ouvrez-le s'il existe déjà) :
odbc-custom-properties.cfg
- ▶ Ajoutez la ligne suivante au fichier :
`cache_connection, N`
- ▶ Enregistrez le fichier, puis fermez-le.
- ▶ Redémarrez IBM® SPSS® Modeler Server pour que les modifications prennent effet.

Emplacement de la licence IBM SPSS Statistics

Pour permettre à IBM® SPSS® Modeler d'utiliser les nœuds Transformation Statistics , Modèle Statistics et Sortie Statistics, vous devez disposer d'une copie de IBM® SPSS® Statistics installée avec licence sur l'ordinateur où le flux est exécuté. Lors de 'exécution sur un serveur, la configuration de la licence doit être effectuée sur ce serveur.

Fenêtres

- ▶ A l'invite de commande, allez dans le répertoire IBM® SPSS® Modeler Server bin et exécutez :

```
statisticsutility -location =<path to IBM SPSS Statistics Server license file>
```

Où <path to SPSS Statistics Server license file> est le répertoire d'installation d'un serveur SPSS Statistics avec licence.

UNIX

- ▶ A l'invite de commande, allez dans le répertoire SPSS Modeler Server bin et exécutez :

```
./statisticsutility -location =<path to IBM SPSS Statistics Server license file>
```

Où <path to SPSS Statistics Server license file> est le répertoire d'installation d'un serveur SPSS Statistics avec licence.

Utilisation du protocole SSL pour sécuriser le transfert de données

Le protocole SSL (Secure Sockets Layer) permet de coder les données transférées entre deux ordinateurs. Le protocole SSL sécurise la communication entre les ordinateurs. SSL peut coder l'authentification d'un nom utilisateur/mot de passe et le contenu d'un échange entre un serveur et un client.

Fonctionnement de SSL

Le protocole SSL repose sur des clés publiques et privées de serveur, en plus d'un certificat de clé publique qui relie l'identité du serveur à sa clé publique.

- ▶ Lorsqu'un client se connecte à un serveur, le client authentifie ce serveur avec le certificat de clé publique.
- ▶ Le client génère ensuite un chiffre au hasard, code ce chiffre avec la clé publique du serveur et renvoie le message codé au serveur.
- ▶ Le serveur décode ce chiffre aléatoire avec sa clé privée.
- ▶ Avec le chiffre aléatoire, le serveur et le client créent les clés de session utilisées pour le codage et le décodage des informations suivantes.

Le certificat de clé publique est généralement signé par une autorité de certification. Les autorités de certification, comme VeriSign et Thawte, sont des organismes qui fournissent, authentifient et gèrent les informations de sécurité contenues dans les certificats de clé publique. L'autorité de certification a pour rôle principal de confirmer l'identité du serveur. Le certificat délivré par l'autorité de certification est généralement payant, mais des certificats signés peuvent également être délivrés.

Sécurisation des communications Client-Serveur et Serveur-Serveur avec le protocole SSL

Les étapes principales de sécurisation des communications client-serveur et serveur-serveur avec le protocole SSL sont :

- ▶ Installez OpenSSL sur l'ordinateur serveur.
- ▶ Obtenez et installez le certificat et les clés SSL.
- ▶ Activez et configurez SSL dans l'application d'administration du serveur (IBM® SPSS® Collaboration and Deployment Services Deployment Manager).
- ▶ Si nécessaire, installez le codage de type Unlimited Strength sur les ordinateurs client.
- ▶ Si vous utilisez un certificat signé, copiez-le sur l'ordinateur client.
- ▶ Ajoutez le certificat au magasin de clés du client.
- ▶ Donnez aux utilisateurs finaux l'instruction d'activer SSL lors de la connexion au serveur.

Remarque : Parfois, un produit serveur agit comme un client. Par exemple, IBM® SPSS® Statistics Server se connectant au IBM® SPSS® Collaboration and Deployment Services Repository. Dans ce cas, SPSS Statistics Server est le *client*.

Installation d'OpenSSL

Si OpenSSL n'est pas disponible sur le serveur, vous devez l'installer.

- ▶ Téléchargez OpenSSL depuis <http://www.openssl.org/>. Vérifiez que vous utilisez la version d'OpenSSL qui correspond à la version du serveur :

Table 4-1
Versions d'OpenSSL

Produit serveur	Version d'OpenSSL compatible
IBM® SPSS® Statistics 17-19 (non Linux® sur System z®) IBM® SPSS® Modeler 14.2 (non Linux® sur System z®)	0.9.8 et ses sous-versions (0.9.8a, 0.9.8b, etc.)
SPSS Statistics 19 (Linux® sur System z®) SPSS Modeler 14.2 (Linux® sur System z®)	1.0.0

Remarque : lorsque IBM® SPSS® Modeler Server pour UNIX, est configuré pour SSL, il cherche à s'associer à deux bibliothèques nommées (l'extension *.so* peut être spécifique à la plateforme) :

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

Lorsqu'une sous-version de SSL est utilisée (0.9.8a, 0.9.8b, etc.), vous devez créer des liens symboliques pour ces bibliothèques avec les versions *xxx.0.9.8* que SPSS Modeler Server attend.

- ▶ Suivez les instructions d'installation et de configuration du logiciel. Si vous créez vous-même OpenSSL, suivez les instructions générales suivantes :

Windows. OpenSSL doit être créé avec des DLL (qui sont multi-thread par défaut).

UNIX. OpenSSL doit prendre en charge plusieurs threads (fonction qui n'est pas toujours activée par défaut) et les bibliothèques partagées.

- ▶ Vérifiez que le système peut localiser les bibliothèques SSL :

Windows. Vérifiez que les modules OpenSSL sont inclus dans le système path.

UNIX. Ajoutez les bibliothèques SSL à la fin du chemin d'accès à la bibliothèque de chargement défini dans *modelersrv.sh*. Par exemple, sous Solaris, ajoutez cette ligne au script de démarrage du serveur :

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

Obtention et installation du certificat et des clés SSL

- ▶ Procurez-vous un certificat et un fichier de clé SSL. Ceci peut être fait de deux manières :
 - Achetez-les auprès d'une autorité de certification publique (comme VeriSign ou Thawte). L'autorité de certification publique signe le certificat pour vérifier le serveur qui l'utilise.
 - Générez les fichiers de clé et de certificat par le biais d'une autorité de certification signée interne. OpenSSL propose un outil de gestion des certificats adapté ou vous pouvez effectuer une recherche sur Internet pour trouver des instructions sur la création d'un certificat SSL signé.

- ▶ Copiez les fichiers de certificat et de clé dans un répertoire local du serveur. Les clés publiques et privées peuvent être stockées dans un seul fichier. Vérifiez que les droits d'accès sont définis pour refuser toute navigation dans ce répertoire.
- ▶ Installez le certificat et les clés SSL sur le serveur d'application. Pour des informations supplémentaires sur la façon dont les clés et le certificat interfonctionnent avec un serveur d'application spécifique, consultez la documentation d'origine du fournisseur. Veuillez noter qu'il pourrait être nécessaire d'ajouter le certificat et les clés au fichier Java keystore.

Activez et configurez SSL dans le IBM SPSS Collaboration and Deployment Services Deployment Manager

- ▶ Démarrez l'application d'administration du serveur (IBM® SPSS® Collaboration and Deployment Services Deployment Manager) et connectez-vous au serveur.
- ▶ Sur la page de configuration, définissez Secure Sockets Layer sur Oui .
- ▶ Dans fichier de clé publique SSL, indiquez le chemin d'accès complet au fichier de clé publique.
- ▶ Dans fichier de clé privée SSL, indiquez le chemin d'accès complet au fichier de clé privée.

Remarque : Si les clés publique et privée sont stockées dans un seul fichier, indiquez le même fichier dans fichier de clé publique SSL et dans fichier de clé privée SSL.

- ▶ A partir des menus, sélectionnez :
Fichier > Enregistrer
- ▶ Redémarrez le service ou le démon du serveur. Au redémarrage, le mot de passe SSL vous sera demandé. Sous Windows, vous pouvez sélectionner Se souvenir de ce mot de passe pour stocker le mot de passe de manière sécurisée. Cette option permet de ne pas avoir à saisir le mot de passe à chaque démarrage du serveur.

Configurer SSL dans IBM SPSS Modeler Server (systèmes UNIX uniquement)

Si un mot de passe est requis pour accéder au fichier de clé privée SSL sur les systèmes UNIX, vous pouvez le fournir d'une des manières suivantes :

- Spécifiez-le dans le fichier *options.cfg* en utilisant l'option `ssl_private_key_password`. [Pour plus d'informations, reportez-vous à la section SSL sur p. 32.](#)
- Ajoutez un commutateur au fichier *modelersrv.sh* pour afficher une invite permettant d'entrer le mot de passe au démarrage.

Pour cette dernière option, recherchez la ligne suivante dans le fichier *modelersrv.sh* :

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

Ajoutez le commutateur `-request_ssl_password` comme suit :

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

Installation du codage de type Unlimited Strength

Le codage de type US Export-Strength est activé dans le logiciel Java Runtime Environment livré avec le produit. Pour améliorer la sécurité de vos données, il est recommandé de mettre à niveau le codage vers le système Unlimited Strength.

- ▶ Téléchargez Java Cryptography Extension (JCE) Unlimited Strength Jurisdiction Policy Files 5.0 depuis <http://java.sun.com/javase/downloads/index.jsp>.
- ▶ Décompressez le fichier téléchargé.
- ▶ Copiez les deux fichiers `.jarlocal_policy.jar` et `US_export_policy.jar` dans `<installation folder>/jre/lib/security`, où `<dossier d'installation>` est le dossier dans lequel vous avez installé le produit.

Ajout du certificat au magasin de clés du client (pour les connexions au référentiel ou à IBM SPSS Modeler Server)

Remarque : Ignorez cette étape si vous utilisez un certificat signé par une autorité de certification.

Si vous utilisez SSL pour vous connecter au référentiel IBM® SPSS® Collaboration and Deployment Services ou à IBM® SPSS® Modeler Server et que vous utilisez des certificats auto-signés, vous devez ajouter le certificat au magasin de clés Java du client. Les étapes suivantes sont exécutées sur l'ordinateur *client*.

N'oubliez pas qu'un ordinateur serveur peut également agir comme un client. Par exemple, IBM® SPSS® Statistics Server se connectant au IBM® SPSS® Collaboration and Deployment Services Repository. Dans ce cas, SPSS Statistics Server est le *client*, et vous devez alors copier le certificat de IBM SPSS Collaboration and Deployment Services Repository Server sur SPSS Statistics Server.

- ▶ Copiez le fichier du certificat (.cer) du serveur SSL au répertoire `<install_dir>\jre\lib\security` sur l'ordinateur client.
- ▶ Dans ce même répertoire, faites un clic droit sur le fichier `cacerts` et sélectionnez Propriétés.
- ▶ Désélectionnez la case Lecture seule.
- ▶ Ouvrez une invite de commande et définissez le répertoire sur l'emplacement suivant où `<rép. install produit>` est le répertoire dans lequel vous avez installé le produit.

```
<rép. install produit>/jre/bin
```

- ▶ Entrez la commande ci-dessous :

```
keytool -import -alias <nom d'alias> -file <chemin du certificat> -keystore  
<chemin du magasin de clés>
```

Où `<nom d'alias>` est un alias arbitraire pour le certificat, `<chemin du certificat>` est le chemin complet vers le certificat, et `<chemin du magasin de clés>` est le chemin complet vers le magasin de clés Java, qui peut être `<rép. install produit>/lib/security/jssecacerts` ou `<rép. install produit>/lib/security/cacerts`.

- ▶ Lors de l'ouverture de l'invite, saisissez le mot de passe du magasin de clés, qui est `changeit` par défaut.
- ▶ Lors de l'ouverture de l'invite concernant l'approbation du certificat, entrez `yes`.

Indiquer aux utilisateurs finaux d'activer SSL

Lorsque les utilisateurs finaux se connectent au serveur via un produit client, ils doivent activer SSL dans la boîte de dialogue pour se connecter au serveur. N'oubliez pas d'avertir vos utilisateurs de sélectionner la case à cocher appropriée. Par exemple dans IBM® SPSS® Modeler, les utilisateurs doivent sélectionner la case *Garantir une connexion sécurisée* dans la boîte de dialogue *Référentiel : Serveur*, accessible depuis l'option *Référentiel* dans le menu *Outils*.

Configuration du préfixe d'URL

Si IBM® SPSS® Collaboration and Deployment Services Repository est configuré pour l'accès SSL, la valeur du paramètre de configuration du préfixe de l'URL doit être modifiée comme suit :

1. Connectez-vous au référentiel à l'aide de la console utilisant le navigateur.
2. Ouvrez l'option de configuration *Préfixe d'URL*.
Configuration > Installation > Préfixe d'URL
3. Définissez la valeur du préfixe sur `https` à la place de `http` et définissez la valeur du port sur le numéro de port SSL. Par exemple :

```
[default]
http://<hostname>:<port>
[SSL-enabled]
https://<hostname>:<SSLport>
```

Sécurisation de LDAP avec SSL

Le protocole LDAP (Lightweight Directory Access Protocol) est une norme IETF (Internet Engineering Task Force) qui permet d'échanger des informations entre les répertoires et les bases de données réseau contenant n'importe quelle type d'informations. Pour les systèmes nécessitant d'avantage de sécurité, les fournisseurs LDAP comme Active Directory de Microsoft, peuvent opérer avec SSL (Secure Socket Layer), tant que le serveur Web ou d'application prend en charge LDAP avec SSL. L'utilisation de SSL avec LDAP permet de s'assurer que les mots de passe de connexion, les informations d'application et d'autres données sensibles ne soient ni piratés, ni compromis, ni volés.

L'exemple suivant explique comment activer les LDAPS utilisant Active Directory de Microsoft comme fournisseur de sécurité. Pour des informations détaillées sur une des étapes ou pour obtenir des détails sur une publication spécifique du fournisseur de sécurité, consultez la documentation d'origine du fournisseur.

1. Vérifiez que Active Directory et que l'autorité de certification d'entreprise sont installés et fonctionnent correctement.

2. Utilisez l'autorité de certification pour générer un certificat et importez ce certificat dans le magasin de certificats de l'installation IBM® SPSS® Collaboration and Deployment Services Deployment Manager. Ceci permet d'établir une connexion LDAPS entre le IBM® SPSS® Collaboration and Deployment Services Repository et un serveur de Active Directory.

Pour configurer Deployment Manager avec des connexions Active Directory sécurisées, vérifiez qu'une connexion au référentiel existe.

3. Lancez le IBM® SPSS® Collaboration and Deployment Services Deployment Manager.
4. Dans le menu Outils, sélectionnez Administration du serveur.
5. Connectez-vous à un serveur administré défini au préalable.
6. Double-cliquez sur l'icône Configuration pour développer la hiérarchie du serveur.
7. Double-cliquez sur l'icône Fournisseurs de sécurité pour développer la hiérarchie.
8. Double-cliquez sur le fournisseur de sécurité Active Directory.
9. Saisissez les valeurs de configuration pour l'instance de Active Directory sur laquelle les certificats de sécurité sont installés.
10. Sélectionnez la case Utilisez SSL.
11. Notez le nom dans le champ Utilisateur du domaine. Les informations de connexions suivantes qui utilisent Active Directory sont authentifiées avec SSL.

Pour des informations supplémentaires sur l'installation, la configuration et l'implémentation des protocoles LDAPS sur un serveur d'application particulier, consultez la documentation d'origine du fournisseur.

Journal du serveur

IBM® SPSS® Modeler Server conserve une trace des opérations importantes dans un fichier journal nommé *messages.log*. Par défaut, ce fichier se trouve dans *[dossier d'installation du serveur]/log*, mais vous pouvez définir un autre emplacement. Lors du démarrage de SPSS Modeler Server, le fichier journal est mis à jour pour indiquer l'heure du démarrage et le numéro du port d'écoute. Le fichier journal est ensuite mis à jour avec le résultat des tentatives de connexion des clients au serveur.

Par défaut, le serveur écrit les événements suivants dans le fichier journal :

- Serveur démarré
- Serveur arrêté
- Connexion acceptée
- Connexion refusée
- Session démarrée
- Session arrêtée
- Erreur système

- Connexion réussie
- Echec de la connexion

Si vous souhaitez enregistrer des informations sur les événements d'accès à la base de données, définissez la consignation des informations sur l'accès à la base de données et la consignation des erreurs de base de données sur *yes* (oui) dans la configuration de SPSS Modeler Server. Les événements d'accès à la base de données suivants seront consignés :

- Connexion à la base de données réussie
- Echec de la connexion à la base de données
- Déconnexion de la base de données

Si vous souhaitez enregistrer des informations sur les événements d'accès aux fichiers, définissez la consignation des informations sur l'accès aux fichiers et la consignation des erreurs d'accès aux fichiers sur *yes* (oui) dans la configuration de SPSS Modeler Server. Les événements d'accès aux fichiers suivants seront consignés :

- Ouverture/création d'un fichier
- Fermeture d'un fichier

Voici un exemple de fichier journal (des numéros de ligne ont été ajoutés) :

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

Tous les éléments du fichier journal sont associés à un horodatage. Les lignes 1 et 2 ont été ajoutées au fichier journal lorsque SPSS Modeler Server a été démarré. La ligne 3 a été ajoutée lorsqu'un utilisateur de l'ordinateur *localhost*, sur le numéro de port *1827*, a tenté de se connecter au serveur. La ligne 4 indique qu'une session de serveur a été démarrée avec cet utilisateur. La ligne 5 indique que la connexion de l'utilisateur a échoué. Un peu plus tard, l'utilisateur a mis fin à la session. Une demi-heure plus tard, une autre connexion a été tentée à partir de *localhost* et une session a été démarrée. Cette fois, la connexion a réussi. Une demi-heure plus tard, l'utilisateur a mis fin à la session et à la connexion.

Considérations relatives aux performances

Dans le cadre de l'analyse des données, les performances réelles dépendent d'un certain nombre de facteurs, par exemple la configuration du serveur et de la base de données ou l'ordre des noeuds individuels au sein d'un flux. De façon générale, vous pouvez obtenir les meilleures performances en effectuant les opérations suivantes :

- Stockez vos données dans un SGBD (système de gestion de base de données) et utilisez, dans la mesure du possible, l'optimisation SQL.
- Utilisez un matériel qui répond au minimum aux recommandations données dans la rubrique [le chapitre 2](#).
- Vérifiez que les paramètres de performances et d'optimisation de votre serveur sont correctement configurés. Certains de ces paramètres sont également disponibles sur le client mais ne s'appliquent que lorsque l'application IBM® SPSS® Modeler est utilisée en mode autonome. Lorsque IBM® SPSS® Modeler est connecté à un serveur, il utilise automatiquement ses paramètres de performances et d'optimisation.
- Optimisez les flux afin d'obtenir des performances maximales. [Pour plus d'informations, reportez-vous à la section Remarques relatives aux performances des flux et des noeuds dans le chapitre 13 dans *Guide de l'utilisateur de IBM SPSS Modeler 14.2*.](#)

Vous trouverez davantage d'informations sur chacun de ces facteurs de performance dans les sections suivantes.

Paramètres de performances et d'optimisation du serveur

Vous pouvez configurer certains paramètres IBM® SPSS® Modeler Server en vue d'optimiser les performances. Vous pouvez ajuster ces paramètres à l'aide de l'application IBM® SPSS® Collaboration and Deployment Services Deployment Manager fournie avec SPSS Modeler Server. [Pour plus d'informations, reportez-vous à la section Utilisation de IBM SPSS Modeler Administration Console dans le chapitre 4 sur p. 24.](#)

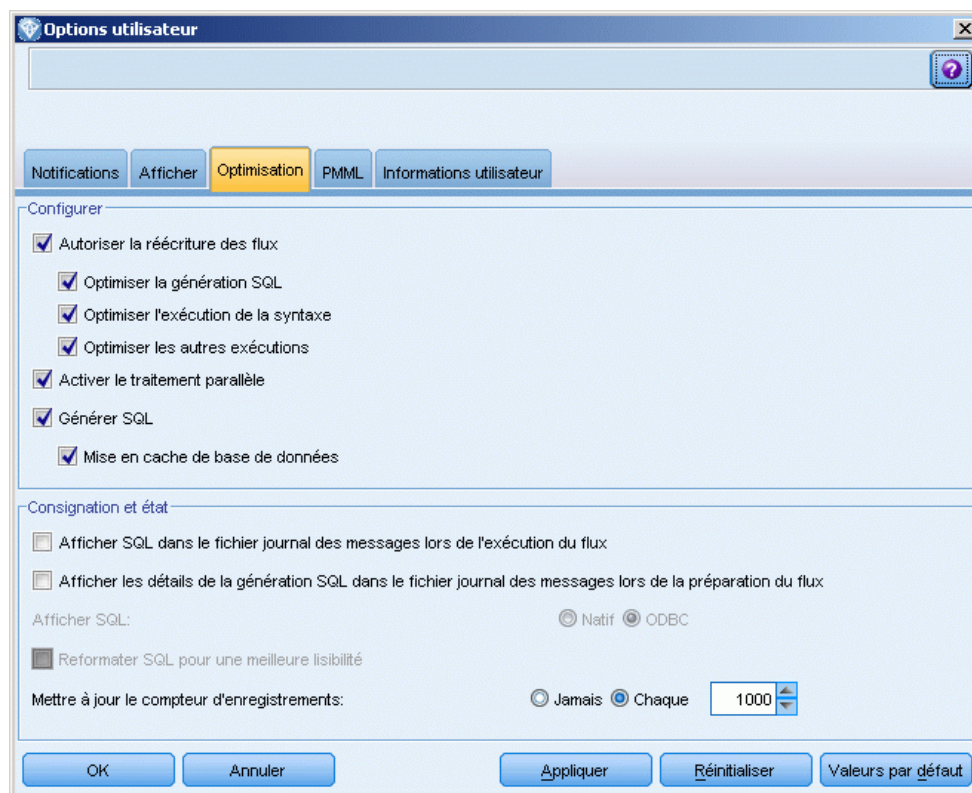
Dans la fenêtre de configuration Deployment Manager, les paramètres sont regroupés sous l'en-tête Performances et optimisation. Les paramètres sont préconfigurés sur des performances optimales pour la plupart des installations. Toutefois, il se peut que vous deviez les ajuster en fonction de votre matériel, de la taille de vos ensembles de données et du contenu de vos flux. [Pour plus d'informations, reportez-vous à la section Performances/Optimisation dans le chapitre 4 sur p. 29.](#)

Paramètres de performances et d'optimisation du client

L'onglet Optimisation de la boîte de dialogue Options utilisateur vous permet d'optimiser les performances du flux. Les paramètres de performances et d'optimisation de IBM® SPSS® Modeler Server (si ce dernier est utilisé) annulent ceux définis pour le client. *Cet onglet n'est disponible que si la connectivité SPSS Modeler Server a été activée sur l'ordinateur IBM® SPSS® Modeler.*

Remarque : La modélisation de base de données et l'optimisation SQL requièrent l'activation de la connectivité à SPSS Modeler Server sur l'ordinateur SPSS Modeler. Avec ce paramètre activé, vous pouvez accéder aux algorithmes de la base de données, effectuer le push back de SQL directement depuis SPSS Modeler et accéder à SPSS Modeler Server. Pour vérifier l'état de la licence actuelle, cliquez sur A propos de dans le menu Aide de SPSS Modeler. [Pour plus d'informations, reportez-vous à la section Connexion au IBM SPSS Modeler Server dans le chapitre 3 dans Guide de l'utilisateur de IBM SPSS Modeler 14.2.](#)

Figure 5-1
Boîte de dialogue Options utilisateur, onglet Optimisation.



Remarque : La prise en charge des répercussions et de l'optimisation SQL dépend du type de base de données utilisé. Pour obtenir les informations les plus récentes sur les bases de données et pilotes ODBC pris en charge et testés pour une utilisation avec IBM® SPSS® Modeler 14.2, consultez les matrices de compatibilité des produits sur le site Web de support technique de l'entreprise à l'adresse <http://www.ibm.com/support>.

Autoriser la réécriture des flux. Sélectionnez cette option pour activer la réécriture de flux dans SPSS Modeler. Deux types de réécriture sont disponibles. Vous pouvez sélectionner l'une de ces options ou les deux. La réécriture de flux réorganise les noeuds d'un flux en arrière-plan, sans pour autant modifier sa sémantique, afin qu'il soit exécuté de manière plus efficace.

- **Optimiser la génération SQL.** Cette option permet de réorganiser les noeuds du flux afin d'augmenter le nombre d'opérations pouvant être répercutées, à l'aide d'une génération SQL, pour y être exécutées dans la base de données. Lorsqu'il détecte un noeud qui ne peut pas être affiché sous forme de code SQL, le programme d'optimisation recherche plus en avant tout noeud qui peut l'être et qu'il est possible de ramener avant le noeud problématique sans risquer de modifier la sémantique du flux. Non seulement la base de données peut effectuer les opérations plus efficacement que SPSS Modeler, mais de telles répercussions contribuent également à réduire le volume des données renvoyées à SPSS Modeler pour être traitées. Vous réduisez ainsi le trafic sur le réseau et accélérez les opérations de flux. Veuillez noter que la case Générer SQL doit être sélectionnée pour que l'optimisation SQL ait le moindre effet. [Pour plus d'informations, reportez-vous à la section Optimisation SQL dans le chapitre 6 sur p. 49.](#)
- **Optimiser l'exécution de syntaxe.** Cette méthode de réécriture de flux augmente l'efficacité des opérations qui incorporent plus d'un noeud contenant une syntaxe IBM® SPSS® Statistics. L'optimisation est réalisée en combinant des commandes de syntaxe dans une opération unique, au lieu d'exécuter chacune d'elle en tant qu'opération séparée.
- **Optimiser les autres exécutions.** Cette méthode de réécriture de flux augmente l'efficacité des opérations exécutées (c'est-à-dire des opérations dont l'exécution ne peut pas être déléguée à la base de données). L'optimisation est obtenue en réduisant dès que possible la quantité de données dans le flux. Tout en conservant l'intégrité des données, le flux est réécrit de façon à rapprocher les opérations de la source de données, ceci afin de réduire le volume des données en aval pour les opérations coûteuses, telles que les jointures.

Activer le traitement parallèle. Lorsque vous exécutez ce programme sur un ordinateur doté de plusieurs processeurs, cette option permet au système d'équilibrer la charge entre les différents processeurs et d'augmenter ainsi les performances. Le traitement parallèle peut s'avérer bénéfique en cas d'utilisation de plusieurs noeuds ou des noeuds individuels suivants : C5.0, Fusionner (par clé), Trier, Discrétiser (méthodes des rangs et des quantiles) et Agréger (à l'aide d'un ou de plusieurs champs-clés).

Générer SQL. Sélectionnez cette option pour activer la génération SQL et répercuter ainsi les opérations de flux dans la base de données à l'aide du code SQL pour générer les processus d'exécution ; cette opération peut améliorer les performances. Pour augmenter encore les performances, vous pouvez également sélectionner l'option Optimiser la génération SQL, afin d'accroître au maximum le nombre d'opérations répercutées dans la base de données. Si les opérations d'un noeud ont été retransmises à la base de données, le noeud apparaît en violet au cours de l'exécution du flux.

- **Mise en cache de base de données.** Dans le cas des flux exécutés dans la base de données qui génèrent du code SQL, les données peuvent être mises en mémoire cache en milieu de flux dans une table temporaire de la base de données, plutôt que dans le système de fichiers. Vous pouvez combiner cela à la fonction d'optimisation SQL et accroître ainsi les performances de manière significative. Par exemple, la sortie d'un flux qui fusionne plusieurs tables afin de créer une vue de Data mining peut être mise en mémoire cache et réutilisée selon les besoins. Une fois la mise en mémoire cache de base de données activée, il vous suffit de cliquer à

l'aide du bouton droit de la souris sur un noeud non terminal afin de mettre les données en mémoire cache à ce stade du flux ; le cache est automatiquement créé dans la base de données lors de l'exécution suivante du flux. Le code SQL peut ainsi être généré pour les noeuds en aval, ce qui améliore encore les performances. Vous pouvez également désactiver cette option si besoin est, par exemple lorsque des règles ou des permissions empêchent l'écriture des données dans la base de données. Si la mise en mémoire cache de base de données ou l'optimisation SQL n'est pas activée, le cache est écrit dans le système de fichiers. [Pour plus d'informations, reportez-vous à la section Options de mise en cache des noeuds dans le chapitre 5 dans *Guide de l'utilisateur de IBM SPSS Modeler 14.2*.](#)

Remarque : En raison de différences d'implémentation SQL mineures, les flux exécutés dans une base de données peuvent renvoyer des résultats légèrement différents que ceux renvoyés lors de leur exécution dans SPSS Modeler. Pour ces mêmes raisons, ces différences peuvent également varier en fonction du fournisseur de base de données.

Afficher SQL dans le fichier journal des messages lors de l'exécution du flux. Détermine l'insertion dans le journal des messages du code SQL généré lors de l'exécution du flux.

Afficher les détails de la génération SQL dans le fichier journal des messages lors de la préparation du flux. Au cours de la prévisualisation du flux, détermine l'insertion dans le journal des messages d'un aperçu du code SQL susceptible d'être généré.

Afficher SQL. Indique si le code SQL affiché dans le journal doit contenir des fonctions SQL natives ou des fonctions ODBC standard (générées par SPSS Modeler au format {fn FUNC(...)}). Les premières reposent sur certaines fonctionnalités du pilote ODBC, qui ne sont peut-être pas implémentées. Ainsi, cette commande n'aurait aucun impact dans SQL Server.

Reformater SQL pour une meilleure lisibilité. Indique si le SQL affiché dans le journal doit être formaté pour être lu.

Mettre à jour le compteur d'enregistrements. Indique le moment où les enregistrements doivent être enregistrés à mesure qu'ils arrivent dans les noeuds terminaux. Indiquez le chiffre déterminant la mise à jour de l'état tous les N enregistrements.

Cliquez sur Valeurs par défaut pour revenir aux paramètres système par défaut de cet onglet.

Utilisation et optimisation de la base de données

Serveur de base de données. Si possible, créez une instance de base de données dédiée au Data mining de façon à ce que le serveur de production ne subisse pas l'impact des requêtes IBM® SPSS® Modeler. Les instructions SQL générées par SPSS Modeler peuvent être exigeantes ; plusieurs tâches sur l'ordinateur IBM® SPSS® Modeler Server peuvent exécuter SQL dans la même base de données.

Exploration des bases de données. De nombreux fournisseurs de base de données offrent des extensions de Data mining pour leurs produits. Ces extensions permettent des activités de Data mining (telles que la création de modèle ou le scoring) à effectuer avec le serveur de base de données ou sur un serveur dédié distinct. Les fonctions d'exploration des bases de données de SPSS Modeler complètent et étendent ses capacités de création SQL et offrent une manière de traiter les extensions de la base de données spécifique du fournisseur. Dans certains cas, cette approche évite le temps système potentiellement coûteux lié au transfert de données

entre SPSS Modeler et la base de données. D'autre part, la mise en cache de base de données peut augmenter davantage les bénéfices. Pour plus d'informations, reportez-vous au fichier *DatabaseMiningGuide.pdf*, disponible dans le dossier *Documentation* sur le disque d'installation IBM® SPSS® Modeler.

Optimisation SQL

Pour obtenir les meilleures performances, tentez systématiquement d'optimiser le volume du code SQL généré afin de tirer parti des performances et de l'évolutivité de la base de données. Seules les parties du flux qui ne peuvent pas être compilées en langage SQL doivent être exécutées au sein de IBM® SPSS® Modeler Server. Pour plus d'informations, consultez [le chapitre 6](#).

Envoi de données basées sur des fichiers

Les données qui ne sont pas stockées dans une base de données ne peuvent bénéficier de l'optimisation SQL. Si les données à analyser ne se trouvent pas déjà dans une base de données, vous pouvez les y envoyer à l'aide d'un noeud de sortie SGBD. Vous pouvez également utiliser ce noeud pour stocker des ensembles de données intermédiaires issus de la préparation des données ainsi que les résultats du déploiement. [Pour plus d'informations, reportez-vous à la section Export SGBD - Options avancées dans le chapitre 7 dans IBM SPSS Modeler 14.2 Noeuds source, exécution et de sortie.](#)

IBM® SPSS® Modeler fonctionne avec les chargeurs externes d'un grand nombre de systèmes de base de données connus. Plusieurs scripts sont inclus dans le logiciel ; ils sont disponibles (avec une documentation) dans le sous-répertoire */scripts* dans votre dossier d'installation de SPSS Modeler.

Le tableau suivant présente les avantages potentiels, en matière de performances, du chargement en masse. Les chiffres illustrent la durée d'exportation de 250 000 enregistrements et de 21 champs vers une base de données Oracle. Le module de chargement externe est l'utilitaire *sqlldr* d'Oracle.

Table 5-1

Avantages, en termes de performances, du chargement en masse

Option d'exportation	Durée (en secondes)
Par défaut (ODBC)	409
Chargement en masse par ODBC	52
Chargement en masse via un module de chargement externe	33

Optimisation SQL

L'une des fonctionnalités les plus efficaces de IBM® SPSS® Modeler réside dans sa capacité à exécuter de nombreuses opérations de préparation des données et de Data mining directement dans la base de données. Grâce à la création de code SQL pouvant être répercuté dans la base de données pour y être exécuté, de nombreuses opérations, comme l'échantillonnage, le tri, le calcul de nouveaux champs et la génération de certains types de graphique, peuvent être effectuées dans la base de données plutôt que sur l'ordinateur IBM® SPSS® Modeler ou IBM® SPSS® Modeler Server. Lorsque vous travaillez avec des ensembles de données volumineux, ces **répercussions** peuvent améliorer considérablement les performances, et ce, de plusieurs manières :

- En réduisant le volume des résultats à transférer du SGBD vers SPSS Modeler. La lecture de résultats trop nombreux via un pilote ODBC risque d'entraîner des problèmes de performances au niveau de l'E/S réseau ou du pilote. C'est la raison pour laquelle les opérations qui tirent le plus parti de l'optimisation SQL sont la sélection et l'agrégation de lignes et de colonnes (noeuds Sélectionner, Echantillon, Agréger), lesquelles réduisent généralement la taille de l'ensemble de données à transférer. Pour améliorer davantage les performances, il est également possible de mettre les données en cache dans une table temporaire de la base de données, au niveau de points critiques du flux (par exemple, après un noeud Fusionner ou Sélectionner).
- En exploitant les performances et l'évolutivité de la base de données. Efficacité accrue : les SGBD bénéficient en effet souvent de la technique du traitement parallèle, d'un matériel plus puissant, d'une gestion plus évoluée du stockage sur disque et de la présence d'index.

Etant donné ces avantages, SPSS Modeler est conçue pour optimiser le volume du code SQL généré par chaque flux de façon à ce que seules les opérations qui ne peuvent pas être compilées en code SQL soient exécutées par SPSS Modeler Server. Toutefois, le langage SQL standard (SQL-92) ne permet pas de tout exprimer. Certaines opérations risquent donc de ne pas être prises en charge. [Pour plus d'informations, reportez-vous à la section Astuces relatives à l'optimisation de la génération SQL sur p. 55.](#)

Remarque : En raison de différences d'implémentation SQL mineures, les flux exécutés dans une base de données peuvent renvoyer des résultats légèrement différents que lors de leur exécution dans SPSS Modeler. Pour ces mêmes raisons, ces différences peuvent également varier en fonction du fournisseur de base de données. Par exemple, en fonction de la configuration de la base de données en ce qui concerne la distinction des majuscules/minuscules pour la comparaison et la collecte de chaînes, les flux SPSS Modeler exécutés en utilisant les répercussions SQL peuvent générer des résultats différents des flux exécutés sans les répercussions SQL. Contactez l'administrateur de votre base de données pour obtenir des conseils sur la configuration de votre base de données. Pour optimiser la compatibilité avec SPSS Modeler, les comparaisons de chaînes des bases de données devraient distinguer les majuscules et les minuscules.

Remarque : La modélisation de base de données et l'optimisation SQL requièrent l'activation de la connectivité à SPSS Modeler Server sur l'ordinateur SPSS Modeler. Avec ce paramètre activé, vous pouvez accéder aux algorithmes de la base de données, effectuer le push back de

SQL directement depuis SPSS Modeler et accéder à SPSS Modeler Server. Pour vérifier l'état de la licence actuelle, cliquez sur A propos de dans le menu Aide de SPSS Modeler. [Pour plus d'informations, reportez-vous à la section Connexion au IBM SPSS Modeler Server dans le chapitre 3 dans *Guide de l'utilisateur de IBM SPSS Modeler 14.2*.](#)

Exigences de base de données

Pour obtenir les informations les plus récentes sur les bases de données et pilotes ODBC pris en charge et testés pour une utilisation avec SPSS Modeler, consultez les matrices de compatibilité des produits sur le site Web de support technique de l'entreprise à l'adresse <http://www.ibm.com/support>.

Vous pouvez obtenir des performances accrues en utilisant la modélisation de base de données. [Pour plus d'informations, reportez-vous à la section Modélisation de base de données - Présentation dans le chapitre 2 dans *Guide d'exploration de base de données IBM SPSS Modeler 14.2*.](#)

Fonctionnement de la génération SQL

Les fragments initiaux d'un flux issu des noeuds source de la base de données sont les principales cibles de la génération SQL. Si un noeud ne peut pas faire l'objet d'une compilation en code SQL, les données sont extraites de la base de données et un traitement consécutif est exécuté par IBM® SPSS® Modeler Server.

Le processus de génération SQL, qui a lieu pendant la préparation du flux et avant l'exécution, se déroule comme suit :

- Le serveur réorganise les flux afin de déplacer les noeuds en aval dans la "zone SQL" s'il est admis que cette action est sûre. (Il est possible de désactiver cette fonctionnalité sur le serveur.)
- Dans le sens noeuds source vers noeuds terminaux, les expressions SQL sont construites de manière incrémentielle. Cette phase s'interrompt lorsqu'elle rencontre un noeud qui ne peut pas faire l'objet d'une conversion SQL ou lorsque le noeud terminal (noeud Table, noeud Graphiques, etc.) est converti en code SQL. A la fin de cette phase, chaque noeud comporte une instruction SQL si le noeud et ses prédécesseurs sont dotés d'un équivalent SQL.
- Dans le sens noeuds dotés des équivalents SQL les plus complexes vers noeuds source, l'instruction SQL fait l'objet d'une vérification de validité. L'instruction SQL validée est choisie pour l'exécution.
- Les noeuds pour lesquels toutes les opérations ont généré un code SQL apparaissent surlignés en violet dans l'espace de travail de flux. Sur la base des résultats, vous pouvez réorganiser davantage le flux, si nécessaire, pour bénéficier pleinement de l'exécution de la base de données. [Pour plus d'informations, reportez-vous à la section Astuces relatives à l'optimisation de la génération SQL sur p. 55.](#)

Cibles des améliorations

L'optimisation SQL améliore les performances au niveau de plusieurs opérations de données :

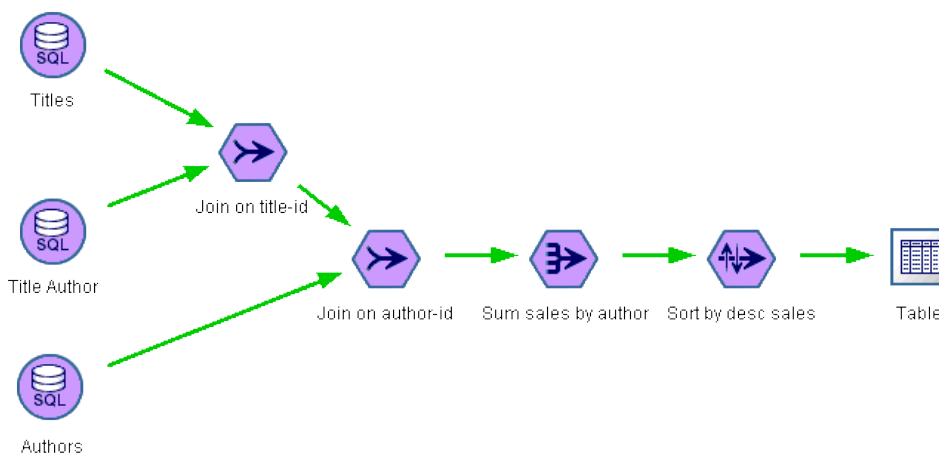
- **Jointures (fusion par clés).** Les opérations de jointure peuvent accroître l'optimisation au sein des bases de données.
- **Agrégation.** Les noeuds Agréger, Proportion et Relations utilisent tous l'agrégation pour produire leurs résultats. Les données récapitulatives utilisent beaucoup moins de bande passante que les données d'origine.
- **Sélection :** La sélection d'enregistrements sur la base de critères spécifiques réduit le nombre d'enregistrements.
- **Tri.** Le tri des enregistrements est une activité très consommatrice de ressources qui est exécutée plus efficacement dans une base de données.
- **Calcul de champ.** Les nouveaux champs sont générés de façon plus efficace dans une base de données.
- **Projection de champ.** SPSS Modeler Server extrait uniquement, de la base de données, les champs nécessaires aux traitements suivants ; les exigences en termes de bande passante et de mémoire sont ainsi réduites. Idem pour les champs superflus des fichiers plats : même si le serveur doit lire ces champs, il ne leur affecte pas d'espace de stockage.
- **Scoring.** Le code SQL peut être généré à partir de modèles d'arbre décision, d'ensembles de règles, de régression linéaire et des modèles factoriels générés.

Exemple de génération SQL

Le flux suivant relie trois tables de base de données par des opérations-clés, puis effectue une agrégation et un tri.

Figure 6-1

Flux optimisé avec noeuds violets indiquant les répercussions SQL (opérations effectuées dans la base de données)



Code SQL généré

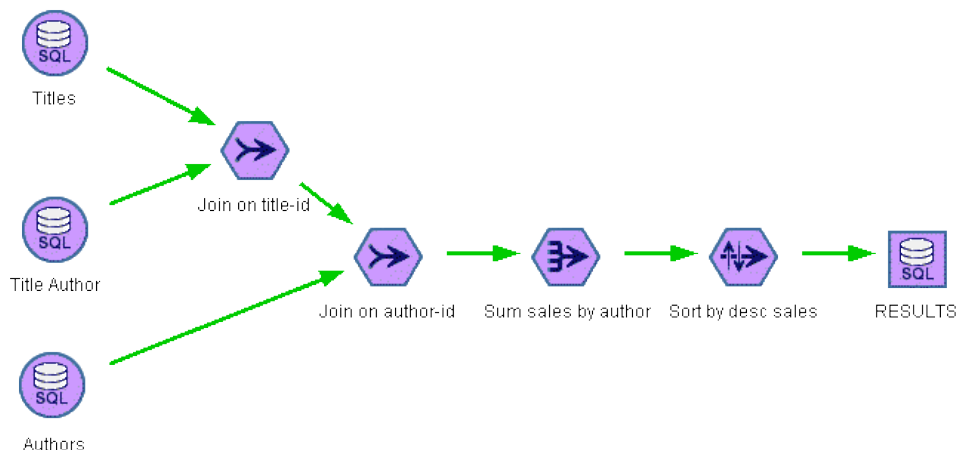
Le code SQL généré pour ce flux est le suivant :

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

Exécution du flux

Lorsque le flux se termine par un nœud Export SGBD, le flux peut être intégralement exécuté dans la base de données.

Figure 6-2
Intégralité du flux exécutée dans la base de données

**Configuration de l'optimisation SQL**

- ▶ Installez un pilote ODBC et configurez une source de données pour la base de données à utiliser. [Pour plus d'informations, reportez-vous à la section Accès aux données dans le chapitre 2 sur p. 11.](#)
- ▶ Créez un flux utilisant un nœud source pour extraire les données de la base de données.
- ▶ Vérifiez que la génération SQL est activée sur le client et le serveur, s'il y a lieu. Elle est activée par défaut sur les deux.

Pour activer l'optimisation SQL sur le client

- ▶ Dans le menu Outils, sélectionnez Options utilisateur.
- ▶ Cliquez sur l'onglet Optimisation. Pour activer l'optimisation SQL, sélectionnez Générer SQL. Vous avez également la possibilité de sélectionner d'autres paramètres pour améliorer les performances. [Pour plus d'informations, reportez-vous à la section Paramètres de performances et d'optimisation du client dans le chapitre 5 sur p. 45.](#)

Pour activer l'optimisation SQL sur le serveur

Etant donné que les paramètres du serveur remplacent ceux du client, vous devez activer les paramètres de réécriture des flux et de génération SQL automatique de la configuration serveur. Pour plus d'informations sur le mode de modification des paramètres de IBM® SPSS® Modeler Server, reportez-vous à la section [Performances/Optimisation dans le chapitre 4.](#)

Pour activer l'optimisation lors du scoring de modèles

Pour les besoins de scoring, la génération SQL doit être activée séparément sur chaque noeud de modélisation, quels que soient les paramètres serveur ou client. Il en est ainsi parce que certains modèles génèrent des expressions SQL extrêmement complexes qui pourraient ne pas être évaluées efficacement dans la base de données. La base de données peut signaler des erreurs lors de l'exécution de l'instruction SQL générée, en raison de la taille ou de la complexité de l'instruction SQL.

Une certaine quantité d'essais et d'erreurs peuvent être nécessaires pour déterminer si la génération d'instructions SQL améliore les performances d'un modèle donné. Activez la génération SQL dans l'onglet Paramètres, après avoir ajouté un modèle généré à un flux. [Pour plus d'informations, reportez-vous à la section Utilisation de nuggets de modèle dans les flux dans le chapitre 3 dans Noeuds de modélisation IBM SPSS Modeler 14.2.](#)

Prévisualisation du code SQL généré

Avant d'exécuter le code SQL généré dans la base de données, vous pouvez le prévisualiser dans le journal des messages. Cette étape peut s'avérer utile pour le débogage. Elle permet en outre d'exporter le code SQL généré pour l'éditer ou l'exécuter ultérieurement dans la base de données. Elle indique également les noeuds qui seront répercutés dans la base de données ; grâce à cette information, vous pouvez déterminer si le flux peut être réorganisé afin d'améliorer les performances.

- ▶ Veillez à sélectionner les options Afficher SQL dans le fichier journal des messages lors de l'exécution du flux et Afficher les détails de la génération SQL dans le fichier journal des messages lors de la préparation du flux dans la boîte de dialogue Options utilisateur. [Pour plus d'informations, reportez-vous à la section Paramètres de performances et d'optimisation du client dans le chapitre 5 sur p. 45.](#)
- ▶ Dans l'espace de travail de flux, sélectionnez le noeud ou le flux à prévisualiser.

- Cliquez sur le bouton de prévisualisation SQL de la barre d'outils.

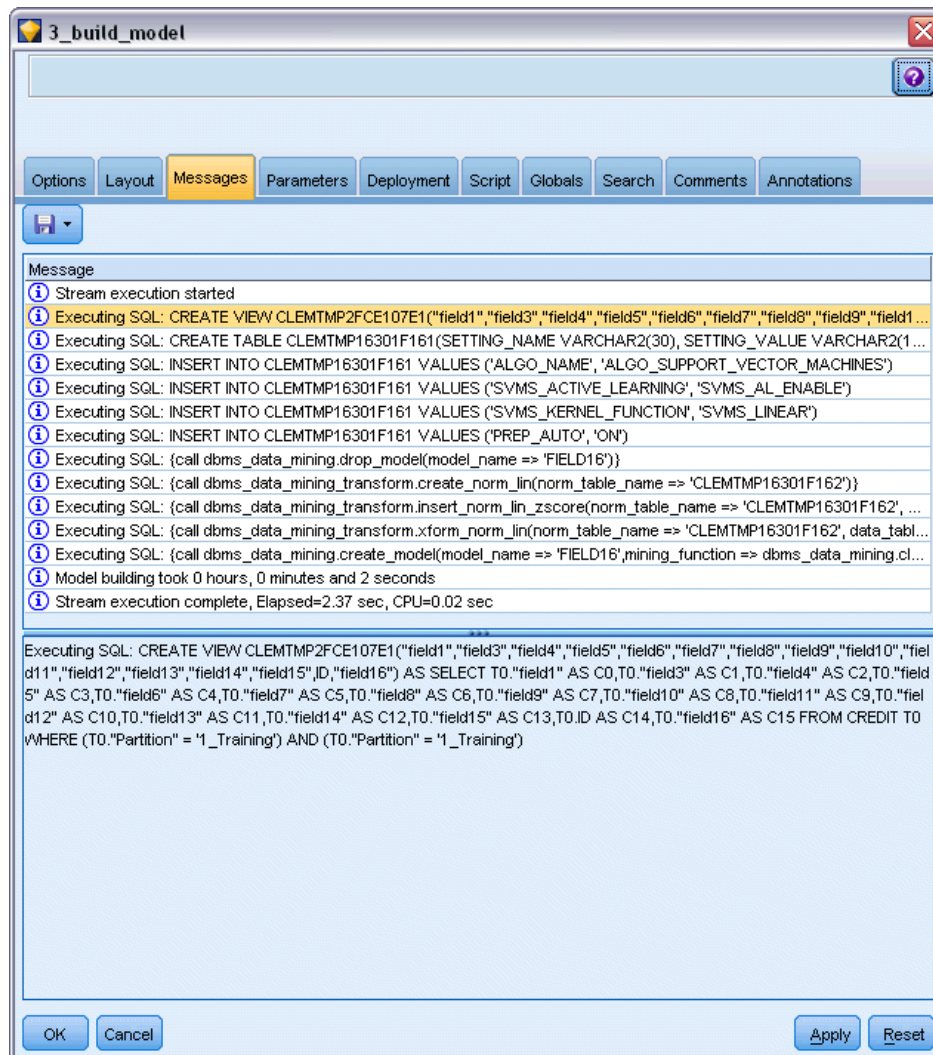
Tous les noeuds pour lesquels du code SQL est généré (et qui seront répercutés dans la base de données lors de l'exécution du flux) apparaissent en violet dans l'espace de travail de flux.

Figure 6-3
Bouton de prévisualisation SQL



- Pour prévisualiser le code SQL généré, choisissez les options suivantes :
Outils > Propriétés du flux > Messages...

Figure 6-4
Code SQL généré affiché dans le journal des messages



Affichage SQL pour des nuggets de modèle

Pour certains modèles, il est possible de générer une instruction SQL pour le nugget de modèle, en répercutant l'étape de scoring de modèle dans la base de données. L'utilisation principale de cette caractéristique n'est pas d'améliorer les performances, mais d'autoriser des flux contenant ces nuggets à voir l'intégralité de leur instruction SQL répercutée. [Pour plus d'informations, reportez-vous à la section Noeuds prenant en charge la génération SQL sur p. 57.](#)

Pour afficher l'instruction SQL d'un nugget de modèle qui prend en charge la génération SQL :

- ▶ Sélectionnez l'onglet Paramètres dans le nugget de modèle.
- ▶ Sélectionnez l'option appropriée, Générer avec/sans prise en charge des valeurs manquantes ou Générer SQL pour ce modèle.
- ▶ Dans le menu du nugget de modèle, sélectionnez :
Fichier > Exporter SQL
- ▶ Enregistrez le fichier.
- ▶ Ouvrez le fichier pour afficher l'instruction SQL.

Astuces relatives à l'optimisation de la génération SQL

Pour obtenir les meilleures performances possibles au niveau de l'optimisation SQL, prenez en considération les éléments suivants.

Ordre du flux. La génération SQL peut être arrêtée lorsque la fonction du noeud n'a pas d'équivalent sémantique dans le langage SQL. Cette absence d'équivalent tient au fait que la fonctionnalité de Data mining de IBM® SPSS® Modeler est plus riche que les opérations classiques de traitement de données prises en charge par le langage SQL standard. Dans ce cas de figure, la génération SQL est également supprimée pour les noeuds en aval. Par conséquent, vous pouvez améliorer considérablement les performances du système en réorganisant les noeuds et en plaçant les opérations qui bloquent le code SQL le plus en aval possible. L'optimiseur SQL peut, dans une certaine mesure, procéder à une réorganisation automatique (vérifiez simplement que la fonction de réécriture de flux est activée), mais d'autres améliorations sont possibles. Le noeud Sélectionner s'applique parfaitement à ce cas ; il est souvent possible de le déplacer vers l'avant du flux. [Pour plus d'informations, reportez-vous à la section Noeuds prenant en charge la génération SQL sur p. 57.](#)

CLEM expressions. S'il n'est pas possible de réorganiser un flux, vous pouvez modifier les options de noeud ou les expressions CLEM, ou bien encore modifier le mode d'exécution de l'opération de façon à ce que cette dernière n'entrave plus la génération SQL. Les noeuds Calculer, Sélectionner, et tout noeud semblable, peuvent généralement être convertis en code SQL, à condition que tous les opérateurs d'expression CLEM présentent des équivalents SQL. La plupart des opérateurs peuvent être convertis, mais certains d'entre eux entravent la génération SQL (en particulier, les fonctions séquentielles [“fonctions @ ”]). Parfois, la génération est arrêtée car la requête générée est devenue trop complexe pour une prise en charge par la base de données. [Pour plus](#)

d'informations, reportez-vous à la section [Opérateurs et expressions CLEM prenant en charge la génération SQL](#) sur p. 60.

Noeuds source multiples. Lorsqu'un flux comporte plusieurs noeuds source de base de données, la génération SQL est appliquée indépendamment à chaque branche d'entrée. Si la génération est arrêtée sur une branche, elle peut se poursuivre sur une autre. Lorsque deux branches fusionnent (et qu'elles peuvent toutes deux être exprimées en code SQL jusqu'au point de fusion), la fusion elle-même peut être souvent remplacée par une jointure de base de données ; la génération peut ainsi continuer en aval.

Algorithmes de base de données. L'estimation du modèle est toujours effectuée sur IBM® SPSS® Modeler Server plutôt que dans la base de données, excepté lors de l'utilisation d'algorithmes natifs de base de données issus de Microsoft, d'IBM ou d'Oracle. [Pour plus d'informations, reportez-vous à la section Modélisation de base de données - Présentation dans le chapitre 2 dans Guide d'exploration de base de données IBM SPSS Modeler 14.2.](#)

Scoring de modèles. Le scoring dans la base de données est pris en charge pour certains modèles par conversion du modèle généré en code SQL. Toutefois, certains modèles génèrent des expressions SQL extrêmement complexes qui ne sont pas toujours évaluées efficacement au sein de la base de données. C'est la raison pour laquelle la génération SQL doit être activée séparément pour chaque noeud de modèle. Si vous constatez qu'un noeud de modèle entrave la génération de SQL, accédez à l'onglet Paramètres dans la boîte de dialogue du noeud et sélectionnez Générer SQL pour ce modèle (certains modèles peuvent offrir des options supplémentaires qui contrôlent la génération). Exécutez des tests pour confirmer que l'option présente des avantages pour votre application. [Pour plus d'informations, reportez-vous à la section Noeuds prenant en charge la génération SQL](#) sur p. 57.

Lors du test de noeuds de modélisation pour vérifier que la génération SQL fonctionne correctement pour les modèles, nous recommandons d'abord d'enregistrer tous les flux depuis IBM® SPSS® Modeler. Certains systèmes de base de données peuvent se bloquer lors des tentatives de traiter le SQL généré (potentiellement complexe), et nécessitent de fermer SPSS Modeler depuis le gestionnaire de tâches de Windows.

Mise en cache de base de données. Si vous utilisez un cache de noeud pour enregistrer les données au niveau des points critiques du flux (par exemple, après un noeud Fusionner ou Agréger), vérifiez que la mise en cache de base de données est activée conjointement avec l'optimisation SQL. Dans la plupart des cas, les données peuvent ainsi être mises en cache dans une table temporaire de la base de données (plutôt que dans le système de fichiers). [Pour plus d'informations, reportez-vous à la section Configuration de l'optimisation SQL](#) sur p. 52. Reportez-vous également à [Options de mise en cache des noeuds](#).

Code SQL propre au fournisseur. La plupart du code SQL généré est conforme aux normes (SQL-92), mais certaines fonctionnalités non standard, propres au fournisseur, sont utilisées le cas échéant. Le degré d'optimisation SQL peut varier en fonction de la source de base de données.

Noeuds prenant en charge la génération SQL

Les tableaux suivants contiennent des noeuds qui représentent des opérations de Data mining prenant en charge la génération SQL. Si un noeud n'apparaît pas dans ces tables, il ne prend pas en charge la génération SQL à l'exception des noeuds de modélisation de base de données.

Vous pouvez avoir un aperçu de l'instruction SQL générée avant de l'exécuter. [Pour plus d'informations, reportez-vous à la section Prévisualisation du code SQL généré sur p. 53.](#)

Sources

Noeud prenant en charge la génération SQL	Remarques
SGBD	Ce noeud permet d'indiquer les tables et vues à utiliser pour approfondir l'analyse. Il autorise la saisie de requêtes SQL. Evitez les résultats comportant des noms de colonne en double. Pour plus d'informations, reportez-vous à la section Rédaction de requêtes SQL sur p. 63.

Opérations sur les lignes

Noeud prenant en charge la génération SQL	Remarques
Sélectionner	Ce noeud prend en charge la génération SQL uniquement si celle-ci est prise en charge pour l'expression sélectionnée elle-même (reportez-vous aux expressions ci-dessous). Si des champs comportent des valeurs nulles, la génération SQL n'offre pas les mêmes résultats pour l'exclusion comme ils sont donnés dans l'application IBM® SPSS® Modeler native.
Echantillonner	L'échantillonnage simple prend en charge la génération SQL comme indiqué dans le tableau suivant. L'échantillonnage complexe ne prend pas en charge la génération SQL.
Agréger	
RFM Agréger	Prend en charge la génération sauf en cas d'enregistrement de la date de la deuxième et troisième transactions les plus récentes, ou en cas de prise en compte des transactions récentes. Toutefois, cette prise en compte fonctionne si la fonction <code>datetime_date(YEAR,MONTH,DAY)</code> est répercutée.
Trier	
Fusionner	Aucun code SQL n'est généré pour la fusion par ordre. La fusion par clé avec les jointures externe complète ou partielle est seulement prise en charge si la base de données/le pilote la prend en charge. Vous pouvez renommer les champs d'entrées qui ne correspondent pas à l'aide du noeud Filtre ou de l'onglet Filtre d'un noeud source. Pour tous les types de fusion, <code>SQL_SP_EXISTS</code> n'est pas pris en charge si les entrées proviennent de bases de données différentes.
Ajouter	Prend en charge la génération si les entrées ne sont pas triées.
Distinguer	

Prise en charge de la génération SQL dans le noeud Echantillonner lorsque l'échantillonnage simple est indiqué dans le tableau suivant.

Mode	Echantillon	File max	Grain	DB2 OS/Z	DB2 OS/400	DB2 Windows	Netezza Windows/UNIX	Oracle	SQL Server	Teradata
Inclure	Premiers	n/a		Y	Y	Y	Y	Y	Y	Y
		désactivé		Y	Y	Y	Y	Y		Y
	% aléatoire	max	désactivé	désactivé		Y	Y	Y		Y
			activé(e)			Y		Y		
		max	désactivé	désactivé		Y	Y	Y		Y
			activé(e)			Y		Y		
Supprimer	Premiers	désactivé					Y	Y		
		max					Y	Y		
	Tous les	désactivé		Y	Y	Y	Y	Y		Y
		max		Y	Y	Y	Y	Y		Y
	% aléatoire	max	désactivé	désactivé		Y	Y	Y		Y
			activé(e)			Y		Y		
	max	désactivé	désactivé		Y	Y	Y		Y	
	activé(e)				Y		Y			

Opérations sur les champs

Noeud prenant en charge la génération SQL	Remarques
Typier	Ce noeud prend en charge la génération SQL si le noeud Typier est instancié, et qu'aucune vérification du type ABORT ou WARN n'est indiquée.
Filtrer	
Calculer	Ce noeud prend en charge la génération SQL si le code SQL généré pour la formule de calcul est pris en charge (reportez-vous aux expressions ci-dessous).
Ensemble	Prend en charge la génération SQL pour les cibles continues. Pour les autres champs cible, prend uniquement en charge la génération si la méthode d'ensemble «La plus grande confiance gagne» est utilisée.
Remplacer	Ce noeud prend en charge la génération SQL si le code SQL généré pour la formule de calcul est pris en charge (reportez-vous aux expressions ci-dessous).
Anonymiser	Prend en charge la génération SQL pour les cibles continues, et une génération SQL partielle pour les cibles nominales et booléennes.
Recoder	
Discrétiser	Prend en charge la génération SQL si la méthode de création d'intervalles de type « quantile (nombre égal) » est utilisée et que l'option « Lire dans l'onglet Valeurs d'intervalle si disponible » est sélectionnée.
Analyse RFM	Prend en charge la génération SQL si l'option « Lire dans l'onglet Valeurs d'intervalle si disponible » est sélectionnée mais les noeuds en aval ne prennent pas en charge la génération SQL.
Partitionner	Prend en charge la génération SQL pour affecter des enregistrements aux partitions.

Noeud prenant en charge la génération SQL	Remarques
Binariser	
Restructurer	

Graphiques

Noeud prenant en charge la génération SQL	Remarques
Représentation graphique	La génération SQL est prise en charge pour les types de graphiques suivants : Graphique de surface, Graphique de surface 3D, Graphique à barres, Graphique à barres 3D, Graphique à barres des effectifs, Visualisation graphique des points chauds, Graphique circulaire, Graphique circulaire 3D, Graphique circulaire des effectifs. Pour les histogrammes, la génération SQL est prise en charge pour les données catégorielles uniquement.
Proportion	
Relations	
Evaluation	

Noeuds de modèle généré

Pour certains modèles, il est possible de générer une instruction SQL pour le nugget de modèle, en répercutant l'étape de scoring de modèle dans la base de données. L'utilisation principale de cette caractéristique n'est pas d'améliorer les performances, mais d'autoriser des flux contenant ces nuggets à voir l'intégralité de leur instruction SQL répercutée. [Pour plus d'informations, reportez-vous à la section Affichage SQL pour des nuggets de modèle sur p. 55.](#)

Nugget de modèle prenant en charge la génération SQL	Remarques
Arbre C&RT	Prend en charge la génération SQL pour l'option d'arbre unique, mais pas pour les options de boosting, de bagging et d'ensembles très volumineux.
QUEST	
CHAID	
C5,0	
Liste de décision	
Linéaire	Prend en charge la génération SQL pour l'option de modèle standard, mais pas pour les options de boosting, de bagging et d'ensembles très volumineux.
R. neurones	Prend en charge la génération SQL pour l'option de modèle standard (perceptron multicouche uniquement), mais pas pour les options de boosting, de bagging et d'ensembles très volumineux.
ACP/Facteur	
Logistique	Prend en charge la génération SQL pour la procédure multinomiale, mais pas la procédure binomiale. Pour la procédure multinomiale, la génération n'est pas prise en charge lorsque les confiances sont sélectionnées, sauf si le type de cible est Booléen.
Ensemble de règles généré	

Sortie

Noeud prenant en charge la génération SQL	Remarques
Table	Ce nœud prend en charge la génération si la génération SQL est prise en charge pour l'expression de mise en évidence (reportez-vous aux expressions ci-dessous).
Matrice	Prend en charge la génération sauf si la fonction « Tous les numériques » est sélectionnée pour l'option Champs.
Analyse	La prise en charge de la génération dépend des options sélectionnées.
Transformation	
Statistiques	Ce nœud prend en charge la génération si l'option Corréler n'est pas utilisée.
Rapport	
V. globales	

Exporter

Noeud prenant en charge la génération SQL	Remarques
SGBD	
Publisher	Le flux publié contient du code SQL généré.

Opérateurs et expressions CLEM prenant en charge la génération SQL

Les tableaux ci-dessous présentent les expressions et opérations mathématiques qui prennent en charge la génération SQL et sont fréquemment utilisées dans le cadre du Data mining. Les opérations absentes de ces tableaux ne prennent pas en charge la génération SQL dans la version actuelle.

Opérateurs

Opérations prenant en charge la génération SQL	Remarques
+	
-	
/	
*	
><	Permet de concaténer les chaînes.

Opérateurs relationnels

Opérations prenant en charge la génération SQL	Remarques
=	
/=	Permet d'indiquer "pas égal".
>	
>=	

Opérations prenant en charge la génération SQL	Remarques
<	
<=	

Fonctions

Opérations prenant en charge la génération SQL	Remarques
abs	
allbutfirst	
allbutlast	
and	
arccos	
arcsin	
arctan	
arctanh	
cos	
div	
exp	
fracof	
hasstartstring	
hassubstring	
integer	
intof	
isalphacode	
islowercode	
isnumbercode	
isstartstring	
issubstring	
isuppercode	
last	
length	
locchar	
log	
log10	
lowertoupper	
max	
member	
min	
negate	
not	
number	
or	
pi	
real	

Opérations prenant en charge la génération SQL	Remarques
rem	
round	
sign	
sin	
sqrt	
string	
strmember	
subscrs	
substring	
substring_between	
uppertolower	
to_string	

Fonctions spéciales

Opérations prenant en charge la génération SQL	Remarques
@NULL	
@GLOBAL_AVE	Les fonctions globales spéciales permettent d'extraire les valeurs globales calculées par le nœud V. globales.
@GLOBAL_SUM	
@GLOBAL_MAX	
@GLOBAL_MEAN	
@GLOBAL_MIN	
@GLOBALSDEV	

Fonctions Agrégation

Opérations prenant en charge la génération SQL	Remarques
Sum	
Mean	
Min	
Max	
Count	
SDev	

Utilisation de fonctions SQL dans les expressions CLEM

Le @SQLFN peut être utilisé pour ajouter des fonctions SQL nommées avec des expressions CLEM à des fins d'exécution de la base de données uniquement. Cela peut s'avérer utile dans des cas particuliers où le SQL propriétaire ou d'autres personnalisations propres au fournisseur sont nécessaires.

L'utilisation de cette fonction n'est pas couverte par l'accord d'assistance IBM® SPSS® Modeler standard. En effet, son exécution repose sur des composants de base de données externes non contrôlés par IBM Corp.. Elle peut néanmoins être déployée dans des cas particuliers, par exemple dans le cadre d'un contrat de services. Contactez <http://www.ibm.com/software/analytics/spss/services/> si vous avez besoin d'informations complémentaires.

Rédaction de requêtes SQL

Lorsque vous utilisez le noeud SGBD, faites particulièrement attention aux requêtes SQL aboutissant à des ensembles de données comportant des noms de colonne en double. Ces derniers empêchent souvent l'optimisation SQL avec les noeuds en aval.

IBM® SPSS® Modeler utilise des instructions **SELECT** imbriquées pour répercuter le code SQL des flux utilisant une requête SQL dans le noeud source SGBD. Autrement dit, le flux imbrique la requête indiquée dans le noeud source SGBD dans une ou plusieurs instructions **SELECT** générées au cours de l'optimisation des noeuds en aval. Par conséquent, si les résultats de la requête contiennent des noms de colonne en double, l'instruction ne peut pas être imbriquée par le SGBDR. Les problèmes d'imbrication surviennent surtout lors des jointures de table : une colonne de même nom est alors sélectionnée dans plusieurs tables jointes. Prenons l'exemple de la requête suivante dans le noeud source :

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

Cette requête empêche toute optimisation SQL ultérieure. En effet, l'instruction **SELECT** aboutirait à un ensemble de données comportant deux colonnes intitulées **ID**.

Afin d'autoriser une optimisation SQL complète, vous devez être plus précis lors de la rédaction des requêtes SQL et indiquer des alias de colonne dès que vous rencontrez le problème des noms de colonne en double. L'instruction ci-dessous présente une requête plus explicite :

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

Configuration d'Oracle pour les plates-formes UNIX

Configuration d'Oracle pour l'optimisation SQL

Lorsque vous exécutez IBM® SPSS® Modeler Server sur des plates-formes UNIX et que la lecture s'effectue à partir d'une base de données Oracle, suivez les conseils ci-dessous pour garantir une optimisation totale du code SQL généré dans la base de données.

Spécification des paramètres régionaux appropriés

Si vous exécutez SPSS Modeler Server avec des paramètres régionaux autres que ceux fournis avec les pilotes Connect ODBC, reconfigurez l'ordinateur afin d'améliorer l'optimisation SQL. Seuls les fichiers de paramètres régionaux *en_US* sont fournis avec les pilotes Connect ODBC. Par conséquent, si l'ordinateur SPSS Modeler Server est exécuté avec d'autres paramètres régionaux ou si les paramètres régionaux du shell dans lequel SPSS Modeler Server a été lancé ne sont pas entièrement définis, le code SQL généré risque de ne pas être totalement optimisé dans Oracle. En voici les raisons :

- SPSS Modeler Server utilise les fichiers de paramètres régionaux ODBC correspondant aux paramètres régionaux avec lesquels il est exécuté pour convertir les codes renvoyés par la base de données en chaînes de texte. Il se sert ensuite de ces chaînes pour identifier la base de données à laquelle il tente de se connecter.
- Si les paramètres régionaux renvoyés à SPSS Modeler Server par la requête \$LANG du système ne sont pas des paramètres *en_US*, IBM® SPSS® Modeler ne parvient pas à convertir en texte les codes reçus du pilote ODBC. C'est donc un code non converti, et non la chaîne *Oracle*, qui est renvoyé à SPSS Modeler Server au début de la connexion à la base de données. SPSS Modeler n'est alors pas en mesure d'optimiser les flux pour Oracle.

Pour vérifier les spécifications des paramètres régionaux et les redéfinir :

- ▶ Dans un shell UNIX, exécutez la chaîne suivante :

```
#locale
```

Vous obtenez alors des informations sur les paramètres régionaux du shell. Par exemple :

```
$ régionaux
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
```

- ▶ Accédez au répertoire Connect ODBC/locale. (Un seul répertoire, *en_US*, y figure.)
- ▶ Créez un lien symbolique avec ce répertoire *en_US*, en indiquant le nom de la configuration des paramètres régionaux dans le shell. Voici un exemple de cette commande :

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

Si les paramètres régionaux concernés ne sont pas anglais (*fr_FR.ISO8859-1*, par exemple), procédez comme suit pour créer le lien symbolique :

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

- ▶ Une fois le lien créé, redémarrez SPSS Modeler Server à partir de ce même shell. (SPSS Modeler Server reçoit les informations relatives aux paramètres régionaux à partir du shell duquel il est lancé.)

Remarques

Lorsque vous optimisez un ordinateur UNIX afin d'effectuer des répercussions SQL dans Oracle, suivez ces conseils :

- Vous devez définir l'intégralité des paramètres régionaux. Dans l'exemple ci-avant, vous devez créer un lien de type *langue_pays.code-page*. Le répertoire de paramètres régionaux existant (*en_US*) n'est pas suffisant.
- Pour une optimisation totale de l'exploration de base de données, définissez les variables **LANG** et **LC_ALL** du shell utilisé pour lancer SPSS Modeler Server. Comme toute autre variable d'environnement, la variable **LANG** peut être définie dans le shell, avant le redémarrage de SPSS Modeler Server. Prenons l'exemple de la définition suivante :

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

- A chaque lancement de SPSS Modeler Server, vérifiez que les paramètres régionaux du shell sont entièrement définis et qu'un lien symbolique approprié figure dans le répertoire ODBC/locale.

Configuration des scripts de démarrage UNIX

Introduction

Cette annexe décrit certains scripts qui accompagnent les versions UNIX de IBM® SPSS® Modeler Server et explique comment les configurer. Les scripts permettent d'exécuter les actions suivantes :

- configurer le démarrage automatique de SPSS Modeler Server lorsque l'ordinateur du serveur est redémarré ;
- arrêter et redémarrer manuellement SPSS Modeler Server.
- modifier les droits d'accès sur les fichiers créés par SPSS Modeler Server.
- configurer SPSS Modeler Server de façon à ce qu'il fonctionne avec les pilotes Connect ODBC fournis avec SPSS Modeler Server. [Pour plus d'informations, reportez-vous à la section IBM SPSS Modeler Server et Data Access Pack sur p. 68.](#)
- Configurer SPSS Modeler Server de façon à ce qu'il utilise les IBM® SPSS® Collaboration and Deployment Services Enterprise View Driver lors de la lecture de données avec un nœud Enterprise View. [Pour plus d'informations, reportez-vous à la section Configuration d'un pilote pour le nœud Enterprise View sur p. 75.](#)

Scripts

IBM® SPSS® Modeler Server utilise plusieurs scripts, notamment :

- **modelersrv.sh.** Le script de démarrage manuel de SPSS Modeler Server est situé dans le répertoire d'installation SPSS Modeler Server. Il configure l'environnement du serveur lorsque le processus démon du serveur est démarré *manuellement*. Exécutez-le lorsque vous souhaitez démarrer et arrêter manuellement le serveur. Editez-le si vous souhaitez modifier la configuration relative au démarrage manuel.
- **auto.sh.** Ce script configure votre système de façon à démarrer le processus démon du serveur automatiquement lors du démarrage. Exécutez-le une fois pour configurer le démarrage automatique. Il n'est pas nécessaire de l'éditer. Le script est situé dans le répertoire d'installation de SPSS Modeler Server.
- **rc.modeler.** Lorsque vous exécutez *auto.sh*, ce script est créé à un emplacement qui dépend du système d'exploitation de votre serveur. Il configure l'environnement du serveur lorsque ce dernier est démarré *automatiquement*. Editez-le si vous souhaitez modifier la configuration relative au démarrage automatique.

Table B-1
Emplacement du script *rc.modeler* en fonction du système d'exploitation

Système d'exploitation	Emplacement
AIX	<i>/etc/rc.modeler</i>
HP-UX	<i>/sbin/init.d/rc.modeler</i>
Solaris	<i>/etc/init.d/rc.modeler</i>

Démarrage et arrêt automatiques de IBM SPSS Modeler Server

IBM® SPSS® Modeler Server doit être démarré en tant que processus démon. Le programme d'installation inclut un script (*auto.sh*) qui permet de configurer votre système sur un arrêt et un redémarrage automatiques de SPSS Modeler Server.

Pour configurer le système sur un démarrage et un arrêt automatiques

- ▶ Connectez-vous en tant que racine (root).
- ▶ Remplacez le répertoire actuel par le répertoire d'installation de SPSS Modeler Server.
- ▶ Exécutez le script. A l'invite UNIX, tapez :

```
./auto.sh
```

Un script de démarrage automatique, *rc.modeler*, est créé à l'emplacement indiqué dans le tableau ci-dessus. Le système d'exploitation utilise *rc.modeler* pour démarrer le processus démon SPSS Modeler Server à chaque redémarrage de l'ordinateur du serveur. Le système d'exploitation utilise également le script *rc.modeler* pour arrêter le démon à chaque arrêt du système.

Démarrage et arrêt manuels de IBM SPSS Modeler Server

Vous pouvez démarrer et arrêter manuellement IBM® SPSS® Modeler Server en exécutant le script *modelersrv.sh*.

Pour démarrer et arrêter IBM SPSS Modeler Server manuellement

- ▶ Remplacez le répertoire actuel par le répertoire d'installation de SPSS Modeler Server.
- ▶ Pour démarrer le serveur, saisissez ce qui suit à l'invite de commande UNIX :

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

- ▶ A l'invite de commande UNIX, pour arrêter le serveur, entrez la ligne suivante :

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

Edition de scripts

Si vous utilisez à la fois le démarrage manuel et le démarrage automatique, procédez à des modifications parallèles dans *modelersrv.sh* et *rc.modeler*. Si vous utilisez uniquement le démarrage manuel, procédez à des modifications dans *modelersrv.sh*. Si vous utilisez uniquement le démarrage automatique, procédez à des modifications dans *rc.modeler*.

Pour éditer les scripts

- ▶ Arrêter IBM® SPSS® Modeler Server. (Pour plus d'informations, reportez-vous à la section [Démarrage et arrêt manuels de IBM SPSS Modeler Server](#) sur p. 67.)
- ▶ Localisez le script approprié. (Pour plus d'informations, reportez-vous à la section [Scripts](#) sur p. 66.)
- ▶ Ouvrez le script dans un éditeur de texte, apportez vos modifications et enregistrez le fichier.
- ▶ Démarrez SPSS Modeler Server, soit automatiquement (en redémarrant l'ordinateur du serveur), soit manuellement.

Contrôle des droits d'accès relatifs à la création de fichier

IBM® SPSS® Modeler Server crée des fichiers temporaires présentant des droits d'accès en lecture, écriture et exécution pour tous les utilisateurs. Vous pouvez ignorer ce paramètre par défaut en éditant le paramètre `UMASK` du script de démarrage, soit dans *modelersrv.sh*, soit dans *rc.modeler* ou dans les deux. (Pour plus d'informations, reportez-vous à [Edition de scripts](#) ci-dessus.) Nous vous recommandons, en tant que restriction maximale, de paramétrer `077` sur `UMASK`. Les paramètres plus restrictifs peuvent engendrer des problèmes en termes de droits d'accès, pour SPSS Modeler Server.

IBM SPSS Modeler Server et Data Access Pack

Pour que vous puissiez utiliser les pilotes Connect for ODBC avec IBM® SPSS® Modeler Server, l'environnement Connect for ODBC doit être configuré par *odbc.sh* lors du démarrage du processus SPSS Modeler Server. Pour cela, éditez le script de démarrage IBM® SPSS® Modeler approprié, soit dans *modelersrv.sh*, soit dans *rc.modeler* ou dans les deux. (Pour plus d'informations, reportez-vous à la section [Edition de scripts](#) sur p. 68.)

Pour plus d'informations, consultez le site Web de l'entreprise, à l'adresse <http://www.spss.com/drivers/>. Si vous avez des questions sur la création ou la définition d'autorisations pour les sources de données ODBC, contactez l'administrateur de votre base de données.

Pour configurer le démarrage de l'environnement Connect for ODBC avec IBM SPSS Modeler Server

- ▶ À partir du site Web de l'entreprise à l'adresse <http://www.spss.com/drivers/clientCLEM.htm>, téléchargez l'archive tar compressée appropriée à la plateforme sur laquelle vous avez installé SPSS Modeler Server. Assurez-vous de télécharger les pilotes adaptés à la version de SPSS

Modeler Server que vous avez installée. Copiez le fichier à l'emplacement où vous souhaitez installer les pilotes ODBC (par exemple, */usr/spss/odbc*).

- ▶ Procédez à l'extraction de l'archive.
- ▶ Exécutez le texte du script *setodbcpath.sh* ayant été extrait de l'archive.
- ▶ Modifiez le script *odbc.sh* pour ajouter la définition de ODBCINI au bas de ce script et exportez-le, par exemple :

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

ODBCINI doit pointer vers le nom du chemin d'accès complet du fichier *odbc.ini* que SPSS Modeler doit lire pour obtenir une liste des sources de données ODBC que vous définissez (un *odbc.ini* par défaut est installé avec les pilotes).

- ▶ Enregistrez *odbc.sh*.
- ▶ (installations SPSS Modeler Server 64 bits uniquement ; pour les autres installations, continuez à partir de l'étape suivante) Définissez et exportez LD_LIBRARY_PATH_64 dans *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
```

où *library_path* est le même que pour la définition de LD_LIBRARY_PATH qui se trouve déjà dans le script et qui a été initialisée avec le chemin de votre installation (par exemple */usr/spss/odbc/lib*). La façon la plus simple de procéder est de copier les instructions `if` et `export` pour LD_LIBRARY_PATH dans votre fichier *odbc.sh*, de les ajouter à la fin du script, puis de remplacer les chaînes "LD_LIBRARY_PATH" dans les nouvelles instructions modifiées `if` et `export` avec "LD_LIBRARY_PATH_64".

Ainsi, votre fichier *odbc.sh* final sur une installation SPSS Modeler Server 64 bits devrait ressembler à cela :

```
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
```

N'oubliez pas d'exporter LD_LIBRARY_PATH_64, et de le définir dans la boucle `if`.

- ▶ Modifiez le fichier *odbc.ini* que vous avez défini précédemment à l'aide de \$ODBCINI. Définissez les noms de la source de données nécessaires (ils dépendent de la base de données à laquelle vous accédez).
- ▶ Enregistrez le fichier *odbc.ini*.
- ▶ Configurez SPSS Modeler Server pour qu'il utilise ces pilotes. Pour ce faire, modifiez *modelersrv.sh* et ajoutez la ligne suivante directement sous la ligne qui définit SCLEMDNAME :

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

où *odbc.sh_path* est le chemin complet vers le fichier *odbc.sh* que vous avez modifié au début de cette procédure, par exemple :

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

Remarque : La syntaxe est importante ; assurez-vous de laisser un espace entre le premier point et le chemin d'accès au fichier.

- ▶ Enregistrez *modelersrv.sh*.

Pour tester la connexion

- ▶ Redémarrez SPSS Modeler Server.
- ▶ Connectez-vous à SPSS Modeler Server à partir d'un client.
- ▶ Sur le client, ajoutez un noeud source de base de données à l'espace de travail.
- ▶ Ouvrez le noeud et vérifiez que vous pouvez voir les noms de la source de données que vous avez définis précédemment dans le fichier *odbc.ini* lors de la configuration.

Si vous ne les voyez pas là où ils devraient être, ou si vous obtenez des erreurs lorsque vous essayez de vous connecter à une source de données que vous avez définie, suivez la procédure de résolution des problèmes. [Pour plus d'informations, reportez-vous à la section Dépannage de la configuration Connect for ODBC sur p. 71.](#)

Pour configurer le démarrage de l'environnement Connect for ODBC avec IBM SPSS Modeler Solution Publisher Runtime

Lorsque vous parvenez à vous connecter à la base de données à partir de SPSS Modeler Server, vous pouvez configurer une installation de IBM® SPSS® Modeler Solution Publisher Runtime sur le même serveur en référençant le même script *odbc.sh* à partir du script de départ de SPSS Modeler Solution Publisher Runtime.

- ▶ Modifiez le script *modelerrun* dans SPSS Modeler Solution Publisher Runtime pour ajouter la ligne suivante immédiatement au-dessus de la dernière ligne du script :

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

où *odbc.sh_path* est le chemin complet vers le fichier *odbc.sh* que vous avez modifié au début de cette procédure, par exemple :

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

Remarque : La syntaxe est importante ; assurez-vous de laisser un espace entre le premier point et le chemin d'accès au fichier.

- ▶ Enregistrez le fichier du script *modelerrun*.

Pour configurer le démarrage de l'environnement Connect for ODBC avec IBM SPSS Modeler Batch

Aucune configuration du script IBM® SPSS® Modeler Batch n'est nécessaire pour ODBC, car vous vous connectez à SPSS Modeler Server à partir de SPSS Modeler Batch pour exécuter les flux. Vérifiez que la configuration ODBC de SPSS Modeler Server a été effectuée et fonctionne correctement, comme décrit précédemment dans cette section.

Pour ajouter ou modifier un nom de source de données

- ▶ Modifiez le fichier *odbc.ini* pour inclure le nouveau nom ou le nom modifié.
- ▶ Testez la connexion comme décrit précédemment dans cette section.

Lorsque la connexion avec SPSS Modeler Server fonctionne correctement, la nouvelle source de données ou la source modifiée, doit également fonctionner correctement avec SPSS Modeler Solution Publisher Runtime et SPSS Modeler Batch.

Dépannage de la configuration Connect for ODBC

Aucune source de donnée n'apparaît ou du texte aléatoire s'affiche

Si vous ouvrez un noeud source de base de données et que la liste des sources de données disponibles est vide ou contient des entrées inattendues, il se peut qu'il y ait un problème avec le script de démarrage.

- ▶ Vérifiez que \$ODBCINI est défini dans *modelersrv.sh*, soit explicitement dans le script lui-même ou dans le script *odbc.sh* qui est référencé dans *modelersrv.sh*.
- ▶ Dans ce dernier cas, vérifiez que ODBCINI pointe vers le chemin d'accès complet du fichier *odbc.ini* que vous avez utilisé pour définir vos sources de données ODBC.
- ▶ Si la spécification du chemin d'accès dans ODBCINI est correcte, vérifiez la valeur de \$ODBCINI utilisée dans l'environnement IBM® SPSS® Modeler Server en renvoyant la variable depuis *modelersrv.sh*. Pour ce faire, ajoutez la ligne suivante à *modelersrv.sh* après le point où vous avez défini ODBCINI :

```
echo $ODBCINI
```

- ▶ Enregistrez puis exécutez *modelersrv.sh*. La valeur de \$ODBCINI définie dans l'environnement SPSS Modeler Server est écrite dans *stdout* pour la vérification.

- Si aucune valeur n'est renvoyée vers *stdout*, et que vous définissez \$ODBCINI dans le script *odbc.sh* auquel vous faites référence depuis *modelersrv.sh*, vérifiez que la syntaxe de référence est correcte. Celle-ci devrait ressembler à :

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

où *odbc.sh_path* est le chemin complet vers le fichier *odbc.sh* que vous avez modifié au début de cette procédure, par exemple :

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

Remarque : La syntaxe est importante ; assurez-vous de laisser un espace entre le premier point et le chemin d'accès au fichier.

Lorsque la valeur correcte est renvoyée vers *stdout* lors de l'exécution de *modelersrv.sh*, vous devez voir les noms de source de données dans le noeud de source de base de données lorsque vous redémarrez SPSS Modeler Server et le connectez au client.

Le client IBM SPSS Modeler attend que vous cliquiez sur *Connecter* dans la boîte de dialogue *Connexions à la base de données*

Ce comportement peut se produire lorsque votre chemin d'accès à la bibliothèque n'est pas correctement défini pour inclure les chemins d'accès aux bibliothèques ODBC. Le chemin d'accès de la bibliothèque est défini par \$LD_LIBRARY_PATH (et \$LD_LIBRARY_PATH_64 sur les versions 64 bits).

Pour consulter la valeur du chemin d'accès de la bibliothèque dans l'environnement daemon de SPSS Modeler Server, renvoyez la valeur de la variable d'environnement appropriée depuis *modelersrv.sh*, après la ligne dans laquelle vous ajoutez le chemin de la bibliothèque ODBC au chemin de la bibliothèque et exécutez le script. La valeur du chemin de la bibliothèque sera renvoyée au terminal la prochaine fois que vous exécuterez le script.

Si vous faites référence à *odbc.sh* depuis *modelersrv.sh* pour configurer votre environnement ODBC de SPSS Modeler Server, renvoyez la valeur de votre chemin de bibliothèque de la ligne après celle où vous avez référencé le script *odbc.sh*. Pour renvoyer la valeur, ajoutez la ligne suivante au script puis enregistrez et exécutez le fichier de script :

```
echo $<library_path_variable>
```

où *<library_path_variable>* est la variable du chemin d'accès à la bibliothèque approprié pour votre système d'exploitation de serveur.

La valeur renvoyée du chemin d'accès de votre bibliothèque contient le chemin d'accès vers le sous-répertoire *lib* de votre installation ODBC. Si ce n'est pas le cas, ajoutez cet emplacement au fichier.

Si vous exécutez la version 64 bits de SPSS Modeler Server, \$LD_LIBRARY_PATH_64 remplacera \$LD_LIBRARY_PATH s'il est configuré pour cela. Si vous rencontrez des problèmes sur l'une de ces plateformes 64 bits, renvoyez LD_LIBRARY_PATH_64 et \$LD_LIBRARY_PATH

depuis *modelersrv.sh* et, si nécessaire, définissez `$LD_LIBRARY_PATH_64` pour qu'il contienne le sous-répertoire *lib* de votre installation ODBC et exportez la définition.

Nom de la source de données introuvable et aucun pilote par défaut spécifié

Si vous voyez cette erreur lorsque vous cliquez sur Connecter dans la boîte de dialogue Connexions à la base de données, cela indique généralement que votre fichier *odbc.ini* n'est pas défini correctement. Vérifiez que le nom de la source de données (DSN) tel que défini dans la section [ODBC Data Sources] en haut du fichier correspond à la chaîne spécifiée entre crochets plus loin dans *odbc.ini* pour définir le DSN. Si des différences existent entre les deux, vous verrez cette erreur apparaître lorsque vous essaieriez de vous connecter à l'aide du DSN depuis IBM® SPSS® Modeler. L'exemple suivant illustre une spécification *incorrecte* :

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

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

Vous devez modifier une des deux chaînes en gras pour qu'elles correspondent parfaitement. Cela devrait résoudre l'erreur.

Impossible de charger le pilote spécifié

Cette erreur indique également que le fichier *odbc.ini* n'est pas correctement défini. Une possibilité est que le paramètre du pilote dans la strophe du pilote ne soit pas correctement défini, par exemple :

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

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

- ▶ Vérifiez que l'objet partagé spécifié par le paramètre Pilote existe.
- ▶ Corrigez le chemin d'accès à l'objet partagé s'il est incorrect.
- ▶ Si le paramètre Pilote est spécifié dans ce format :

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

cela indique que vous n'avez pas initialisé vos scripts ODBC. Exécutez le script *setodbcpath.sh* installé avec les pilotes. [Pour plus d'informations, reportez-vous à la section IBM SPSS Modeler Server et Data Access Pack sur p. 68.](#) Lorsque vous avez exécuté ce script, la chaîne "ODBCHOME" doit être remplacée par le chemin d'accès à votre installation ODBC. Cela devrait résoudre le problème.

Une autre cause peut concerner la bibliothèque de votre pilote. Utilisez l'outil *ivtestlib* fourni avec l'environnement Connect for ODBC pour confirmer l'impossibilité de charger le pilote. Pour Connect64, utilisez l'outil *ddtestlib*. Corrigez le problème en configurant la variable du chemin de la bibliothèque dans le script de démarrage.

Par exemple, si le pilote Oracle ne peut être chargé pour une installation 32 bits, suivez ces étapes :

- Utilisez l'outil *ivtestlib* pour confirmer l'impossibilité de charger le pilote. Par exemple, à l'invite UNIX, entrez la ligne suivante :

```
sh
cd REPODBC
. odbcs.sh
./bin/ivtestlib MFor815
```

où *ODBCDIR* est remplacé par le chemin d'accès à votre répertoire d'installation Connect for ODBC.

- Lisez le message pour savoir s'il existe une erreur. Par exemple, le message :

```
Echec du chargement de MFor815.so : ld.so.1 : bin/ivtestlib : fatal : libclntsh.so : échec de l'ouverture :
aucun fichier ni répertoire de ce type
```

indique que la bibliothèque client Oracle, *libclntsh.so*, est manquante ou qu'elle ne correspond pas au chemin (par exemple, qu'elle ne correspond pas à *LD_LIBRARY_PATH* sous Solaris).

- Confirmez l'existence de la bibliothèque. Si elle n'existe pas, réinstallez le client Oracle. Si elle existe, saisissez la séquence de commandes suivante à l'invite de commandes UNIX :

```
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/disque/oracle/produit/8.1.6/lib
export LD_LIBRARY_PATH
./bin/ivtestlib Mfor815
```

où */disque/oracle/produit/8.1.6/lib* est remplacé par le chemin d'accès à *libclntsh.so* et *LD_LIBRARY_PATH* correspond à la variable du chemin de bibliothèque pour votre système d'exploitation.

Remarque: si vous utilisez SPSS Modeler 64 bits sous Linux ou Solaris, la variable du chemin de la bibliothèque contient le suffixe *_64*. Par conséquent, les deux premières lignes de l'exemple précédent dévient :

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

- Lisez le message pour vérifier qu'il est désormais possible de charger le pilote. Par exemple, le message :

```
Chargement de MFor815.so réussi ; qehandle : 0xFF3A1BE4
```


indique qu'il est possible de charger la bibliothèque client Oracle.

- ▶ Corrigez le chemin de bibliothèque dans le script de démarrage SPSS Modeler.
- ▶ Redémarrez le SPSS Modeler Server à l'aide du script de démarrage édité (*modelersrv.sh* ou *rc.modeler*).

Chemins de bibliothèque

Le nom de la variable du chemin de bibliothèque varie en fonction du système d'exploitation. Le tableau suivant vous aide à procéder aux substitutions adéquates lorsque vous configurez ou dépannez votre système.

Table B-2
Chemin de bibliothèque en fonction du système d'exploitation

Système d'exploitation	Nom de la variable du chemin de bibliothèque
AIX	LIBPATH
HP-UX	SHLIB_PATH
Solaris ou Linux 32 bits	LD_LIBRARY_PATH
Solaris ou Linux 64 bits	LD_LIBRARY_PATH_64

Configuration d'un pilote pour le noeud Enterprise View

Le noeud Enterprise View vous permet de créer et de conserver une connexion entre une session IBM® SPSS® Modeler et un IBM SPSS Collaboration and Deployment Services Enterprise View dans un IBM® SPSS® Collaboration and Deployment Services Repository partagé. [Pour plus d'informations, reportez-vous à la section Stockage et et déploiement des objets IBM SPSS Collaboration and Deployment Services Repository dans le chapitre 9 dans *Guide de l'utilisateur de IBM SPSS Modeler 14.2*.](#)

Pour utiliser le noeud Enterprise View avec IBM® SPSS® Modeler Server, un IBM® SPSS® Collaboration and Deployment Services Enterprise View Driver doit être installé et configuré. Pour plus de détails sur l'installation du pilote, contactez votre administrateur local. Après l'installation, une référence à *pev.sh* doit être ajoutée au script de démarrage—dans *modelersrv.sh*, *rc.modeler* ou dans les deux—au format :

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



Exécution en tant que processus non-racine sous UNIX

Introduction

Ces instructions fournissent des informations sur l'exécution de IBM® SPSS® Modeler Server en tant que processus non-racine sur les systèmes UNIX.

Exécution en tant que processus racine. Dans l'installation de SPSS Modeler Server par défaut, le processus démon du serveur est exécuté en tant que processus racine. L'exécution en tant que processus racine permet à IBM® SPSS® Modeler d'authentifier de manière fiable chaque connexion utilisateur et de démarrer chaque session utilisateur sur le compte utilisateur UNIX correspondant. Les utilisateurs n'ont ainsi accès qu'à leurs propres fichiers et répertoires. Toutefois, il est possible d'exécuter le processus démon sur un compte non-racine en suivant les instructions ci-dessous.

Exécution en tant que processus non-racine. S'il n'existe pas de privilège racine, SPSS Modeler authentifie les utilisateurs grâce à une base de données de mots de passe privés, différente de la base de données de mots de passe UNIX, et démarre toutes les sessions utilisateur sur le même compte utilisateur UNIX. Cela signifie que toutes les données de fichier lues ou écrites par SPSS Modeler sont partagées par tous les utilisateurs de SPSS Modeler. L'accès aux données de la base de données n'est pas concerné car les utilisateurs doivent s'authentifier chaque fois qu'ils souhaitent accéder à l'une des sources de données.

Configuration de IBM SPSS Modeler Server en tant que processus non-racine

Exécution de IBM® SPSS® Modeler Server en tant que processus non-racine signifie que les ID utilisateur réels et effectifs du processus démon du serveur sont définis sur un compte de votre choix. Vous devez toujours vous connecter en tant qu'utilisateur root, ou définir su sur root, pour démarrer et arrêter le serveur, et pour effectuer des opérations d'administration générale.

Pour configurer SPSS Modeler Server en vue d'une exécution sur un compte non-racine, suivez les étapes ci-dessous :

1. Créez un groupe qui contienne tous vos utilisateurs. Vous pouvez nommer ce groupe comme vous le souhaitez, mais, dans cet exemple, appelons-le *modelerusers*.
2. Créez le compte utilisateur sur lequel exécuter SPSS Modeler Server. Ce compte est réservé au processus démon SPSS Modeler Server. Par conséquent, vous ne pouvez pas utiliser ce compte pour démarrer et arrêter le serveur, ou pour effectuer des opérations d'administration générale dans IBM® SPSS® Modeler. Dans cet exemple, appelons-le *modelerserv*.

Lorsque vous créez le compte, notez que :

- Le compte ne nécessite pas de fonctionnalité de connexion.
- Le groupe principal doit être le groupe `<modelers>` créé précédemment.
- Le répertoire personnel peut être le répertoire d'installation de SPSS Modeler ou tout autre répertoire par défaut pratique (utilisez un répertoire autre que le répertoire d'installation si vous souhaitez conserver votre compte d'une mise à niveau à une autre).

3. Ensuite, configurez des scripts de démarrage pour démarrer SPSS Modeler Server à l'aide du compte nouvellement créé. Recherchez le script de démarrage approprié et ouvrez-le dans un éditeur de texte. [Pour plus d'informations, reportez-vous à la section Scripts dans l'annexe B sur p. 66.](#)

- ▶ Dans le script approprié, modifiez la ligne qui définit la variable `MODELERUSER` pour désigner le nom du compte utilisateur créé pour exécuter SPSS Modeler Server. Par exemple :

```
MODELERUSER=modelerserv
```

- ▶ Modifiez le paramètre `umask` pour donner au groupe un accès en lecture sur les fichiers créés :

```
umask 027
```

4. Editez le fichier d'options du serveur, `config/options.cfg`, pour spécifier que l'authentification doit s'effectuer en fonction de la base de données de mots de passe privés. Pour ce faire, ajoutez la ligne :

```
authentication_methods, "pasw_modeler"
```

5. Vous devez ensuite créer une base de données de mots de passe privés, stockée dans le fichier `config/passwords.cfg`. Le fichier de mots de passe définit les combinaisons nom d'utilisateur/mot de passe permettant de se connecter à SPSS Modeler. *Remarque* : Ces données propres à SPSS Modeler n'ont aucun lien avec les noms d'utilisateur et les mots de passe utilisés pour se connecter à UNIX. Vous pouvez utiliser les mêmes noms d'utilisateur pour des raisons pratiques, mais pas les mêmes mots de passe.

Pour créer le fichier de mots de passe, utilisez l'utilitaire de création de mots de passe, `pwutil`, qui se trouve dans le répertoire `bin` de l'installation SPSS Modeler Server. La syntaxe de ce programme est la suivante :

```
pwutil [ nom d'utilisateur [ mot de passe ] ]
```

A partir d'un nom d'utilisateur et d'un mot de passe en texte brut, le programme écrit le nom d'utilisateur et le mot de passe codé dans une sortie standard dont le format est adapté au fichier de mots de passe. Par exemple, pour définir un utilisateur `modeler` avec le mot de passe "data mining", entrez la ligne suivante :

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

La définition d'un seul nom d'utilisateur est suffisante dans la plupart des cas. Les utilisateurs se connectent avec les mêmes nom et mot de passe. Toutefois, vous pouvez créer des utilisateurs supplémentaires à l'aide de l'opérateur >>, qui permet de les ajouter à la fin du fichier. Par exemple :

```
bin/pwutil modeler "data miner2" >> config/passwords.cfg
```

Remarque : Si un seul signe > est utilisé, le contenu de *passwords.cfg* est écrasé à chaque fois, remplaçant ainsi les utilisateurs précédemment définis. N'oubliez pas que tous les utilisateurs partagent le même compte utilisateur UNIX.

6. Attribuez, de façon récursive, la propriété du répertoire d'installation de SPSS Modeler et de son contenu à l'utilisateur <modelerserv> et au groupe <modelerusers>, les noms mentionnés étant ceux créés précédemment. Par exemple :

```
chown -R -h modelerserv:modelerusers .
```

7. Dans le répertoire de données, créez des sous-répertoires afin que vos utilisateurs SPSS Modeler disposent d'emplacements où stocker les données de travail sans risque. Ces répertoires doivent appartenir au groupe <modelerusers>, qui doit y avoir accès en lecture et en écriture, et pouvoir y effectuer des recherches. Par exemple, pour créer un répertoire de travail pour l'utilisateur *bob* :

```
mkdir data/bob
chown bob:modelerusers data/bob
chmod ug=rwx,o= data/bob
```

En outre, vous pouvez définir le bit set-group-ID du répertoire afin que tous les fichiers de données copiés dans le répertoire appartiennent automatiquement au groupe <modelerusers> :

```
chmod g+s data/bob
```

Exécution de IBM SPSS Modeler Server en tant qu'utilisateur non-racine sous AIX

Lors du démarrage de SPSS Modeler Server en tant qu'utilisateur non-racine sous AIX, il est possible que l'erreur *lsattr introuvable* soit imprimée sur la console. Cela s'explique par le fait que *lsattr* est situé dans le répertoire */usr/sbin*, mais lorsque SPSS Modeler Server connecte un nouvel utilisateur (pour une exécution en tant que non-racine) il réinitialise le chemin pour inclure simplement */usr/bin* et */bin*.

Si cette erreur se produit, la solution consiste à créer le lien suivant sous */usr/bin* :

```
su
cd /usr/bin
ln -s /usr/sbin/lsattr lsattr
```

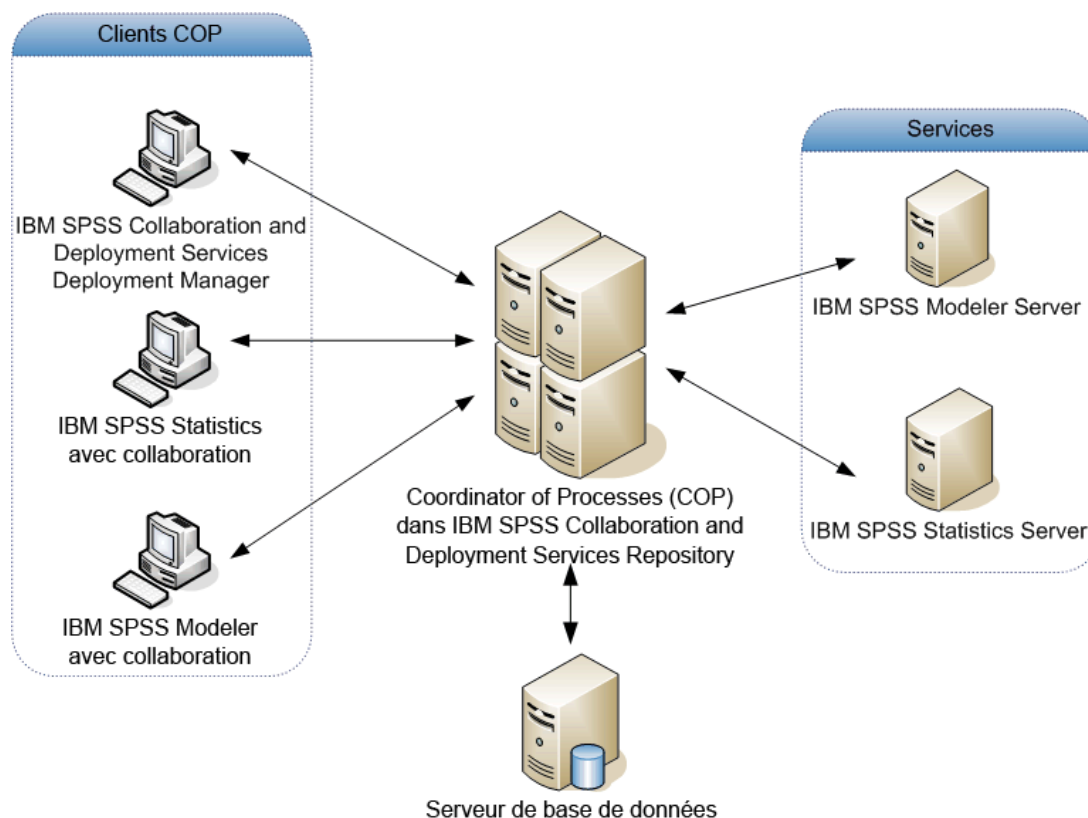
Une fois le lien créé, redémarrez SPSS Modeler Server.

Equilibrage de charge avec classes de serveur

Depuis IBM® SPSS® Collaboration and Deployment Services, un plug-in appelé Coordinateur de processus (COP) peut être utilisé pour gérer les services sur le réseau. Le Coordinateur de processus fournit des fonctionnalités de gestion de serveur conçues pour optimiser la communication client-serveur et le traitement.

Les services à gérer, tels IBM® SPSS® Statistics Server ou IBM® SPSS® Modeler Server, sont enregistrés avec le Coordinateur de processus au démarrage et envoient périodiquement des messages de statut mis à jour. Les services peuvent aussi stocker tout fichier de configuration nécessaire dans le IBM® SPSS® Collaboration and Deployment Services Repository et le récupérer lors de l'initialisation.

Figure D-1
Architecture du Coordinateur de processus



Exécuter vos flux IBM® SPSS® Modeler sur un serveur peut augmenter les performances. Dans certains cas, vous pouvez n'avoir le choix qu'entre un ou deux serveurs. Dans d'autres, vous pouvez avoir un choix de serveurs plus important car il existe une différence appréciable entre

chacun, tels le propriétaire, les droits d'accès, les données du serveur, les serveurs de tests ou les serveurs de production, etc. De plus, si votre réseau possède le Coordinateur de processus, vous pouvez vous voir proposer un groupe de serveurs.

Un groupe de serveurs contient plusieurs serveurs qui sont interchangeables en termes de configuration et de ressources. Le Coordinateur de processus détermine quel serveur convient le mieux à une demande de traitement grâce à un algorithme qui équilibre la charge en fonction de plusieurs critères, parmi lesquels la pondération des serveurs, les priorités de l'utilisateur et la charge des traitements en cours. Pour plus d'informations, consultez le *Guide des développeurs du service Coordinateur de processus*, disponible sur demande chez .

A chaque fois que vous vous connectez à un serveur ou à un groupe de serveurs SPSS Modeler, vous pouvez saisir un serveur manuellement ou rechercher un serveur ou un groupe de serveurs utilisant le Coordinateur de processus. [Pour plus d'informations, reportez-vous à la section Connexion au IBM SPSS Modeler Server dans le chapitre 3 dans *Guide de l'utilisateur de IBM SPSS Modeler 14.2*.](#)



Avis

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.

Le paragraphe suivant ne s'applique pas au Royaume Uni ni à aucun pays dont la réglementation locale est incompatible avec les dispositions suivantes : SPSS INC., AN IBM COMPANY, FOURNIT CETTE PUBLICATION « EN L'ÉTAT » SANS GARANTIE D'AUCUNE SORTE, EXPLICITE OU IMPLICITE, Y COMPRIS, MAIS SANS SE LIMITER À, LES GARANTIES TACITES DE NON CONTREFAÇON, DE CONFORMITÉ OU D'APTITUDE À UN USAGE PARTICULIER. Certains états n'autorisent pas les clauses de non garantie explicites ou implicites lors de certaines transactions, par conséquent, il est possible que cette disposition ne vous concerne pas.

Ces informations peuvent contenir des erreurs techniques ou typographiques. Ces informations sont modifiées périodiquement ; ces modifications seront incluses dans les nouvelles versions de cette publication. SPSS Inc. a le droit d'apporter des améliorations et/ou des modifications au(x) produit(s) et/ou au(x) programme(s) décrits dans cette publication à tout moment et sans avertissement préalable.

Toute référence, contenue dans ces informations, à des sites Internet autres que ceux de SPSS et IBM est fournie pour des raisons de commodité uniquement et ne peut, en aucune sorte, être considérée comme de la publicité pour ces sites Internet. Le matériel de ces sites Internet ne fait pas partie du matériel de ce produit SPSS Inc. et l'utilisation de ces sites Internet se fait à vos risques et périls.

Lorsque vous envoyez des informations à IBM ou SPSS, vous octroyez à IBM et SPSS le droit non exclusif d'utiliser ou de distribuer ces informations de la façon dont IBM ou SPSS le juge approprié sans aucune obligation envers vous.

Les informations sur des produits autres que SPSS ont été obtenues auprès des fournisseurs de ces produits, à partir de leurs annonces publiées ou d'autres sources publiques disponibles. SPSS n'a pas effectué de test sur ces produits et ne peut confirmer la validité de leurs performances, de leur compatibilité ou de toute autre prétention associée aux produits autres que SPSS. Les questions concernant les performances des produits autres que SPSS devront être adressées aux fournisseurs de ces produits.

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.

Ces informations contiennent des exemples de données et de rapports utilisées dans les opérations quotidiennes de l'entreprise. Pour les illustrer le plus précisément possible, ces exemples contiennent les noms de personnes, d'entreprises, de marques et de produits. Tous ces noms sont fictifs et toute ressemblance avec des noms et adresses utilisés par une entreprise réelle ne serait que pure coïncidence.

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

Marques commerciales

IBM, le logo IBM et [ibm.com](http://www.ibm.com) sont des marques commerciales de IBM Corporation, déposées dans de nombreuses juridictions à travers le monde. Une liste mise à jour des marques commerciales IBM est disponible sur Internet à l'adresse <http://www.ibm.com/legal/copytrade.shtml>.

SPSS est une marque commerciale de SPSS Inc., an IBM Company, déposées dans de nombreuses juridictions à travers le monde.

Adobe, le logo Adobe, PostScript et le logo PostScript sont des marques déposées ou des marques commerciales de Adobe Systems Incorporated aux États-Unis et/ou dans d'autres pays.

IT Infrastructure Library est une marque déposée du Central Computer and Telecommunications Agency qui fait désormais partie du Office of Government Commerce.

Intel, le logo Intel, Intel Inside, le logo Intel Inside, Intel Centrino, le logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium et Pentium sont des marques commerciales ou déposées de Intel Corporation ou de ses filiales aux États-Unis ou dans d'autres pays.

Linux est une marque déposée de Linus Torvalds aux États-Unis, dans d'autres pays, ou les deux.

Microsoft, Windows, Windows NT et le logo Windows sont des marques de Microsoft Corporation aux États-Unis, dans d'autres pays, ou les deux.

ITIL est une marque déposée et une marque communautaire déposée du Office of Government Commerce et a été enregistrée auprès du Bureau américain des brevets et des marques de commerce.

UNIX est une marque déposée de The Open Group aux États-Unis et dans d'autres pays.

Cell Broadband Engine est une marque de Sony Computer Entertainment, Inc. aux États-Unis, dans d'autres pays ou les deux et est par conséquent sous contrat de licence.

Java et les marques et logo Java sont des marques de Sun Microsystems, Inc. aux États-Unis, dans d'autres pays, ou les deux.

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.

D'autres noms de produits ou de services peuvent être des marques de IBM, SPSS ou d'autres entreprises.

Index

- accès administrateur
 - pour IBM SPSS Modeler Server, 25
- accès aux données, 11
- administration
 - de IBM SPSS Modeler Server, 23
 - de l'emplacement de la licence IBM SPSS Statistics, 36
- ajout de connexions à IBM SPSS Modeler Server, 16–17
- allow_modelling_memory_override
 - fichier *options.cfg*, 29
- aperçu
 - génération SQL, 53
- architecture
 - composants, 7
- arrêt de IBM SPSS Modeler Server
 - sous UNIX, 21
 - sous Windows, 21
- authentification, 19
- authentification de l'utilisateur, 19
- auto.sh* (UNIX)
 - emplacement, 66
- avis juridiques, 81

- base de données
 - niveaux de prise en charge, 11
- Bases de données
 - accès, 11

- cache_compression
 - fichier *options.cfg*, 29
- chemins, 13
- chemsrv.sh* (UNIX)
 - emplacement, 66
- codage
 - SSL, 36
- codage de données SSL
 - activation pour IBM SPSS Modeler Server, 32
- compression du cache, 29
- comptes utilisateur
 - IBM SPSS Modeler Server, 19
- configuration du coordinateur de processus
 - pour IBM SPSS Modeler Server, 32
- Configuration du COP
 - pour IBM SPSS Modeler Server, 32
- Connect for ODBC
 - configuration sous UNIX, 68
- connexion à IBM SPSS Modeler Server, 14
- connexion unique, 16
- connexions
 - à IBM SPSS Modeler Server, 14, 16–17
 - groupe de serveurs, 17
- connexions à la base de données
 - fermeture, 35
- Coordinateur de processus, 17
 - équilibre de charge, 79
 - groupes de serveurs, 79

- COP, 17
 - équilibre de charge, 79
 - groupes de serveurs, 79
- cop_enabled
 - fichier *options.cfg*, 32
- cop_host
 - fichier *options.cfg*, 32
- cop_password
 - fichier *options.cfg*, 32
- cop_port_number
 - fichier *options.cfg*, 32
- cop_service_description
 - fichier *options.cfg*, 32
- cop_service_host
 - fichier *options.cfg*, 32
- cop_service_name
 - fichier *options.cfg*, 32
- cop_service_weight
 - fichier *options.cfg*, 32
- cop_update_interval
 - fichier *options.cfg*, 32
- cop_user_name
 - fichier *options.cfg*, 32

- Data Access Pack
 - chemins de bibliothèque UNIX, 75
 - configuration d'UNIX, 68
 - Connect for ODBC, configuration sous UNIX, 68
 - dépannage de la configuration Connect for ODBC sous UNIX, 71
- data_file_path
 - fichier *options.cfg*, 28
- data_files_restricted
 - fichier *options.cfg*, 28
- DB2
 - optimisation SQL, 49–50
- démarrage automatique du serveur
 - configuration sous UNIX, 67
- démarrage de IBM SPSS Modeler Server
 - sous UNIX, 21
 - sous Windows, 21
- disques durs, 11
- documentation, 2
- droits d'accès aux fichiers
 - sous IBM SPSS Modeler Server, 19
- droits d'accès sur les fichiers
 - configuration sous UNIX, 68

- emplacement de la licence IBM SPSS Statistics, 36
- erreur d'exécution de flux, 31
- Espace disque, 10
 - calcul, 10
- état
 - de IBM SPSS Modeler Server sous UNIX, 21
 - de IBM SPSS Modeler Server sous Windows, 21

- exécution de flux multiples, 31
- exemples
 - Aperçu, 4
- Exemples
 - Guide des applications, 2
- exemples d'application, 2
- exploration des bases de données, 47
- expressions CLEM
 - génération SQL, 60

- fichier *messages.log*, 42
- fichier *options.cfg*, 34
- fichiers de données
 - IBM SPSS Statistics, 13
 - importation et exportation, 13
- fichiers journaux
 - affichage du code SQL généré, 47, 53
 - pour IBM SPSS Modeler Server, 27, 42
- fichiers temporaires, 10
 - autorisations relatives à (IBM SPSS Modeler Server), 19
- fonction @SQLFN, 62

- génération SQL, 45, 49–50, 57
 - activation, 52
 - activation pour IBM SPSS Modeler Server, 31
 - affichage pour des nuggets de modèle, 55
 - astuces, 55
 - consignation, 47, 53
 - expressions CLEM, 55, 60
 - prévisualisation, 47, 53
 - réécriture des flux, 55
- gestion de la mémoire
 - options d'administration, 29

- IBM SPSS Modeler, 1
 - documentation, 2
- IBM SPSS Modeler Administration Console, 24
 - accès administrateur, 25
- IBM SPSS Modeler Server
 - accès administrateur, 25
 - administration, 24
 - authentification de l'utilisateur, 19
 - comptes utilisateur, 19
 - configuration du coordinateur de processus, 32
 - Configuration du COP, 32
 - création de fichiers, 19
 - fichiers journaux, 27
 - ID utilisateur, 14
 - informations pour les utilisateurs finaux, 18
 - mot de passe, 14
 - nom de domaine (Windows), 14
 - nom d'hôte, 14, 16
 - numéro de port, 14, 16, 27
 - options d'administration, 23
 - options de configuration, 27
 - processus qui ne répondent pas, 23
 - processus serveur, 34
 - répertoire temporaire, 28
 - résultats différents de ceux de Clementine Client, 20
 - surveillance de l'utilisation, 34
- IBM SPSS Statistics fichiers de données
 - importation et exportation, 13
- IBM SPSS Text Analytics, 2
- ID utilisateur
 - IBM SPSS Modeler Server, 14
- io_buffer_size
 - fichier *options.cfg*, 29

- LDAP, 41
 - sécurisation, 41
- limites du noyau sous UNIX, 22
- limites du noyau UNIX, 22
- log_database_access
 - fichier *options.cfg*, 27
- log_database_errors
 - fichier *options.cfg*, 27
- log_directory
 - fichier *options.cfg*, 27
- log_file_access
 - fichier *options.cfg*, 27
- log_file_errors
 - fichier *options.cfg*, 27

- marques commerciales, 82
- max_file_size
 - fichier *options.cfg*, 28
- max_login_attempts
 - fichier *options.cfg*, 27
- max_parallelism
 - fichier *options.cfg*, 29
- max_sessions
 - fichier *options.cfg*, 27
- max_sql_string_length
 - fichier *options.cfg*, 31
- Mémoire, 11
- memory_usage
 - fichier *options.cfg*, 29
- messages
 - affichage du code SQL généré, 47, 53
- Microsoft SQL Server
 - optimisation SQL, 49–50
- mise en cache de base de données
 - controlling from options.cfg, 35
 - génération SQL, 56
- mise en mémoire cache des bases de données, 35
- mise en mémoire cache des noeuds
 - écriture dans la base de données, 56
 - génération SQL, 56
- mise en mémoire cache, bases de données, 35
- modélisation
 - gestion de la mémoire, 29
- modelling_memory_limit_percentage
 - fichier *options.cfg*, 29

- mot de passe
 - IBM SPSS Modeler Server, 14
- niveaux, prise en charge de la base de données, 11
- noeuds
 - prenant en charge la génération SQL, 57
- noeuds violets, 45
- optimisation SQL, 51
- nom de domaine (Windows)
 - IBM SPSS Modeler Server, 14
- nom d'hôte
 - IBM SPSS Modeler Server, 14, 16
- noms de fichier
 - UNIX, 13
 - Windows, 13
- noms de fichier UNC, 13
- nuggets de modèle
 - affichage SQL pour, 55
- numéro de port
 - IBM SPSS Modeler Server, 14, 16, 27
- opérateurs
 - génération SQL, 60
- optimisation, 45
 - génération SQL, 49–50, 52
- option `cache_connection`, 35
- options de configuration
 - Accès aux fichiers de données, 28
 - Aperçu, 27
 - codage de données SSL, 32
 - connexions et sessions, 27
 - consignation, 27
 - coordinateur de processus, 32
 - COP, 32
 - de IBM SPSS Modeler Server, 23
 - génération SQL automatique, 31
 - gestion de la mémoire, 29
 - longueur de chaîne SQL, 31
 - numéro de port, 27
 - performances et optimisation, 29
 - réécriture des flux, 29
 - répertoire temporaire, 28
 - tentatives de connexion, 27
 - traitement parallèle, 29
- Oracle
 - optimisation SQL, 49–50, 64
- paramètres de port
 - fichier *options.cfg*, 29
- Paramètres de port du serveur
 - fichier *options.cfg*, 29
- paramètres du pare-feu
 - fichier *options.cfg*, 31
- Performances
 - de IBM SPSS Modeler Server, 44
- port_number
 - fichier *options.cfg*, 27
- Préfixe d'URL, 41
- processeurs, 9
 - Multiple, 29
- processus zombies, IBM SPSS Modeler Server, 23
- processus, ne répondant pas, 23
- program_file_path
 - fichier *options.cfg*, 28
- program_files_restricted
 - fichier *options.cfg*, 28
- RAM, 11
- rc.modeler* (UNIX)
 - emplacement, 66
- recherche de connexions dans COP, 17
- recommandations matérielles
 - pour IBM SPSS Modeler Server, 8
- réécriture des flux, 55
 - activation, 45
- répercussions, 45, 49–50, 52, 57
 - expressions CLEM, 60
- Répercussions SQL. *Consultez également* Génération SQL, 49
- répertoire temporaire
 - pour IBM SPSS Modeler Server, 28
- résultats
 - arrondi, 20
 - différences entre Clementine Client et Clementine Server, 20
 - ordre des enregistrements, 20
- scripts UNIX
 - auto.sh*, 66
 - modelersrv.sh*, 66
 - modification, 68
 - rc.modeler*, 66
- sécurisation
 - LDAP, 41
- sécurité
 - configuration de la création de fichier sous UNIX, 68
 - création de fichiers, 19
 - SSL, 36
- serveur
 - ajout de connexions, 16
 - connexion, 14
 - recherche de serveurs dans COP, 17
- serveurs de base de données, 47
- sources de données ODBC
 - Connect for ODBC et scripts UNIX, 68
 - et UNIX, 68
- SPSS Modeler Server, 1
- SQL
 - affichage pour des nuggets de modèle, 55
 - noms de colonne en double, 63
 - optimisation d'Oracle, 64

-
- prévisualisation générée, 53
 - requêtes, 63
 - SQL Server
 - optimisation SQL, 49–50
 - sql_generation_enabled
 - fichier *options.cfg*, 31
 - SSL, 36
 - Aperçu, 37
 - sécurisation des communications, 37
 - SSL (Secure Sockets Layer), 36
 - ssl_certificate_file
 - fichier *options.cfg*, 32
 - ssl_enabled
 - fichier *options.cfg*, 32
 - ssl_private_key_file
 - fichier *options.cfg*, 32
 - ssl_private_key_password
 - fichier *options.cfg*, 32
 - stream_rewriting_enabled
 - fichier *options.cfg*, 29
 - systèmes d'exploitation
 - 64 bits, 9
 - systèmes d'exploitation 64 bits, 9
-
- technologie IBM SPSS Statistics Data Access, 11
 - temp_directory
 - fichier *options.cfg*, 28
 - traitement parallèle
 - activation, 45
 - contrôle, 29
-
- UNIX
 - authentification de l'utilisateur, 19
 - chemins de bibliothèque, 75
 - configuration des droits d'accès aux fichiers, 68