

*Guide des performances et
d'administration d'IBM SPSS Modeler
Server 17.1*



Remarque

Avant d'utiliser le présent document et le produit associé, prenez connaissance des informations générales figurant à la section «Remarques», à la page 95.

Certaines illustrations de ce manuel ne sont pas disponibles en français à la date d'édition.

LE PRESENT DOCUMENT EST LIVRE EN L'ETAT SANS AUCUNE GARANTIE EXPLICITE OU IMPLICITE. IBM DECLINE NOTAMMENT TOUTE RESPONSABILITE RELATIVE A CES INFORMATIONS EN CAS DE CONTREFACON AINSI QU'EN CAS DE DEFAUT D'APTITUDE A L'EXECUTION D'UN TRAVAIL DONNE.

Ce document est mis à jour périodiquement. Chaque nouvelle édition inclut les mises à jour. Les informations qui y sont fournies sont susceptibles d'être modifiées avant que les produits décrits ne deviennent eux-mêmes disponibles. En outre, il peut contenir des informations ou des références concernant certains produits, logiciels ou services non annoncés dans ce pays. Cela ne signifie cependant pas qu'ils y seront annoncés.

Pour plus de détails, pour toute demande d'ordre technique, ou pour obtenir des exemplaires de documents IBM, référez-vous aux documents d'annonce disponibles dans votre pays, ou adressez-vous à votre partenaire commercial.

Vous pouvez également consulter les serveurs Internet suivants :

- <http://www.fr.ibm.com> (serveur IBM en France)
- <http://www.ibm.com/ca/fr> (serveur IBM au Canada)
- <http://www.ibm.com> (serveur IBM aux Etats-Unis)

*Compagnie IBM France
Direction Qualité
17, avenue de l'Europe
92275 Bois-Colombes Cedex*

Cette édition s'applique à la version 17.1.0 d'IBM(r) SPSS(r) Modeler et à toutes les éditions et modifications ultérieures sauf mention contraire dans les éditions suivantes.

Table des matières

Avis aux lecteurs canadiens	v	Traitement des processus de serveurs qui ne répondent pas (systèmes UNIX)	28
Préface	vii	Configuration de profils de serveur	29
Chapitre 1. A propos d'IBM SPSS Modeler	1	Utilisation des profils de serveur	29
Produits IBM SPSS Modeler	1	Structure de profil	31
IBM SPSS Modeler	1	Scripts de profil	33
IBM SPSS Modeler Server	1	Administration	37
IBM SPSS Modeler Administration Console	2	Utilisation de IBM SPSS Modeler Administration Console	38
IBM SPSS Modeler Batch	2	Démarrage de Modeler Administration Console	38
IBM SPSS Modeler Solution Publisher	2	Configuration de l'accès avec Modeler Administration Console	38
Adaptateurs IBM SPSS Modeler Server pour IBM SPSS Collaboration and Deployment Services	2	SPSS Modeler Server Connexions	39
Editions de IBM SPSS Modeler	2	Configuration de SPSS Modeler Server	39
Documentation de IBM SPSS Modeler	3	SPSS Modeler Server Surveillance	46
Documentation de SPSS Modeler Professional	3	Utilisation du fichier options.cfg	46
Documentation de SPSS Modeler Premium	4	Fermeture de connexions inutilisées à une base de données	47
Exemples d'application	5	Utilisation du protocole SSL pour sécuriser le transfert de données	48
Dossier Demos	5	Fonctionnement de SSL	48
Chapitre 2. Recommandations relatives à l'architecture et au matériel	7	Sécurisation des communications client-serveur et serveur-serveur avec le protocole SSL	48
Architecture de IBM SPSS Modeler	7	Connexion SSL Cognos BI	52
Description de l'architecture	7	Connexion SSL Cognos TM1.	52
Recommandations matérielles	9	Configuration des groupes	53
Exigences en termes d'espace disque temporaire et de mémoire RAM	10	Journal du serveur	55
Accès aux données	11	Chapitre 5. Considérations relatives aux performances	57
Référencement des fichiers de données	12	Paramètres de performances et d'optimisation du serveur	57
Importation des fichiers de données IBM SPSS Statistics	13	Paramètres de performances et d'optimisation du client	57
Instructions d'installation	13	Utilisation et optimisation de la base de données	59
Chapitre 3. Support IBM SPSS Modeler	15	Optimisation SQL	59
Connexion à IBM SPSS Modeler Server	15	Chapitre 6. Optimisation SQL	61
Configuration de la connexion unique	16	Fonctionnement de la génération SQL	62
Ajout et modification d'une connexion à IBM SPSS Modeler Server	21	Exemple de génération SQL	63
Recherche de serveurs dans IBM SPSS Collaboration and Deployment Services	22	Configuration de l'optimisation SQL	64
Systèmes de données et de fichiers	23	Prévisualisation du code SQL généré.	65
Authentification de l'utilisateur	23	Affichage SQL pour des nuggets de modèle	65
Droits	24	Astuces relatives à l'optimisation de la génération SQL	65
Création de fichiers.	24	Noeuds prenant en charge la génération SQL	66
Différences au niveau des résultats	25	Opérateurs et expressions CLEM prenant en charge la génération SQL	70
Chapitre 4. Administration de IBM SPSS Modeler	27	Utilisation de fonctions SQL dans les expressions CLEM	73
Démarrage et arrêt de IBM SPSS Modeler Server	27	Rédaction de requêtes SQL	73
Pour démarrer ou arrêter le serveur sous Windows, ou vérifier son statut.	27	Adaptateur de scoring pour Teradata - Lignes dupliquées	74
Pour démarrer ou arrêter le serveur sous UNIX, ou vérifier son statut	27		

Annexe A. Configuration d'Oracle pour les plates-formes UNIX.	75
Configuration d'Oracle pour l'optimisation SQL	75

Annexe B. Configuration des scripts de démarrage UNIX.	77
---	-----------

Introduction	77
Scripts	77
Démarrage et arrêt automatiques de IBM SPSS Modeler Server	77
Démarrage et arrêt manuels de IBM SPSS Modeler Server	78
Edition de scripts	78
Contrôle des droits d'accès relatifs à la création de fichier	78
IBM SPSS Modeler Server et Data Access Pack	78
Dépannage de la configuration ODBC	81
Chemins de bibliothèque	84
Configuration d'un pilote pour le noeud Enterprise View	84

Annexe C. Configuration et exécution de SPSS Modeler Server en tant que processus non-racine sous UNIX	85
---	-----------

Introduction	85
Configuration en tant qu'utilisateur non racine sans base de données de mots de passe privés	85

Configuration en tant qu'utilisateur non-racine avec une base de données de mots de passe privés	86
Exécution de SPSS Modeler Server en tant qu'utilisateur non-racine	87
Dépannage des échecs d'authentification de l'utilisateur	88

Annexe D. Configuration et exécution de SPSS Modeler Server avec un fichier de mots de passe privés sous Windows.	89
--	-----------

Introduction	89
Configuration d'une base de données de mots de passe privés	89

Annexe E. Equilibrage de charge avec clusters de serveur	91
---	-----------

Annexe F. Authentification LDAP	93
--	-----------

Remarques	95
Marques	96

Index	99
--------------	-----------

Avis aux lecteurs canadiens

Le présent document a été traduit en France. Voici les principales différences et particularités dont vous devez tenir compte.

Illustrations

Les illustrations sont fournies à titre d'exemple. Certaines peuvent contenir des données propres à la France.

Terminologie

La terminologie des titres IBM peut différer d'un pays à l'autre. Reportez-vous au tableau ci-dessous, au besoin.

IBM France	IBM Canada
ingénieur commercial	représentant
agence commerciale	succursale
ingénieur technico-commercial	informaticien
inspecteur	technicien du matériel

Claviers

Les lettres sont disposées différemment : le clavier français est de type AZERTY, et le clavier français-canadien de type QWERTY.








OS/2 et Windows - Paramètres canadiens

Au Canada, on utilise :

- les pages de codes 850 (multilingue) et 863 (français-canadien),
- le code pays 002,
- le code clavier CF.

Nomenclature

Les touches présentées dans le tableau d'équivalence suivant sont libellées différemment selon qu'il s'agit du clavier de la France, du clavier du Canada ou du clavier des États-Unis. Reportez-vous à ce tableau pour faire correspondre les touches françaises figurant dans le présent document aux touches de votre clavier.

France	Canada	Etats-Unis
 (Pos1)		Home
Fin	Fin	End
 (PgAr)		PgUp
 (PgAv)		PgDn
Inser	Inser	Ins
Suppr	Suppr	Del
Echap	Echap	Esc
Attn	Intrp	Break
Impr écran	ImpEc	PrtSc
Verr num	Num	Num Lock
Arrêt défil	Défil	Scroll Lock
 (Verr maj)	FixMaj	Caps Lock
AltGr	AltCar	Alt (à droite)

Brevets

Il est possible qu'IBM détienne des brevets ou qu'elle ait déposé des demandes de brevets portant sur certains sujets abordés dans ce document. Le fait qu'IBM vous fournisse le présent document ne signifie pas qu'elle vous accorde un permis d'utilisation de ces brevets. Vous pouvez envoyer, par écrit, vos demandes de renseignements relatives aux permis d'utilisation au directeur général des relations commerciales d'IBM, 3600 Steeles Avenue East, Markham, Ontario, L3R 9Z7.

Assistance téléphonique

Si vous avez besoin d'assistance ou si vous voulez commander du matériel, des logiciels et des publications IBM, contactez IBM direct au 1 800 465-1234.

Préface

IBM® SPSS Modeler est le puissant utilitaire d'exploration de données 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 offre 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 d'IBM Business Analytics

Le logiciel IBM Business Analytics propose des informations complètes, cohérentes et précises auxquelles les preneurs de décisions peuvent se fier pour améliorer les performances de leur entreprise. Un porte-feuilles étendu de veille économique, d'analyses prédictives, de gestion des performances et de stratégie financières et d'applications analytiques vous offre des informations claires, immédiates et décisionnelles sur les performances actuelles et vous permet de prévoir les résultats futurs. Ce logiciel intègre des solutions dédiées à l'industrie, des pratiques éprouvées et des services professionnels qui permettent aux organisations de toute taille de maximiser leur productivité, d'automatiser leurs décisions sans risque et de proposer de meilleurs résultats.

Ce porte-feuilles intègre le logiciel IBM SPSS Predictive Analytics qui aide les organisations à prévoir les événements à venir et à réagir en fonction des informations afin d'améliorer leurs résultats. Les clients de l'industrie du commerce, de l'éducation et des administrations du monde entier font confiance à la technologie IBM SPSS qui offre un avantage concurrentiel en attirant et fidélisant les clients et en améliorant la base de données de la clientèle tout en diminuant la fraude et en réduisant les risques. En utilisant le logiciel IBM SPSS dans leurs opérations quotidiennes, les organisations deviennent des entreprises prédictives, capables de diriger et d'automatiser les décisions pour répondre aux objectifs commerciaux et obtenir un avantage concurrentiel mesurable. Pour des informations supplémentaires ou pour joindre un représentant, consultez <http://www.ibm.com/spss>.

Assistance technique

L'assistance technique est disponible pour les clients du service de maintenance. 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 contacter l'assistance technique, rendez-vous sur le site Web IBM Corp. à l'adresse <http://www.ibm.com/support>. Lorsque vous contactez l'assistance technique, soyez prêt à indiquer votre identité, le nom de votre société et votre contrat d'assistance.

Chapitre 1. A propos d'IBM SPSS Modeler

IBM SPSS Modeler est un ensemble d'outils d'exploration de données 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, IBM SPSS Modeler prend en charge l'intégralité du processus d'exploration de données, des données à l'obtention de meilleurs résultats commerciaux.

IBM 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 tant que client en combinaison avec SPSS Modeler Server. Plusieurs autres options sont également disponibles, telles que décrites dans les sections suivantes. Pour plus d'informations, voir <http://www.ibm.com/software/analytics/spss/products/modeler/>.

Produits IBM SPSS Modeler

La famille des produits IBM SPSS Modeler et les logiciels associés sont composés des éléments suivants.

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- Adaptateurs IBM SPSS Modeler Server pour IBM SPSS Collaboration and Deployment Services

IBM SPSS Modeler

SPSS Modeler est une version complète du produit que vous installez et exécutez sur votre ordinateur personnel. Pour obtenir de meilleures performances lors du traitement de jeux de données volumineux, vous pouvez exécuter SPSS Modeler en mode local, comme produit autonome, ou l'utiliser en mode réparti, en association avec IBM SPSS Modeler Server.

Avec SPSS Modeler, vous pouvez créer des modèles prédictifs précis rapidement et de manière intuitive, sans aucune programmation. L'interface visuelle unique vous permet de visualiser facilement le processus d'exploration de données. Grâce aux analyses avancées intégrées au produit, vous pouvez découvrir des motifs et tendances masqués dans vos données. Vous pouvez modéliser les résultats et comprendre les facteurs qui les influencent, afin d'exploiter les opportunités commerciales et de réduire les risques.

SPSS Modeler est proposé dans deux éditions : SPSS Modeler Professional et SPSS Modeler Premium. Pour plus d'informations, voir la rubrique «Editions de IBM SPSS Modeler», à la page 2.

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 jeux de données plus volumineux.

SPSS Modeler Server est un produit avec licence distincte qui s'exécute en permanence en mode d'analyse réparti sur un hôte de serveur en combinaison avec une ou plusieurs installations de IBM SPSS Modeler. Ainsi, SPSS Modeler Server fournit des performances supérieures sur de grands jeux de données car les opérations nécessitant beaucoup de mémoire peuvent être effectuées sur le serveur sans télécharger de données sur l'ordinateur client. IBM SPSS Modeler Server prend également en charge l'optimisation SQL et propose des capacités de modélisation dans la base de données pour des performances et une automatisation améliorées.

IBM SPSS Modeler Administration Console

Modeler Administration Console est une application graphique permettant de gérer de nombreuses options de SPSS Modeler Server qui peuvent également être configurées au moyen d'un fichier d'options. Cette application 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.

IBM SPSS Modeler Batch

Alors que l'exploration de données est généralement un processus interactif, il est également possible d'exécuter SPSS Modeler à partir d'une ligne de commande sans recourir à l'interface utilisateur graphique. Par exemple, vous pouvez avoir des tâches longue durée ou répétitives à exécuter sans intervention de l'utilisateur. SPSS Modeler Batch est une version spécifique du produit qui prend en charge toutes les capacités d'analyse de SPSS Modeler sans avoir besoin d'accéder à l'interface utilisateur standard. SPSS Modeler Server est requis pour utiliser SPSS Modeler Batch.

IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher est un outil qui permet de créer une version « packagée » d'un flux SPSS Modeler qui peut être exécutée par un moteur Runtime externe ou intégrée dans une application externe. Ainsi, vous pouvez publier et déployer des flux SPSS Modeler complets dans des environnements où SPSS Modeler n'est pas installé. SPSS Modeler Solution Publisher est fourni avec le service IBM SPSS Collaboration and Deployment Services - Scoring et nécessite une licence distincte. Avec cette licence, vous recevez SPSS Modeler Solution Publisher Runtime qui vous permet d'exécuter les flux publiés.

Pour plus d'informations sur SPSS Modeler Solution Publisher, reportez-vous à la documentation IBM SPSS Collaboration and Deployment Services. IBM SPSS Collaboration and Deployment Services Knowledge Center contient des sections appelées "IBM SPSS Modeler Solution Publisher" et "IBM SPSS Analytics Toolkit."

Adaptateurs IBM SPSS Modeler Server pour IBM SPSS Collaboration and Deployment Services

Différents adaptateurs pour IBM SPSS Collaboration and Deployment Services sont disponibles et permettent à SPSS Modeler et SPSS Modeler Server d'interagir avec un référentiel IBM SPSS Collaboration and Deployment Services. Ainsi, un flux SPSS Modeler déployé sur le référentiel peut être partagé par différents utilisateurs ou peut être accessible depuis l'application client léger IBM SPSS Modeler Advantage. Installez l'adaptateur sur le système qui héberge le référentiel.

Editions de IBM SPSS Modeler

SPSS Modeler est disponible dans les éditions suivantes.

SPSS Modeler Professional

SPSS Modeler Professional offre tous les outils nécessaires à l'utilisation de la plupart des types de données structurées, tels que les comportements et interactions suivis dans les systèmes CRM, les

caractéristiques sociodémographiques, les comportements d'achat et les données de vente.

SPSS Modeler Premium

SPSS Modeler Premium est un produit avec licence distincte qui étend le champ d'applications de SPSS Modeler Professional afin de pouvoir traiter des données spécialisées telles que celles utilisées pour les analyses d'entités ou les réseaux sociaux ainsi que des données de texte non structurées. SPSS Modeler Premium comprend les composants suivants :

IBM SPSS Modeler Entity Analytics ajoute une dimension supplémentaire aux analyses prédictives IBM SPSS Modeler. Alors que les analyses prédictives essaient de prévoir les comportements futurs à partir de données passées, les analyses d'entités se concentrent sur l'amélioration de la cohérence des données actuelles en résolvant les conflits d'identités dans les enregistrements eux-mêmes. Une identité peut être celle d'un individu, d'une organisation, d'un objet ou d'une autre entité pour laquelle une ambiguïté peut exister. La résolution d'identité peut être vitale dans de nombreux domaines, y compris la gestion de la relation client, la détection de la fraude, le blanchiment d'argent et la sécurité nationale et internationale.

IBM SPSS Modeler Social Network Analysis transforme les informations sur les relations en champs qui caractérisent le comportement social des individus et des groupes. Grâce aux données qui décrivent les relations qui sous-tendent les réseaux sociaux, IBM SPSS Modeler Social Network Analysis identifie les chefs sociaux qui influencent le comportement des autres individus du réseau. De plus, il est possible de déterminer les individus qui sont le plus influencés par les autres participants du réseau. En combinant ces résultats avec d'autres mesures, il est possible de créer des profils détaillés des individus sur lesquels baser vos modèles prédictifs. Les modèles qui contiennent ces informations sociales seront plus efficaces que les modèles qui en sont dépourvus.

IBM SPSS Modeler Text Analytics utilise des technologies linguistiques avancées et le traitement du langage naturel pour traiter rapidement une large variété de données textuelles non structurées, en extraire les concepts clés et les organiser pour les regrouper dans des 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 d'exploration de données de IBM SPSS Modeler, afin de favoriser une prise de décision précise et efficace.

Documentation de IBM SPSS Modeler

Une documentation 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, ainsi que le Guide des applications (aussi appelé Tutoriel) et d'autres documentations utiles.

La documentation complète de chaque produit (y compris les instructions d'installation) au format PDF est disponible dans le dossier *\Documentation* de chaque DVD de produit. Les documents d'installation peuvent également être téléchargés depuis le Web sur le site <http://www.ibm.com/support/docview.wss?uid=swg27043831>.

La documentation sous ces deux formats est également disponible depuis le site SPSS Modeler Knowledge Center à l'adresse http://www-01.ibm.com/support/knowledgecenter/SS3RA7_17.0.0.0.

Documentation de SPSS Modeler Professional

La suite de documentation SPSS Modeler Professional (à l'exception des instructions d'installation) est la suivante.

- **Guide d'utilisation d'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 d'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.
- **Noeuds modélisation d'IBM SPSS Modeler.** Descriptions de tous les noeuds utilisés pour créer des modèles d'exploration de données. IBM SPSS Modeler propose différentes méthodes de modélisation issues des domaines de l'apprentissage automatique, de l'intelligence artificielle et des statistiques.
- **Guide des algorithmes d'IBM SPSS Modeler.** Descriptions des fondements mathématiques des méthodes de modélisation utilisées dans IBM SPSS Modeler. Ce guide est disponible au format PDF uniquement.
- **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, voir la rubrique «Exemples d'application», à la page 5.
- **Guide de génération de scripts Python et d'automatisation IBM SPSS Modeler.** Ce manuel fournit des informations sur l'automatisation du système via des scripts Python, notamment sur les propriétés pouvant être utilisées pour manipuler les noeuds et les flux.
- **Guide de déploiement d'IBM SPSS Modeler.** Informations sur l'exécution des scénarios et des flux IBM SPSS Modeler comme étapes des travaux d'exécution sous IBM SPSS Collaboration and Deployment Services Deployment Manager.
- **Guide du développeur IBM SPSS Modeler CLEF.** 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 IBM 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 capacités d'analyse via des algorithmes tiers.
- **Guide des performances et d'administration d'IBM SPSS Modeler Server.** Informations sur le mode de configuration et d'administration de IBM SPSS Modeler Server.
- **Guide d'utilisation de la console d'administration d'IBM SPSS Modeler.** Informations concernant l'installation et l'utilisation de l'interface utilisateur de la console permettant de surveiller et de configurer IBM SPSS Modeler Server. La console est implémentée en tant que plug-in à l'application Deployment Manager.
- **Guide CRISP-DM d'IBM SPSS Modeler.** Guide détaillé sur l'utilisation de la méthodologie CRISP-DM pour l'exploration de données avec SPSS Modeler
- **Guide d'utilisation d'IBM SPSS Modeler Batch.** Guide complet sur l'utilisation de IBM SPSS Modeler en mode de traitement par lots, avec des détails sur l'exécution en mode de traitement par lots et les arguments de ligne de commande. Ce guide est disponible au format PDF uniquement.

Documentation de SPSS Modeler Premium

La suite de documentation SPSS Modeler Premium (à l'exception des instructions d'installation) est la suivante.

- **IBM SPSS Modeler Entity Analytics - Guide d'utilisation.** Informations sur l'utilisation des analyses d'entités avec SPSS Modeler, notamment l'installation et la configuration du référentiel, les noeuds d'analyses d'entités et les tâches administratives.
- **IBM SPSS Modeler Social Network Analysis - Guide d'utilisation.** Guide sur l'exécution des analyses de réseaux sociaux avec SPSS Modeler, y compris les analyses de groupe et analyses de diffusion.
- **SPSS Modeler Text Analytics - Guide d'utilisation.** Informations sur l'utilisation des analyses de texte avec SPSS Modeler, notamment sur les noeuds Text Mining, l'espace de travail interactif, les modèles et d'autres ressources.

Exemples d'application

Tandis que les outils d'exploration de données de SPSS Modeler peuvent vous aider à résoudre une grande variété de problèmes métier et organisationnels, les exemples d'application fournissent des introductions brèves et ciblées aux méthodes et aux techniques de modélisation. Les jeux 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, voir la rubrique «Dossier Demos».

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 IBM SPSS Modeler 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.

Chapitre 2. Recommandations relatives à l'architecture et au matériel

Architecture de IBM SPSS Modeler

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 IBM 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 IBM SPSS Modeler Server (plutôt que de IBM SPSS Modeler en mode autonome) présentent de nombreux avantages, tout particulièrement lorsque vous traitez des jeux de données volumineux :

- IBM SPSS Modeler Server peut être exécuté, non seulement sous Windows, mais également sous UNIX, 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 d'exploration de données.
- IBM SPSS Modeler Server est optimisé pour assurer des performances élevées. Lorsqu'il est impossible de répercuter les opérations 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. Etant donné que les serveurs disposent généralement d'une grande quantité d'espace disque disponible, IBM SPSS Modeler Server peut effectuer des opérations de tri, de fusion et d'agrégation sur des jeux de données très volumineux.
- Grâce à l'architecture client-serveur, vous pouvez centraliser les processus d'exploration de données de votre entreprise. Cette fonction de centralisation vous permet de formaliser le rôle de l'exploration de données dans vos processus métier.
- Grâce à des outils d'administration comme le IBM SPSS Modeler Administration Console (fourni avec IBM SPSS Modeler Server) et IBM SPSS Collaboration and Deployment Services (vendu séparément), vous pouvez contrôler les processus d'exploration de données et vérifier que les ressources de calcul appropriées sont disponibles. IBM SPSS Collaboration and Deployment Services vous permet d'automatiser certaines tâches d'exploration de données, de gérer l'accès aux modèles de données et de 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 graphique "IBM SPSS Modeler Server Architecture".

- **IBM SPSS Modeler.** Le logiciel client est installé sur l'ordinateur de l'utilisateur final. Il fournit une interface utilisateur et affiche les résultats de l'exploration de données. La version client consiste en une installation complète de IBM SPSS Modeler, mais se distingue par le fait que lorsqu'elle est connectée à IBM SPSS Modeler Server pour une analyse répartie, son moteur d'exécution est inactif. IBM SPSS Modeler s'exécute sur les systèmes d'exploitation Windows uniquement.
- **IBM SPSS Modeler Server.** Le logiciel serveur est installé sur un ordinateur serveur, et comporte une connexion réseau avec les programmes IBM SPSS Modeler et la base de données. IBM SPSS Modeler Server s'exécute en tant que service (sous Windows) ou en tant que 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 IBM SPSS Modeler.

- **Serveur de bases 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 Server sous Windows, par exemple).

Architecture IBM SPSS Modeler Server

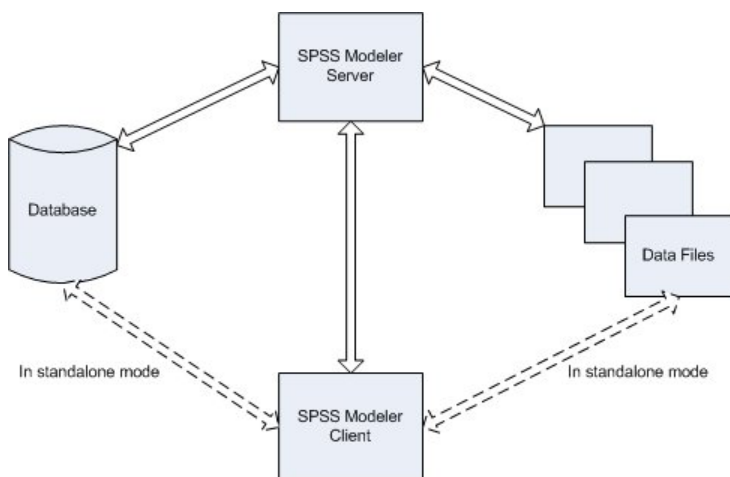


Figure 1. Architecture IBM SPSS Modeler Server

Dans le cas d'une architecture répartie, la plupart des processus s'effectuent sur l'ordinateur serveur. Lorsque l'utilisateur final exécute un flux, IBM 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, IBM SPSS Modeler Server peut exécuter toutes les opérations IBM 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, IBM SPSS Modeler Server est entièrement compatible avec les fichiers plats.

L'équilibrage de la charge est aussi disponible grâce à un cluster de serveurs pour le traitement. Le groupement est disponible à partir de IBM SPSS Collaboration and Deployment Services 3.5 via le plug-in du coordinateur de processus. Pour plus d'informations, voir la rubrique Annexe E, «Équilibrage de charge avec clusters de serveur», à la page 91. Vous pouvez vous connecter à un serveur ou à un cluster de serveurs gérés dans le coordinateur de processus directement grâce à la boîte de dialogue Connexion au serveur de IBM SPSS Modeler. Pour plus d'informations, voir la rubrique «Connexion à IBM SPSS Modeler Server», à la page 15.

Client autonome

Vous pouvez également configurer IBM SPSS Modeler tel qu'il s'exécute en tant qu'application de bureau indépendante, comme l'illustre le graphique ci-dessous. Consultez Chapitre 3, «Support IBM SPSS Modeler», à la page 15 pour plus d'informations.

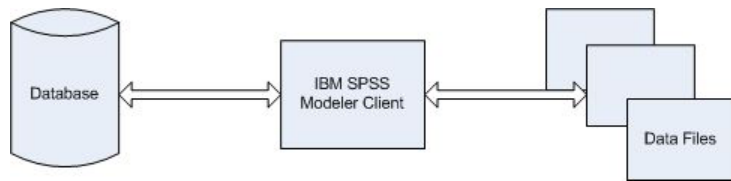


Figure 2. Application IBM SPSS Modeler autonome

Recommandations matérielles

Lors de la phase de préparation de votre installation IBM SPSS Modeler Server, pensez au matériel que vous utiliserez. IBM SPSS Modeler Server est certes conçu pour être rapide, mais pour profiter au maximum de son efficacité, utilisez du matériel adapté à vos tâches d'exploration de données. 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 IBM SPSS Modeler Server sur un serveur dédié, dont il ne disputera pas les ressources avec d'autres applications, notamment les bases de données IBM 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 IBM 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 IBM 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.
- Les capacités de multi-unité d'exécution permettent 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-unité d'exécution 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 une machine 64 bits comme plateforme IBM SPSS Modeler Server, et augmentez au maximum la quantité de mémoire RAM de l'ordinateur. En effet, dans le cas de jeux 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.

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 capacités d'unités d'exécution multiples améliorées de IBM 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 le traitement 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, qui calcule le paramètre de décalage.

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 le jeu 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, IBM 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 jeu 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 Réseau de neurones, Kohonen et k moyenne 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 `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 contient un ensemble de pilotes ODBC qui peuvent être utilisés dans ce but, et ces pilotes sont disponibles sur le DVD ou depuis le site de téléchargement. 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.

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 IBM SPSS Modeler 17.1, 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

Remarque : 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 IBM 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 IBM SPSS Modeler Server. Pour IBM SPSS Modeler Server sur des systèmes UNIX, consultez la rubrique « Configuration des pilotes ODBC sur les systèmes UNIX » plus loin dans cette section.
- Si vous devez accéder aux mêmes sources de données provenant de IBM SPSS Modeler et de IBM SPSS Modeler Server, les pilotes ODBC doivent être installés sur les deux ordinateurs.
- Si vous exécutez IBM SPSS Modeler sur Terminal Services, vous devez installer les pilotes ODBC sur le serveur Terminal Services sur lequel vous disposez de IBM SPSS Modeler.

Configuration des pilotes ODBC sur les systèmes UNIX

Par défaut, le gestionnaire de pilote DataDirect n'est pas configuré pour IBM SPSS Modeler Server sur les systèmes UNIX. Pour configurer le chargement du gestionnaire de pilote DataDirect sur UNIX, saisissez les commandes suivantes :

```
cd <modeler_server_install_directory>/bin
rm -f libspssodbc.so
ln -s libspssodbc_datadirect.so libspssodbc.so
```

Le lien par défaut est alors supprimé et un lien vers le gestionnaire de pilote DataDirect est créé.

Remarque : L'encapsuleur de pilote UTF16 est obligatoire pour utiliser les pilotes SAP HANA ou IBM DB2 CLI de certaines bases de données. Pour créer un lien vers l'encapsuleur de pilote UTF16, entrez plutôt la commande suivante :

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

Pour configurer SPSS Modeler Server :

1. Configurez le script de démarrage SPSS Modeler Server `modelersrv.sh` pour fournir le fichier d'environnement IBM SPSS Data Access Pack `odbc.sh` en ajoutant la ligne suivante à `modelersrv.sh` :
.
./<chemin_installation_SDAP>/odbc.sh

Où <chemin_installation_SDAP> est le chemin d'accès complet à votre installation IBM SPSS Data Access Pack.

2. Redémarrez SPSS Modeler Server.

En outre, pour SAP HANA et IBM DB2 uniquement, ajoutez la définition de paramètre suivante au nom de système par défaut dans le fichier `odbc.ini` afin d'éviter les dépassements de mémoire tampon durant la connexion :

```
DriverUnicodeType=1
```

Remarque : L'encapsuleur `libspssodbc_datadirect_utf16.so` est également compatible avec les autres pilotes ODBC pris en charge pour SPSS Modeler Server.

Remarque : Les règles précisées ci-dessus s'appliquent spécifiquement 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 et SPSS Statistics

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 Annexe B, «Configuration des scripts de démarrage UNIX», à la page 77.

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 IBM 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 IBM 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 IBM 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 nœuds IBM SPSS Statistics dans IBM SPSS Modeler peuvent démarrer automatiquement le client IBM SPSS Statistics, mais les utilisateurs doivent d'abord s'assurer que le client IBM SPSS Statistics présente le même mode d'exécution que IBM 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 DVD correspondant. Des documents différents sont disponibles pour Windows et UNIX.

Pour obtenir des informations complètes sur l'installation et l'utilisation du client IBM SPSS Modeler, voir le DVD du client.

Chapitre 3. Support 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 IBM SPSS Modeler Server
- Différences au niveau des résultats obtenus avec IBM SPSS Modeler Server et IBM SPSS Modeler

Connexion à IBM SPSS Modeler Server

Il est possible d'exécuter IBM SPSS Modeler comme une application autonome ou un comme un client connecté directement à IBM SPSS Modeler Server ou à IBM SPSS Modeler Server ou à un cluster de serveurs par le biais du Coordinateur des processus connecté à partir de IBM SPSS Collaboration and Deployment Services. Le statut de la connexion apparaît en bas à gauche de la fenêtre IBM 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 clusters de serveurs à partir de la boîte de dialogue Connexion au serveur. Vous pouvez naviguer via les services Statistiques s'exécutant sur un réseau grâce au Coordinateur des processus.

Pour vous connecter à un serveur

1. Dans le menu Outils, cliquez sur **Connexion au serveur**. La boîte de dialogue Connexion au serveur s'affiche. Vous pouvez également cliquer deux fois sur la zone d'état de la connexion dans la fenêtre IBM SPSS Modeler.
2. 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, voir la rubrique «Ajout et modification d'une connexion à IBM SPSS Modeler Server», à la page 21.
 - Cliquez sur **Rechercher** pour accéder au serveur ou à un cluster de serveurs dans le Coordinateur de processus. Pour plus d'informations, voir la rubrique «Recherche de serveurs dans IBM SPSS Collaboration and Deployment Services», à la page 22.

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 données 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.

ID 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.

3. Cliquez sur **OK** pour terminer la connexion.

Pour se déconnecter d'un serveur

1. Dans le menu Outils, cliquez sur **Connexion au serveur**. La boîte de dialogue Connexion au serveur s'affiche. Vous pouvez également cliquer deux fois sur la zone d'état de la connexion dans la fenêtre IBM SPSS Modeler.
2. Dans la boîte de dialogue, sélectionnez le serveur local, puis cliquez sur **OK**.

Configuration de la connexion unique

Vous pouvez vous connecter à un serveur IBM SPSS Modeler Server qui s'exécute sur une plateforme prise en charge à l'aide de la connexion unique (SSO). Pour vous connecter à l'aide de la connexion unique, configurez au préalable le serveur et le client IBM SPSS Modeler.

Si vous utilisez une connexion unique pour vous connecter à IBM SPSS Modeler Server et à IBM SPSS Collaboration and Deployment Services, vous devez vous connecter à IBM SPSS Collaboration and Deployment Services avant de vous connecter à IBM SPSS Modeler.

IBM SPSS Modeler Server utilise Kerberos pour la connexion unique.

Kerberos est un composant de base de Windows Active Directory et les informations suivantes supposent une infrastructure Active Directory. Notamment :

- L'ordinateur client est un ordinateur Windows connecté à un domaine Active Directory
- L'utilisateur client s'est connecté à l'ordinateur à l'aide d'un compte de domaine. Le mécanisme utilisé pour la connexion n'est pas important ; il peut faire intervenir une carte à puce, les empreintes digitales, etc.
- IBM SPSS Modeler Server peut valider les données d'identification de l'utilisateur client par référence au contrôleur de domaine Active Directory

Cette documentation explique comment configurer les serveurs Windows et UNIX pour effectuer une authentification de cette manière. D'autres configurations peuvent être possibles, mais n'ont pas été testées.

Pour pouvoir interagir avec les installations Active Directory les plus récentes et les plus sécurisées, vous devez installer le pack de chiffrement haute résistance pour Java, car les algorithmes de chiffrement requis ne sont pas pris en charge par défaut. Vous devez installer le pack à la fois pour le client et pour le serveur. Un message d'erreur tel que `Illegal key size` s'affiche sur le client en cas d'échec de la connexion serveur due à l'absence d'installation du pack. Voir «Installation du chiffrement de type Unlimited Strength», à la page 51.

Nom principal du service

Chaque instance de serveur doit enregistrer un *nom principal de service (SPN)* unique pour s'identifier, et le client doit spécifier le même SPN lorsqu'il se connecte au serveur.

Un SPN pour une instance de SPSS Modeler Server a le format suivant :

```
modelerserver/<hôte>:<port>
```

Par exemple :

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

Notez que le nom d'hôte doit être qualifié avec son domaine DNS (spss.com dans cet exemple), et le domaine doit correspondre au domaine Kerberos.

L'association du nom d'hôte et du numéro de port rend le SPN unique (car chaque instance sur un hôte donné doit utiliser un port d'écoute différent). Le client et le serveur disposent déjà du nom d'hôte et du numéro de port, ils peuvent donc construire le SPN approprié pour l'instance. L'étape de configuration supplémentaire requise consiste à enregistrer le SPN dans la base de données Kerberos.

Enregistrement du SPN sous Windows

Si vous utilisez Active Directory comme implémentation Kerberos, utilisez la commande `setspn` pour enregistrer le SPN. Pour exécuter cette commande, les conditions ci-dessous doivent être satisfaites.

- Vous devez être connecté à un contrôleur de domaine.
- Vous devez exécuter l'invite de commande avec des privilèges élevés (en tant qu'administrateur).
- Vous devez être membre du groupe Domain Admins (ou avoir reçu la permission appropriée d'un administrateur de domaine).

Pour plus d'informations, voir les articles suivants :

- [Setspn Command-Line Reference](#)
- [Delegating Authority to Modify SPNs](#)

Pour l'instance par défaut, en écoute sur le port standard (28054 pour IBM SPSS Modeler version 17, par exemple) et exécutée sur le compte Local System, vous devez enregistrer le SPN par rapport au nom du serveur. Par exemple :

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

Pour chaque instance (profil) suivante, en écoute sur le port personnalisé (par exemple, 29000) et s'exécutant sous un compte utilisateur arbitraire (par exemple, jdoe) avec l'option `start_process_as_login_user` définie sur Y, vous devez enregistrer le SPN en fonction du nom de compte utilisateur du service :

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

Notez que dans ce cas (lorsque le compte de service n'est pas Local System), l'enregistrement du SPN ne suffit pas pour permettre au client de se connecter. D'autres étapes de configuration sont décrites dans la section suivante.

Pour voir les SPN enregistrés sur le compte jdoe :

```
setspn -l jdoe
```

Enregistrement du SPN sous UNIX

Si vous utilisez Active Directory comme implémentation Kerberos, vous pouvez utiliser la commande `setspn` comme indiqué dans la section Windows précédente ; cela suppose que vous avez déjà créé le compte utilisateur ou ordinateur dans le répertoire. Vous pouvez aussi utiliser `ktpass`, comme expliqué dans «Configuration de IBM SPSS Modeler Server sous UNIX et Linux», à la page 18.

Si vous utilisez une autre implémentation Kerberos, utilisez l'outil d'administration Kerberos pour ajouter le principal du service à la base de données Kerberos. Pour convertir le SPN à un principal Kerberos, vous devez ajouter le nom du domaine Kerberos. Par exemple :

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

Ajoutez ce principal et le mot de passe au fichier `keytab` du serveur. Le fichier `keytab` doit contenir une entrée pour chaque instance qui s'exécute sur l'hôte.

Configuration d'IBM SPSS Modeler Server sous Windows

Dans le scénario par défaut où le service SPSS Modeler Server s'exécute sous le compte Système local, il utilise les API Windows natives pour authentifier les données d'identification de l'utilisateur et aucune configuration supplémentaire n'est requise sur le serveur.

Dans l'autre scénario où le service SPSS Modeler Server s'exécute sous un compte utilisateur dédié et où `start_process_as_login_user` a pour valeur Y, il utilise les API Java pour authentifier les données d'identification de l'utilisateur et une configuration supplémentaire est requise sur le serveur.

Assurez-vous dans un premier temps que le scénario par défaut fonctionne. Le client doit pouvoir utiliser la connexion unique SSO pour se connecter à l'instance par défaut qui s'exécute sous le compte Local System. Cela valide la configuration côté client (qui est inchangée). Vous devrez enregistrer le SPN pour l'instance par défaut comme décrit précédemment.

Exécutez ensuite la procédure suivante :

1. Créez le répertoire `<MODELERSERVER>\config\sso`.
2. Créez un fichier appelé `krb5.conf` dans le dossier `sso` créé à l'étape 1. Pour plus d'instructions sur la façon de créer ce fichier, voir l'étape 3 sous «Configuration du client IBM SPSS Modeler», à la page 19. Le fichier doit être le même sur le serveur et sur le client.

3. Utilisez la commande suivante pour créer le fichier `krb5.keytab` dans le répertoire SSO serveur :

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

Par exemple :

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

Vous êtes invité à entrer un mot de passe. Il doit s'agir du mot de passe du compte de service. Ainsi, si le compte de service est `jdoe`, par exemple, vous devez entrer le mot de passe de l'utilisateur `jdoe`.

Le compte de service lui-même n'est pas mentionné dans le fichier `keytab`, mais vous avez précédemment enregistré le SPN dans ce compte à l'aide de `setspn`. Cela signifie que le mot de passe du principal du service et celui du compte du service sont identiques.

Pour chaque nouvelle instance (profil) que vous créez, vous devez enregistrer le SPN de cette instance (à l'aide de `setspn`) (voir «Configuration de profils de serveur», à la page 29 et «**Nom principal du service**», à la page 16) et ajouter une entrée au fichier `keytab` (à l'aide de `jre\bin\ktab`). Il n'existe qu'un seul fichier `keytab`, qui doit contenir une entrée pour chaque instance qui ne s'exécute pas sous Local System. L'instance par défaut, ou toute autre instance s'exécutant sous Local System, ne doit pas nécessairement se trouver dans le fichier `keytab` car elle utilise les API Windows pour l'authentification. Les API Windows n'utilisent pas le fichier `keytab`.

Pour vérifier si une instance est incluse dans le fichier `keytab` :

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

Vous pouvez voir plusieurs entrées pour chaque principal avec différents types de chiffrement, ce qui est normal.

Configuration de IBM SPSS Modeler Server sous UNIX et Linux Éléments prérequis

IBM SPSS Modeler Server s'appuie sur Windows Active Directory (AD) pour activer la connexion unique (SSO) pour laquelle les prérequis suivants sont essentiels :

- L'ordinateur SPSS Modeler Client (Windows) est membre d'un domaine Active Directory (AD).
- L'utilisateur client se connecte à l'ordinateur à l'aide d'un compte de domaine Active Directory.

- L'ordinateur SPSS Modeler Server (UNIX) est identifié par un nom de domaine qualifié complet dont la racine est dans le domaine DNS Active Directory. Par exemple, si le domaine DNS est modelersso.com, le nom d'hôte du serveur peut être myserver.modelersso.com.
- Le domaine DNS Active Directory prend en charge les recherches en aval et en amont de nom d'hôte SPSS Modeler Server.

Si la machine SPSS Modeler Server n'est pas membre du domaine Active Directory, vous devez créer un compte d'utilisateur de domaine pour représenter le service dans l'annuaire. Par exemple, vous pouvez créer un compte de domaine appelé ModelerServer.

Pour configurer SPSS Modeler Server sous UNIX ou Linux

1. Dans le dossier SPSS Modeler Serverconfig, créez un sous-dossier sso.
2. Dans le dossier sso, créez un fichier keytab. La génération du fichier keytab peut être effectuée côté Active Directory. Toutefois, il existe des exigences différentes selon que la machine SPSS Modeler Server est membre du domaine Active Directory ou non :

- Si la machine SPSS Modeler Server **est** membre du domaine Active Directory, utilisez le nom de compte de l'ordinateur comme nom d'utilisateur du service :

```
ktpass -princ <spn>@<realm> -mapUser <domaine>\<compte_ordinateur> -pass <mot_de_passe> -out <fichier_sortie> -ptype KRB5_NT_PRINCIPAL
```

Par exemple :

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

- Si la machine SPSS Modeler Server **n'est pas** membre du domaine Active Directory, indiquez le compte d'utilisateur de domaine, que vous avez créé comme prérequis, comme utilisateur du service :

```
ktpass -princ <spn>@<realm> -mapUser <domaine>\<compte_util> -mapOp set -pass <mot_de_passe> -out <fichier_sortie> -ptype KRB5_NT_PRINCIPAL
```

Par exemple :

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

Pour plus d'informations, voir Ktpass Command-Line Reference.

3. Renommez le fichier keytab dans le dossier sso en krb5.keytab.

Remarque : Si vous ré-associez la machine serveur au domaine, générez un nouveau fichier keytab.

4. Créez un fichier appelé krb5.conf dans le dossier sso créé à l'étape 1. Pour plus d'instructions sur la façon de créer ce fichier, voir l'étape 3 sous «Configuration du client IBM SPSS Modeler». Le fichier doit être le même sur le serveur et sur le client.

Configuration du client IBM SPSS Modeler

1. Configurez Java de manière à autoriser l'accès à la clé de session TGT :
 - a. Dans le menu **Démarrer**, cliquez sur l'option **Exécuter**.
 - b. Entrez regedit et cliquez sur **OK** pour ouvrir l'éditeur de registre.
 - c. Accédez à l'emplacement de registre correspondant au système d'exploitation utilisé par la machine locale :
 - Sous Windows XP : My Computer\HKEY_LOCAL_MACHINE\System\CurrentControlSet\Control\Lsa\Kerberos
 - Sous Windows Vista ou Windows 7 : My Computer\HKEY_LOCAL_MACHINE\System\CurrentControlSet\Control\Lsa\Kerberos\Parameters
 - d. Cliquez avec le bouton droit de la souris sur le dossier et sélectionnez **Nouveau** > **DWORD**. Le nom de la nouvelle valeur doit être allowtgtsessionkey.

- e. Définissez la valeur de `allowtgtsessionkey` à la valeur hexadécimale de 1, c'est-à-dire 0x0000001.
 - f. Fermez l'éditeur de registre.
 - g. **Remarque** : Il existe un problème connu qui se produit lorsque le compte utilisateur est membre du groupes Administrateurs local et que le contrôle de compte utilisateur (UAC) est activé. La clé de session du ticket de service extrait est alors vide et l'authentification SSO (connexion unique) échoue. Pour éviter ce problème, exécutez l'une des actions suivantes :
 - Exécutez l'application en tant qu'administrateur
 - v
 - Utilisez un compte qui n'est pas un compte Administrateur
2. Dans le dossier `config` de l'emplacement d'installation d'IBM SPSS Modeler, créez un dossier appelé `sso`.
 3. Dans le dossier `sso`, créez un fichier `krb5.conf`. Les instructions à suivre pour créer un fichier `krb5.conf` sont disponibles sur le site http://web.mit.edu/kerberos/krb5-current/doc/admin/conf_files/krb5_conf.html. Voici un exemple de fichier `krb5.conf` :


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

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

[domain_realm]
    .modelersso.com = MODELERSSO.COM
    modelersso.com = MODELERSSO.COM
```
 4. Redémarrez la machine locale et le serveur.

Obtention de l'appartenance au groupe de l'utilisateur SSO

Lorsqu'un utilisateur se connecte à SPSS Modeler Server via une connexion unique SSO et que le serveur s'exécute en tant que processus non racine, le nom de l'utilisateur authentifié n'est pas associé à un compte utilisateur du système d'exploitation. Le serveur ne peut pas se procurer l'appartenance au groupe du système d'exploitation de l'utilisateur. Par conséquent, comment s'effectue la configuration de groupe dans ce cas ?

Considérons que l'utilisateur est enregistré dans un annuaire LDAP (qui peut être Active Directory) et qu'une demande peut porter sur l'appartenance au groupe à partir du serveur LDAP. SPSS Modeler Server peut demander l'appartenance au groupe au fournisseur LDAP dans IBM SPSS Collaboration and Deployment Services.

Il existe deux propriétés dans `options.cfg` sur SPSS Modeler Server qui contrôlent l'accès du serveur au référentiel IBM SPSS Collaboration and Deployment Services :

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

Pour activer la recherche de groupe, vous devez définir ces deux propriétés. Par exemple :

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

La connexion à un référentiel est utilisée uniquement pour la recherche de groupe SSO, vous n'avez donc pas besoin de changer les paramètres de ces propriétés sauf si vous avez besoin de cette fonctionnalité.

Pour que la recherche de groupe fonctionne correctement, vous devez d'abord configurer votre référentiel pour ajouter un fournisseur LDAP ou Active Directory, puis pour activer la connexion unique SSO à l'aide de ce fournisseur :

1. Démarrez le client IBM SPSS Collaboration and Deployment Services Deployment Manager et sélectionnez **Fichier > Nouveau > Connexion au serveur administré** pour créer une connexion au serveur administré pour votre référentiel (si vous n'en avez pas déjà une).
2. Connectez-vous via la connexion au serveur administré et développez le dossier **Configuration**.
3. Cliquez avec le bouton droit de la souris sur **Fournisseurs de sécurité**, sélectionnez **Nouveau > Définition du fournisseur de sécurité**, et entrez les valeurs appropriées. Pour plus d'informations, cliquez sur **Aide** dans la boîte de dialogue.
4. Développez le dossier **Fournisseurs de connexion unique**, cliquez avec le bouton droit de la souris sur **Fournisseur SSO Kerberos**, et sélectionnez **Ouvrir**.
5. Cliquez sur **Activer**, sélectionnez votre fournisseur de sécurité, puis cliquez sur **Enregistrer**. Il n'est pas nécessaire d'indiquer d'autres détails ici, sauf si vous souhaitez utiliser SSO (il suffit que le fournisseur soit activé pour permettre la recherche du groupe).

Important : Pour que la recherche de groupe fonctionne correctement, le fournisseur Kerberos que vous configurez ici doit être le même que le fournisseur configuré pour SPSS Modeler Server. En particulier, ils doivent fonctionner dans le même domaine Kerberos. Ainsi, si un utilisateur se connecte à SPSS Modeler Server via une connexion SSO et qu'il est identifié en tant que `jdoe@SPSS.COM` (où `SPSS.COM` est le domaine), le fournisseur de sécurité de IBM SPSS Collaboration and Deployment Services reconnaîtra le nom principal de cet utilisateur et renverra l'appartenance au groupe correspondante à partir de l'annuaire LDAP.

Connexion unique pour sources de données

Vous pouvez vous connecter à des bases de données à partir de IBM SPSS Modeler via une connexion unique. Si vous souhaitez créer une connexion de base de données via une connexion unique, vous devez d'abord utiliser votre logiciel de gestion ODBC pour configurer correctement une source de données et un jeton de connexion unique. Ensuite, lorsque vous vous connecterez à une base de données dans IBM SPSS Modeler, IBM SPSS Modeler utilisera ce jeton de connexion unique et l'utilisateur ne sera pas invité à se connecter à la source de données.

Cependant, si la source de données n'a pas été configurée correctement pour une connexion unique, IBM SPSS Modeler invitera l'utilisateur à se connecter à la source de données. L'utilisateur pourra toujours accéder à la source de données après avoir fourni des données d'identification valides.

Pour les détails complets de la configuration des sources de données ODBC sur votre système avec une connexion unique, consultez la documentation du fournisseur de base de données. Vous trouverez ci-dessous un exemple de la procédure générale à suivre.

1. Configurez votre base de données pour qu'elle prenne en charge la connexion unique Kerberos.
2. Sur la machine IBM SPSS Modeler Server, créez une source de données ODBC et testez-la. La connexion DSN ne doit nécessiter ni ID utilisateur, ni mot de passe.
3. Connectez-vous à IBM SPSS Modeler Server via une connexion unique et commencez à utiliser la source de données ODBC créée et validée à l'étape 2.

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. Les pratiques recommandées indiquent d'utiliser les mêmes ports pour communiquer avec IBM SPSS Collaboration and Deployment Services et avec SPSS Modeler Client. Ces ports peuvent être définis via les paramètres `max_server_port` et `min_server_port` dans le fichier `options.cfg`.

Pour ajouter des connexions au serveur

1. Dans le menu Outils, cliquez sur **Connexion au serveur**. La boîte de dialogue Connexion au serveur s'affiche.
2. Dans la boîte de dialogue, cliquez sur **Ajouter**. La boîte de dialogue Ajouter/Modifier un serveur pour la connexion au serveur s'ouvre.
3. 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.
 - **Chiffrer la connexion (utiliser SSL).** Indique 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

1. Dans le menu Outils, cliquez sur **Connexion au serveur**. La boîte de dialogue Connexion au serveur s'affiche.
2. 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.
3. 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 cluster de serveurs disponible sur le réseau par le biais du Coordinateur de processus, disponible dans IBM SPSS Collaboration and Deployment Services. Un cluster de serveurs contient plusieurs serveurs, et permet au Coordinateur de processus de déterminer le serveur qui répond le mieux à la demande de traitement.

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, voir la rubrique «Ajout et modification d'une connexion à IBM SPSS Modeler Server», à la page 21.

Pour rechercher des serveurs et des clusters de serveurs

1. Dans le menu Outils, cliquez sur **Connexion au serveur**. La boîte de dialogue Connexion au serveur s'affiche.
2. 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.
3. Sélectionnez le serveur ou le cluster de serveurs dans la liste.
4. Cliquez sur **OK** pour fermer la boîte de dialogue et ajouter cette connexion au tableau de la boîte de dialogue Connexion au serveur.

Systèmes 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 IBM 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 IBM 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 IBM SPSS Modeler Server en utilisant l'authentification par défaut sous Windows.

UNIX. Par défaut, IBM 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 IBM 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, voir rubrique «Configuration en tant qu'utilisateur non-racine avec une base de données de mots de passe privés», à la page 86.

Configuration du module PAM

Sur les plateformes Solaris, Linux et AIX, IBM SPSS Modeler Server utilise PAM pour l'authentification.

Pour utiliser l'authentification PAM, les modules PAM appropriés doivent être configurés correctement sur le système hôte. Par exemple, pour que PAM fonctionne avec LDAP, un module LDAP PAM doit

exister sur le système d'exploitation de l'hôte et il doit être correctement configuré. Pour plus d'informations, consultez la documentation du système d'exploitation. Cette condition est requise pour que SPSS Modeler Server puisse utiliser un module PAM.

Pour configurer SPSS Modeler Server afin d'utiliser un module, éditez le fichier "options.cfg" de SPSS Modeler Server et ajoutez (ou éditez) la ligne **authentication_methods, pam**.

Vous pouvez utiliser le nom de service modelerserver pour configurer les modules PAM pour SPSS Modeler Server, si nécessaire. Par exemple, les étapes suivantes illustrent la configuration requise pour Red Hat Linux :

1. Accédez au répertoire de configuration PAM. Par exemple : **/etc/pam.d**.
2. A l'aide d'un éditeur de texte, créez un nouveau fichier nommé "modelerserver".
3. Ajoutez les informations de configuration PAM que vous souhaitez utiliser. Par exemple :

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

Remarque : Ces lignes peuvent varier en fonction de la configuration. Pour plus d'informations, consultez la documentation Linux.

4. Enregistrez le fichier et redémarrez le service Modeler.

Droits

Windows. Un utilisateur qui se connecte au logiciel serveur installé sur une unité *NTFS* doit se connecter avec un compte doté des droits suivants :

- Droits de lecture et d'exécution sur le répertoire d'installation du serveur et sur ses sous-répertoires.
- Droits de lecture, d'exécution et d'écriture sur l'emplacement du répertoire de fichiers temporaires.

Sous Windows Server 2008 et versions suivantes, vous ne pouvez pas considérer que les utilisateurs disposent de ces droits. Veillez à définir ces droits explicitement selon les besoins.

Si le logiciel serveur est installé sur une unité *FAT*, vous n'avez pas besoin de définir ces droits car tous les fichiers permettent aux utilisateurs de disposer d'un contrôle complet.

UNIX. Si vous n'utilisez pas une authentification interne, un utilisateur qui se connecte au logiciel serveur doit se connecter avec un compte doté des droits suivants :

- Droits de lecture et d'exécution sur le répertoire d'installation du serveur et sur ses sous-répertoires.
- Droits de lecture, d'exécution et d'écriture sur l'emplacement du répertoire de fichiers temporaires.

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, voir la rubrique «Exigences en termes d'espace disque temporaire et de mémoire RAM», à la page 10.

UNIX. Les versions UNIX de IBM 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, voir la rubrique «Contrôle des droits d'accès relatifs à la création de fichier», à la page 78.

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 IBM 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 IBM 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.

Chapitre 4. Administration de IBM SPSS Modeler

Ce chapitre contient des informations sur le démarrage et l'arrêt de IBM SPSS Modeler Server, sur la configuration de diverses options du serveur et la configuration de groupes, ainsi que sur l'interprétation du fichier journal. Il décrit l'utilisation de l'application IBM SPSS Modeler Administration Console, qui facilite la configuration et la surveillance du serveur. Pour obtenir des instructions d'installation concernant ce composant, reportez-vous à celles de IBM SPSS Modeler Server, fournies avec ce 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 IBM 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 statut

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

1. **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**.
2. Sélectionnez le service **IBM SPSS Modeler Server <nn.n>**. Vous pouvez maintenant vérifier le statut 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 statut

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

1. Passez dans le répertoire d'installation d'IBM SPSS Modeler Server. Par exemple, à l'invite de commande UNIX, entrez

```
cd /usr/modelersrv
```

où *modelersrv* est le répertoire d'installation de IBM SPSS Modeler Server.
2. A l'invite de commande, pour démarrer le serveur, entrez la ligne suivante :

```
./modelersrv.sh start
```
3. Pour arrêter le serveur, à l'invite de commande, entrez

```
./modelersrv.sh stop
```
4. Pour vérifier le statut de IBM 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 IBM SPSS Modeler Server, les autres sont les sessions IBM SPSS Modeler.

Le programme d'installation de IBM 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, voir la rubrique «Démarrage et arrêt automatiques de IBM SPSS Modeler Server», à la page 77.

Limites du noyau UNIX

Vous devez vous assurer que les limites du noyau sur le système sont suffisantes pour utiliser IBM 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 IBM SPSS Modeler Server. Pour ce faire :

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

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

En outre, définissez la limite de pile sur la valeur maximum admise par votre système (`ulimit -s XXXX`), par exemple :

```
ulimit -s 64000
```

2. Redémarrez IBM 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 IBM 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 ("zombie"), nous recommandons d'arrêter IBM SPSS Modeler Server et de le redémarrer à intervalles réguliers, à l'aide de la séquence de commandes suivantes :

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

Ces processus IBM 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 IBM SPSS Modeler, comme partie intégrale du processus de redémarrage de IBM SPSS Modeler Server à intervalles réguliers.

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

Configuration de profils de serveur

Les profils de serveur vous permettent d'exécuter plusieurs instances indépendantes de SPSS Modeler Server à partir d'une seule installation. Pour un client, ils apparaissent comme des serveurs distincts localisés sur le même hôte mais utilisant différents numéros de ports d'écoute. Le fait d'avoir plusieurs instances partageant une même installation est un atout pour les administrateurs car la maintenance s'en trouve simplifiée. Les instances qui suivent la première peuvent être créées et supprimées plus rapidement que dans le cas d'une installation et d'une désinstallation complètes, et les groupes de correctifs ne doivent être appliqués qu'une seule fois.

L'exécution de plusieurs instances de serveur sur le même hôte permet de configurer chaque instance séparément. Si toutes les instances sont identiques, cette situation n'apporte aucun avantage. En particulier, si les instances s'exécutent en mode *non racine* (toutes les sessions partagent alors le même compte utilisateur), chaque instance peut utiliser un compte utilisateur différent pour assurer l'isolement des données entre les groupes d'utilisateurs. Par exemple, un utilisateur qui se connecte à une instance A recevra une session appartenant à un utilisateur User-A particulier et il ne pourra accéder qu'aux fichiers et dossiers de cet utilisateur. Par contre, un utilisateur se connectant à une instance B verra des fichiers et des dossiers différents accessibles à l'utilisateur User-B. Cette situation peut être utilisée avec une configuration de groupe de sorte que la connexion à une instance en particulier est restreinte à des groupes spécifiques, ce qui signifie que les utilisateurs finals ne peuvent se connecter qu'à l'instance (ou les instances) appropriée à leur rôle. Voir «Configuration des groupes», à la page 53.

Dans le cadre d'une installation SPSS Modeler Server standard, les dossiers config, data et tmp sont spécifiques à une instance de serveur. Le but du dossier config est que l'instance dispose d'une configuration privée, et que les dossiers data et tmp prennent en charge l'isolement des données. Chaque instance possède une copie privée de ces dossiers, et tout le reste est partagé.

Notez qu'une grande partie de la configuration du serveur peut rester commune (les paramètres de base de données, par exemple), une configuration de profil remplacera donc la configuration commune. Le serveur consultera d'abord la configuration du profil, puis la configuration par défaut. Les fichiers les plus susceptibles d'être modifiés pour un profil sont options, groups et passwords.

Pour plus d'informations, voir «Structure de profil», à la page 31.

Pour plus d'informations sur la configuration d'un profil pour qu'il utilise une connexion unique, voir «Configuration de la connexion unique», à la page 16. Pour cela, vous devez enregistrer un nom principal de service (SPN), effectuer certaines tâches de configuration si le compte de service Windows n'est pas compte local et, dans certains cas, activer la recherche de groupe.

Utilisation des profils de serveur

Vous trouverez ci-après des exemples d'utilisation courantes des profils de serveur. Certains sont pris en charge via l'utilisation de scripts (voir «Scripts de profil», à la page 33) et peuvent nécessiter des privilèges administrateur/root.

Création d'un profil de serveur

Un administrateur SPSS Modeler Server appelé Jane utilise un script pour créer un nouveau profil :

- Jane doit spécifier un nom unique pour le profil (qui ne doit pas être un nom de profil existant). Si le répertoire profiles n'existe pas déjà, il est créé pour Jane. Ensuite, un nouveau sous-répertoire est créé dans le répertoire profiles sous le nom spécifié par Jane, contenant les répertoires config, data, log et tmp.
- Si Jane le choisit, elle peut également spécifier le nom d'un profil existant à utiliser comme modèle. Dans ce cas, le contenu du dossier config au sein du profil existant est copié dans le nouveau profil. Si Jane n'indique aucun modèle, ou si le profil existant n'inclut pas de fichier d'options, même s'il le devrait, un fichier d'options vide est créé dans le nouveau profil.

- Jane peut également choisir de spécifier un numéro de port pour le profil. Dans ce cas, ce numéro de port est écrit comme la valeur de la propriété `numero_port` dans le fichier d'options du profil. Si elle ne spécifie pas de numéro de port, une valeur est choisie pour elle et écrite dans le fichier d'options.
- Jane peut également choisir de spécifier le nom d'un groupe de systèmes d'exploitation qui aura un accès exclusif au profil. Dans ce cas, la configuration de groupe est activée dans le fichier d'options. Un fichier de groupes est créé, qui refuse la connexion à tous les groupes sauf le groupe spécifié.

Configuration d'un profil de serveur

Jane, l'administrateur de serveur, configure un profil soit manuellement en éditant les fichiers configuration du profil, soit en utilisant IBM SPSS Modeler Administration Console pour se connecter au service de profil.

Création d'un service Windows pour un profil de serveur

Sous Windows, l'administrateur utilise un script pour créer un service pour un profil spécifié :

- Jane doit spécifier le nom d'un profil existant, puis une instance de service est créée pour ce profil. La ligne de commande du service comportera l'argument `profile`. Le nom du service suivra un modèle standard notamment le nom du profil.
- Jane devra peut-être utiliser la console d'administration de service ultérieurement et éditer les propriétés de service si elle doit changer le nom d'utilisateur et le mot de passe pour le service (lors d'une exécution non root).

Sous UNIX, il est possible de créer des "services" qui démarrent automatiquement lors de l'amorçage du système. L'administrateur peut créer des services de profil à l'aide de ces mécanismes, mais notez que ces derniers ne sont pas officiellement pris en charge par IBM SPSS Modeler.

Gestion des services Windows pour les profils de serveur

Les administrateurs peuvent utiliser un script pour effectuer les tâches suivantes :

- Voir les services de profil de serveur qui sont en cours d'exécution.
- Démarrer un service en particulier.
- Démarrer tous les services.
- Arrêter un service en particulier.
- Arrêter tous les services.

Lorsque vous démarrez ou arrêtez tous les services, la liste des profils est obtenue en recherchant les sous-répertoires du répertoire `profiles`.

Suppression du service Windows d'un profil de serveur

Sous Windows, les administrateurs peuvent utiliser un script pour supprimer un service pour un profil spécifié (si un service existe pour le profil). Le nom du profil doit être spécifié.

Suppression d'un profil de serveur

Après avoir arrêté le service du profil, les administrateurs peuvent supprimer un profil en supprimant son dossier du répertoire `profiles`.

Mise à jour de SPSS Modeler Server

Lorsque vous appliquez un groupe de correctifs à SPSS Modeler Server, il est appliqué à tous les profils de serveur. Sous Windows, tous les services de profil sont arrêtés et redémarrés automatiquement. Sous UNIX, vous devez les arrêter et les redémarrer manuellement.

Désinstallation de SPSS Modeler Server

Lorsque SPSS Modeler Server est désinstallé, tous les profils de serveur sont désinstallés. Notez que le répertoire `profiles` et tous les profils qu'il contient ne sont pas supprimés automatiquement. Ils doivent être supprimés manuellement. Sous Windows, tous les services de profil sont désinstallés automatiquement. Sous UNIX, vous devez les supprimer manuellement.

Installation d'une nouvelle version de SPSS Modeler Server

Lorsque vous installez une nouvelle version de SPSS Modeler Server, les profils de serveur existants ne sont pas migrés automatiquement. Un administrateur doit copier les profils manuellement d'une installation à l'autre (et éditer les configurations, si besoin) pour recréer les services.

Structure de profil

Répertoire de profils

Les profils de serveur sont stockés dans un emplacement de répertoire sélectionné par l'administrateur de serveur. L'emplacement par défaut est un répertoire nommé `profiles` dans le répertoire `[chemin_installation_serveur]\config\` sur le serveur SPSS Modeler Server, mais nous vous recommandons d'utiliser un répertoire différent pour le stockage des profils pour les raisons suivantes :

- Les profils peuvent être partagés entre les noeuds dans un cluster.
- Les profils peuvent être préservés au fur et à mesure des mises à niveau.
- Les administrateurs et les utilisateurs qui configurent des profils n'ont pas besoin de posséder les droits d'écriture dans le répertoire d'installation SPSS Modeler Server.

Le répertoire `profiles` n'existe pas après une toute nouvelle installation de SPSS Modeler Server. Il est créé lorsque le premier profil est créé.

Le répertoire `profiles` contient un sous-répertoire pour chaque profil, et le nom du sous-répertoire correspond à celui du profil. Etant donné que le répertoire et le profil portent le même nom, le nom du profil ne doit pas inclure des caractères non valides dans les noms de fichier. De même, les noms de profil ne doivent pas contenir d'espaces car ils sont susceptibles d'entraîner des problèmes dans les scripts. N'oubliez pas que les noms de profil doivent être uniques au sein d'une installation.

La seule façon d'identifier tous les profils pour une installation est d'identifier les sous-répertoires du répertoire `profiles`. Aucune liste séparée des profils n'existe. Le nombre des profils pouvant être créés pour une installation est illimité, dans les limites de la capacité du système hôte.

Dans le répertoire `profiles`, le sous-répertoire d'un profil donné doit contenir au moins un répertoire appelé `config`, lequel doit contenir au moins un fichier appelé `options.cfg` qui définit la configuration de profil. Ce fichier contient un sous-ensemble de paramètres dans le fichier `options.cfg` standard de SPSS Modeler Server (situé dans `[chemin d'installation du serveur]/config`), car de nombreux paramètres sont requis pour le profil. Les paramètres qui ne sont pas présents dans la configuration de profil doivent être définis à partir du fichier d'options commun contenu dans le répertoire d'installation `config`. La configuration de profil doit contenir au moins un paramètre pour `numero_port` car chaque service de profil doit utiliser un numéro de port d'écoute différent.

La configuration de profil peut inclure d'autres fichiers `*.cfg` que l'on trouve habituellement dans le répertoire d'installation `config`. Ils sont dans ce cas lus à la place des fichiers standard (seul le fichier d'options est cumulatif). Les fichiers supplémentaires `groups` et `passwords` sont susceptibles d'être ajoutés à la configuration de profil. Les fichiers ignorés dans une configuration de profil sont les fichiers de configuration JVM et SSO qui sont partagés entre tous les profils.

Un répertoire de profil peut également contenir les répertoires data et tmp qui remplacent les emplacements de fichiers data et tmp communs, sauf si d'autres emplacements sont spécifiés dans la configuration de profil.

Si vous utilisez des profils pour isoler des données, assurez-vous que les droits sont bien définis dans les répertoires concernés.

Fichier de configuration de profil

L'emplacement du répertoire profiles est spécifié dans un nouveau fichier de configuration appelé [chemin d'installation du serveur]\config\profiles.cfg. Il partage un format commun avec d'autres fichiers de configuration dans le même répertoire, et la clé permettant de définir le répertoire profiles est profiles_directory. Par exemple :

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

Un fichier distinct est utilisé pour la configuration de profil (à la place de l'ajout de paramètres au fichier d'options standard) pour deux raisons :

- La configuration de profil détermine la façon dont les fichiers d'options sont lus, leur définition est donc intrinsèquement difficile.
- Le fichier de configuration de profil est conçu pour être géré automatiquement, à l'aide de scripts, de sorte que dans des cas simples les utilisateurs n'ont pas du tout à s'en préoccuper (mais il peut être édité manuellement en toute sécurité pour prendre en charge des scénarios plus complexes).

Mis à part l'emplacement du répertoire profiles, la seule autre entrée du fichier profiles.cfg est un numéro de port. Par exemple :

```
profile_port, 28501
```

Il s'agit du numéro de port par défaut du prochain profil à créer, qui est incrémenté automatiquement chaque fois qu'un profil est créé à l'aide d'un script. Le fichier profiles.cfg est créé uniquement s'il est nécessaire, il n'existe donc pas dans une toute nouvelle installation.

Démarrage d'un profil

L'exécutable du service (modelerserver.exe) accepte un argument supplémentaire, profile, qui identifie le profil du service :

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

Plusieurs services peuvent s'exécuter à partir de la même installation si chaque service utilise un profil différent. Si l'argument de profil est omis, le service utilise les paramètres par défaut de l'installation commune sans remplacement de profil.

Lorsqu'il est appelé avec l'argument profile, le service :

- Lit [chemin d'installation du serveur]\config\profiles.cfg pour obtenir l'emplacement du répertoire de profils.
- Lit [répertoire de profils]\[nom du profil]\config\options.cfg pour obtenir la configuration de profil (en particulier, le numéro de port).

Si une étape échoue pour une raison quelconque, le service génère un message dans le journal et s'arrête. Si le service est appelé avec un profil et qu'il ne peut pas le charger, il ne s'exécute pas.

Variables d'environnement

Le service définit des variables d'environnement supplémentaires de sorte que les noms de chemin, etc., puissent être exprimés sans connaissance du profil actuel.

Tableau 1. Variables d'environnement

Variable	Valeur
PROFILE_NAME	Nom du profil actuel, ou chaîne vide si aucun profil n'a été spécifié.
MODELERPROFILE	Chemin complet vers le répertoire du profil actuel (par exemple, \$MODELERSERVER\profiles\%PROFILE_NAME). Si aucun profil n'a été spécifié, la valeur est la même que \$MODELERSERVER.
MODELERDATA	Chemin complet vers le répertoire data du profil actuel (par exemple, \$MODELERPROFILE\data). Si aucun profil n'a été spécifié, la valeur pointe sur le répertoire de données standard \$MODELERSERVER\data

Ces variables d'environnement sont définies par le processus du service, ils ne sont donc visibles que dans ce processus et dans les processus enfant qu'il crée. Si vous définissez ces variables hors du processus de service, elles seront ignorées et redéfinies dans le processus, comme décrit.

Consignation

Chaque service de profil attend un dossier privé, séparé, dans lequel il placera ses fichiers journaux. Il existe une copie de `server_logging.log`, etc., pour chaque profil.

La configuration `log4cxx.properties` par défaut dans le répertoire d'installation config utilise la variable d'environnement `PROFILE_NAME` pour identifier le répertoire de journaux pour le service :

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

Vous pouvez changer l'emplacement des journaux pour tous les profils en modifiant la ligne ci-dessus et en incluant l'une des deux variables d'environnement spécifiques aux profils (`PROFILE_NAME` ou `MODELERPROFILE`). Par exemple, pour relocaliser le répertoire de journaux dans le répertoire de profils :

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

Sinon, vous pouvez changer l'emplacement des journaux pour un profil particulier en créant et en éditant une copie du fichier `log4cxx.properties` dans la configuration de profil.

Scripts de profil

Les scripts décrits dans cette section sont fournis pour vous aider à créer et à gérer les profils SPSS Modeler Server. Tous les scripts sont inclus dans le répertoire `scripts/profiles` du répertoire d'installation de SPSS Modeler Server (par exemple, `C:\Program Files\IBM\SPSS\ModelerServer\17\scripts\profiles`).

Script commun (à toutes les plateformes)

Le script suivant aide à créer et gérer des profils. Des variantes de ce script sont fournis avec différentes extensions pour différentes plateformes (`.bat` pour Windows et `.sh` pour UNIX). L'opération est la même dans chaque cas.

Création d'un profil

create_profile [options] <nom_profil>

Crée un nouveau profil portant le nom spécifié. Le nom du profil doit être approprié pour être utilisé comme nom de répertoire sur l'hôte de serveur (car le script créera un répertoire sous ce nom) et ne doit pas contenir d'espaces. Le nom doit être différent des noms de profil existants.

Options :

-d, --profiles-directory<répertoire_profils>

Indique le répertoire de profils dans lequel ce profil et tous les profils suivants doivent être créés. Cette option n'est requise que pour le premier profil, mais il est recommandé de la spécifier à chaque fois. Si vous l'omettez la première fois, un emplacement par défaut sera choisi. Si vous changez le répertoire de profils lors d'un appel suivant, le nouveau profil sera créé dans le nouvel emplacement, mais les profils existants seront ignorés sauf s'ils sont déplacés séparément dans le nouvel emplacement.

-t, --template <nom_profil>

Indique le nom d'un profil existant à utiliser comme modèle. La configuration de profil est copiée à partir du profil existant vers le nouveau profil, et seul le numéro de port est changé.

-p, --port-number <numéro_port>

Indique le numéro de port pour le service de profil. Le numéro de port doit être unique pour ce profil. Si vous omettez le numéro de port, un numéro par défaut est choisi.

-g, --group-name <nom_groupe>

Indique le nom d'un groupe de système d'exploitation qui disposera d'un accès exclusif à ce profil. Le profil est configuré de manière à réserver la connexion aux membres de ce groupe.

Les droit d'accès au système de fichiers ne sont pas changés, vous devez donc effectuer cette action séparément.

Exemples :

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

Crée un nouveau profil appelé comet dans le répertoire C:\Modeler\Profiles. Le profil utilisera un numéro de port d'écoute par défaut. Pour déterminer le numéro de port, ouvrez le fichier options.cfg généré pour le profil (dans cet exemple, C:\Modeler\Profiles\comet\config\options.cfg).

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

Crée un nouveau profil appelé meteor dans le répertoire C:\Modeler\Profiles (mémorisé depuis la commande précédente). Ce profil utilisera le port d'écoute 28510 et la connexion sera réservée aux membres du groupe Meteor Users. Toutes les autres options de configuration seront copiées à partir du profil existant comet.

Scripts Windows

Ces scripts sont fournis pour vous aider à créer et à gérer les services Windows pour les profils SPSS Modeler Server. Ils utilisent le programme Windows Service Control (SC.EXE) pour effectuer les opérations demandées, et le résultat du script est fourni par SC.EXE sauf indication contraire. Vous devez disposer des privilèges administrateur sur la machine locale pour effectuer la plupart de ces tâches.

Pour plus d'informations sur SC.EXE, voir Microsoft's TechNet documentation.

Création d'un service Windows pour un profil

```
create_windows_service [options] <nom_profil>
```

Crée le service Windows pour un profil spécifié. Vous devez disposer des privilèges administrateur pour créer un service. Utilisez la console de gestion des services pour définir des propriétés supplémentaires pour le service après sa création (par exemple, pour définir les détails du compte pour la connexion au service).

Options :

-u, --service-user <nom_compte>

Indique le compte utilisé pour la connexion au service (passim). Il peut s'agir d'un compte utilisateur local, d'un compte utilisateur de domaine ou du nom de l'ordinateur local (qui fait office de compte système local). Par défaut, il s'agit du compte du système local. Si vous spécifiez un compte autre que le compte du système local, vous devez accéder à la console de gestion des services et définir le mot de passe du compte avant le démarrage du service.

-s, --register-spn

Enregistre un nom principal de service (SPN) pour le service afin que les clients puissent se connecter via la connexion unique Kerberos. Vous devez préciser le compte utilisateur du service dans ce cas (-u) pour enregistrer le SPN sur ce compte. Vous devez disposer des privilèges d'administrateur de domaine pour pouvoir utiliser cette option (ou avoir reçu le droit d'enregistrer un SPN).

-H, --service-host <nom_hôte>

Indique le nom d'hôte à utiliser dans la construction du SPN. Il doit s'agir du nom de l'hôte auquel les clients se connecteront, et il doit être qualifié avec un nom de domaine correspondant au domaine Kerberos (dans une configuration Active Directory simple, le nom de domaine et le nom de domaine Kerberos sont les mêmes).

Exemples :

```
scripts\profiles\create_windows_service.bat comet
```

Crée un service Windows pour le profil comet. Le service appartient au compte système local, et les clients sont supposés se connecter à l'aide d'un nom d'utilisateur et d'un mot de passe.

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

Crée un service Windows pour le profil meteor. Le service appartient au compte de domaine ProjectMeteor, et les clients peuvent se connecter à l'aide d'une connexion unique. Le service ne démarre pas tant que vous n'avez pas accédé à la console de gestion des services et défini le mot de passe du compte ProjectMeteor. Le compte recevra automatiquement le droit de se connecter en tant que service.

Suppression d'un service Windows pour un profil

delete_windows_service [options] <noms_profils...>

Supprime les services Windows pour les profils spécifiés. Vous devez disposer des privilèges administrateur pour supprimer un service.

Options :

-s, --summary

Répertorie les noms des services supprimés. Les services qui n'existent pas ou qui ne peuvent pas être supprimés ne sont pas répertoriés. Sans cette option, l'état de suppression de tous les services spécifiés sont répertoriés.

-a, --all

Supprime les services pour tous les profils.

Exemples :

```
scripts\profiles\delete_windows_service.bat comet
```

Supprime les services Windows pour le profil comet.

```
scripts\profiles\delete_windows_service.bat --all
```

Supprime les services Windows pour tous les profils.

Démarrage d'un service Windows pour un profil

start_windows_service [options] <noms_profils...>

Démarre les services Windows pour les profils spécifiés. Vous devez disposer des privilèges administrateur pour démarrer un service.

Options :

-s, --summary

Répertorie les noms des services démarrés. Les services déjà en cours d'exécution ou qui ne peuvent pas être démarrés ne sont pas répertoriés. Sans cette option, l'état des services répertoriés est listé.

-a, --all

Démarre les services pour tous les profils.

Exemples :

scripts\profiles\start_windows_service.bat -s comet meteor

Tente de démarrer les services Windows pour les profils comet et meteor, et répertorie les noms des services dont le démarrage a abouti.

Arrêt d'un service Windows pour un profil

stop_windows_service [options] <noms_profils...>

Arrête les services Windows pour les profils spécifiés. Vous devez disposer des privilèges administrateur pour arrêter un service.

Options :

-s, --summary

Répertorie les noms des services arrêtés. Les services déjà arrêtés ou qui ne peuvent pas l'être ne sont pas répertoriés. Sans cette option, l'état des services répertoriés est listé.

-a, --all

Arrête les services pour tous les profils.

Exemples :

scripts\profiles\stop_windows_service.bat -a -s

Tente d'arrêter les services Windows pour tous les profils, et imprime les noms de ceux dont l'arrêt a abouti. L'ensemble de tous les profils est obtenu à partir du répertoire profiles.

Analyse de l'état d'un service Windows pour un profil

query_windows_service [options] <noms_profils...>

Affiche l'état des services Windows pour les profils spécifiés. Vous n'avez pas besoin de disposer des privilèges administrateur pour analyser un service.

Options :

-s, --summary

Répertorie les noms des services et leur état actuel (RUNNING, STOPPED, etc.). Si un service ne peut pas être analysé pour une raison quelconque (par exemple, s'il n'existe pas), l'état renvoyé est UNKNOWN. Sans cette option, l'état complet des services répertoriés est listé.

-a, --all

Analyse l'état des services pour tous les profils.

Exemples :

scripts\profiles\query_windows_service.bat -a

Signale l'état complet des services pour tous les profils.

Script UNIX

Le script UNIX existant qui gère le service SPSS Modeler Server accepte désormais un argument `profile` supplémentaire qui permet aux services de profil SPSS Modeler Server d'être gérés indépendamment.

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

Gère le service SPSS Modeler Server principal. Pour plus d'informations, voir Chapitre 4, «Administration de IBM SPSS Modeler», à la page 27.

Options :

-p, --profile <nom_profil>

Gère l'instance de service pour le profil spécifié. Lorsque cet argument est utilisé, la commande spécifiée s'applique uniquement à l'instance du profil spécifié. Lorsque cet argument est absent, la commande `start` démarre uniquement l'instance par défaut (un service sans profil), mais les commandes `stop`, `kill` et `list` s'appliquent à toutes les instances actives.

Exemples :

./modelersrv.sh --profile comet start

Démarre le service pour le profil comet.

./modelersrv.sh --profile meteor start

Démarre le service pour le profil meteor.

./modelersrv.sh list

Liste les processus pour tous les services actifs.

./modelersrv.sh --profile comet stop

Arrête le service pour le profil comet.

./modelersrv.sh stop

Arrête tous les services actifs.

Actuellement aucune méthode de démarrage des services de profil SPSS Modeler Server n'est prise en charge automatiquement sous UNIX. Le script `auto.sh` standard est disponible pour configurer le système pour qu'il démarre et arrête le service SPSS Modeler Server principal avec le système d'exploitation, mais cela ne s'applique qu'au service par défaut, et pas à n'importe quel service de profil.

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 Modeler Administration Console, qui est disponible gratuitement pour les clients actuels de IBM SPSS Modeler. Pour plus d'informations, voir la rubrique «Utilisation de IBM SPSS Modeler Administration Console», à la page 38.
- Utilisez le fichier texte `options.cfg`, qui se trouve dans le répertoire `[dossier d'installation du serveur]/config`. Pour plus d'informations, voir la rubrique «Utilisation du fichier options.cfg», à la page 46.

Nous vous conseillons d'installer IBM SPSS Modeler Administration Console 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 IBM SPSS Modeler Server, alors que IBM SPSS Modeler Administration Console vous permet d'autoriser toute personne disposant d'un compte utilisateur à modifier ces options. De plus, IBM SPSS Modeler Administration Console fournit des informations supplémentaires sur les processus serveur.

Vous pouvez donc en contrôler l'utilisation et les performances. La plupart des options de configuration peuvent être modifiées sans redémarrer IBM SPSS Modeler Server, ce qui n'est pas le cas du fichier de configuration.

Pour plus d'informations sur l'utilisation de IBM SPSS Modeler Administration Console 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 noeud pour chaque SPSS Modeler Server à administrer. Les options de configuration correspondant au serveur sélectionné apparaissent dans le panneau situé à droite. Vous devez d'abord paramétrer une connexion pour chaque serveur à administrer.

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(administrateurs)` du 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, "jdupont,mcharlier,asimon"
```

- 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.

Configuration de l'accès avec contrôle d'accès utilisateur

Pour utiliser la console Modeler Administration Console afin d'effectuer des mises à jour dans une configuration SPSS Modeler Server installée sur une machine Windows dotée du contrôle d'accès utilisateur (User Access Control - UAC) vous devez disposer des droits de lecture, d'écriture et

d'exécution définis sur le répertoire *config* et sur le fichier *options.cfg*. Ces droits (NTFS) doivent être définis au niveau de l'utilisateur et non au niveau du groupe, en raison du type d'interaction entre UAC et les droits NTFS.

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

1. Vérifiez que le service IBM SPSS Modeler Server est démarré.
2. Dans le menu Fichier, choisissez **Nouveau**, puis **Connexion du serveur administré**.
3. 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 **Administré IBM SPSS Modeler Server**, puis cliquez sur **Suivant**.
4. 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

1. Dans le panneau Administrateur du serveur, double-cliquez pour sélectionner le serveur auquel vous souhaitez vous connecter.
2. 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, voir la rubrique «Configuration de l'accès avec Modeler Administration Console», à la page 38.

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 IBM 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 IBM 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

Les options de configuration de SPSS Modeler Server figurent dans la sous-fenêtre Configuration. 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.

Remarque : Si un utilisateur non racine (non superutilisateur, non-root) souhaite modifier ces options, il doit disposer d'un droit d'accès en écriture pour le répertoire **config** de SPSS Modeler Server.

Connexions/Sessions

Nombre maximal de connexions. (`max_sessions`) Nombre maximal de sessions de serveur simultanément. La valeur -1 indique qu'il n'y a pas de limite.

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

Connexion à Analytic Server

Activer SSL pour Analytic Server (`as_ssl_enabled`) Indiquez Y pour chiffrer les communications entre Analytic Server et SPSS Modeler, sinon N.

Hôte (`as_host`) Adresse IP d'Analytic Server.

Numéro de port (`as_port`) Numéro de port d'Analytic Server.

Racine de contexte (`as_context_root`) Racine de contexte d'Analytic Server.

Locataire (`as_tenant`) Locataire dont l'installation SPSS Modeler Server est membre.

Invite de saisie du mot de passe (`as_prompt_for_password`) Indiquez N si SPSS Modeler Server est configuré avec le même système d'authentification des utilisateurs et des mots de passe que celui utilisé sur Analytic Server ; par exemple, lorsque vous utilisez l'authentification Kerberos, sinon, indiquez Y.

Remarque : Si vous avez l'intention d'utiliser la connexion unique (SSO) Kerberos, vous devez définir des options supplémentaires dans le fichier `options.cfg`. Pour plus d'informations, voir la rubrique "Options visibles dans `options.cfg` " dans la suite de ce chapitre.

Accès aux fichiers de données

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. Sous Windows, pour indiquer plusieurs répertoires, utilisez des points-virgules (par exemple, [dossier d'installation du serveur]/données;c:/données;c:/temp). Sous Linux et UNIX, utilisez des signes deux-points (:) au lieu des points-virgules. Le chemin d'accès des fichiers de données doit inclure tous les chemins spécifiés par le paramètre `temp_directory` décrit ci-dessous.

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. A partir de l'édition 17, le seul fichier de programme pour lequel l'accès est restreint est l'exécutable Python (voir **Chemin de l'exécutable Python** 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`) Maxs, créés durant l'exécution du flux (ne s'applique pas aux fichiers de données SAS et SPSS Statistics). La valeur -1 indique qu'il n'y a pas 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. Si l'accès aux fichiers de programme est restreint, vous devez ajouter le répertoire contenant l'exécutable Python aux chemins des fichiers de programme (voir **Restreindre l'accès au chemin des fichiers de programme uniquement** ci-dessus).

Performances/Optimisation

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 jeux 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. Si ce paramètre est désactivé sur le serveur, le client ne peut pas l'activer. Mais s'il est activé sur le serveur, le client peut choisir de le désactiver.

Parallélisme. (`max_parallelism`) Décrit le nombre d'unités d'exécution 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 d'unités d'exécution 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 de mémoire 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 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, IBM 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, IBM SPSS Modeler aura besoin de deux ports ouverts (en plus du port du serveur principal) disponibles par flux simultané, plus 3 ports supplémentaires pour chaque connexion ODBC issue d'un client connecté (2 ports pour la connexion ODBC pendant la durée de cette connexion ODBC, et un port temporaire supplémentaire pour l'authentification).

Remarque : Une connexion ODBC est une entrée dans la liste des connexions de la base de données, qui peut être partagée entre plusieurs noeuds de base de données spécifiés avec la même connexion de base de données.

Remarque : Il est possible que les ports d'authentification puissent être partagés si les connexions sont établies à différents moments).

Remarque : Les pratiques recommandées indiquent d'utiliser les mêmes ports pour communiquer avec IBM SPSS Collaboration and Deployment Services et avec SPSS Modeler Client. Ces ports peuvent être définis via les paramètres `max_server_port` et `min_server_port`.

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

Longueur de chaîne SQL maximale. (`max_sql_string_length`) Pour une chaîne importée de la base de données avec SQL, le nombre maximum de caractères qui sera passé avec succès. En fonction du système d'exploitation, les valeurs de chaîne plus longues supérieures à cette valeur peuvent être tronquées à droite sans avertissement. La plage valide est comprise entre 1 et 65 535 caractères. Cette propriété s'applique aussi au noeud d'exportation de base de données.

Remarque : La valeur par défaut de ce paramètre est 2048. Si le texte que vous analysez comporte plus de 2048 caractères (ce qui peut arriver, par exemple, si vous utilisez le noeud Fil de nouvelles SPSS Modeler Text Analytics), il est conseillé d'augmenter cette valeur si vous travaillez en mode natif, sinon vos résultats risquent d'être tronqués. Si vous utilisez une base de données et des fonctions définies par l'utilisateur (UDF), cette restriction ne s'applique pas. Il peut s'agir de différences dans les résultats entre les modes natif et UDF.

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 ; veuillez 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. Si ce paramètre est désactivé sur le serveur, le client ne peut pas l'activer. Mais s'il est activé sur le serveur, le client peut choisir de le désactiver.

Longueur de chaîne SQL par défaut. (`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 Typier en amont n'existe. Si vous disposez de valeurs plus importantes que celles de vos données, instantiez un noeud type en amont avec ces valeurs ou définissez ce paramètre sur une valeur assez importante pour accommoder ces valeurs de chaîne.

Consignation du type SQL pris en charge. (`supported_sql_type_logging`). Pour le débogage uniquement ; ne pas modifier sauf si le représentant technico-commercial vous a donné des instructions spécifiques.

Activer l'UDF de base de données. (`db_udf_enabled`). Lorsque cette option est définie sur Y (valeur par défaut), elle a pour effet que l'option de génération SQL génère du SQL de fonction définie par l'utilisateur (FDU) au lieu de SQL pur SPSS Modeler. Le SQL FDU surpasse généralement le SQL pur.

SSL

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

Magasin de clés. (`ssl_keystore`) Fichier de la base de données de clés SSL à charger lorsque le serveur démarre (chemin complet ou chemin relatif vers le répertoire d'installation SPSS Modeler).

Fichier de dissimulation du magasin de clés. (`ssl_keystore_stash_file`) Nom du fichier de mot de passe secret de la base de données de clés à charger lorsque le serveur démarre (chemin complet ou chemin relatif vers le répertoire d'installation SPSS Modeler). Si vous souhaitez laisser vide ce paramètre et être invité à saisir le mot de passe lors du démarrage de SPSS Modeler Server, reportez-vous aux instructions suivantes :

- Sous Windows :

1. Vérifiez que le paramètre `ssl_keystore_stash_file` dans `options.cfg` ne comporte pas de valeur.

2. Redémarrez SPSS Modeler Server. Vous êtes invité à entrer un mot de passe. Entrez le mot de passe correct et cliquez sur **OK**. Le serveur démarre.
- Sous Linux/UNIX :
 1. Vérifiez que le paramètre `ssl_keystore_stash_file` dans `options.cfg` ne comporte pas de valeur.
 2. Recherchez la ligne suivante dans le fichier `modelersrv.sh` :


```
if "$INSTALLEDPATH/$SCLEMDNAME" -server $ARGS; then
```
 3. Ajoutez le commutateur `-request_ssl_password` comme ceci :


```
if "$INSTALLEDPATH/$SCLEMDNAME" -request_ssl_password -server $ARGS; then
```
 4. Redémarrez SPSS Modeler Server. Vous êtes invité à entrer un mot de passe. Entrez le mot de passe correct et cliquez sur **OK**. Le serveur démarre.

Libellé du certificat du magasin de clés. (`ssl_keystore_label`) Libellé du certificat spécifié.

Remarque : Pour utiliser la console d'administration avec un serveur configuré pour SSL, vous devez importer les certificats requis par SPSS Modeler Server dans le magasin de clés de confiance Deployment Manager (sous `../jre/lib/security`).

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

Configuration du coordinateur de processus

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.

Racine de contexte. (`cop_context_root`) URL du Service de coordinateur de processus.

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 pour l'authentification sur le service du Coordinateur de processus.

Remarque : Si vous mettez à jour le fichier `options.cfg` manuellement au lieu d'utiliser Modeler Administration Console, vous devez manuellement chiffrer la valeur `cop_password` que vous indiquez dans le fichier. Les mots de passe en texte normal ne sont pas valides et entraînent l'échec de l'enregistrement auprès du Coordinateur de processus.

Suivez la procédure ci-dessous pour chiffrer manuellement le mot de passe :

1. Ouvrez une invite de commande et accédez au répertoire SPSS Modeler `./bin`.
2. Lorsque vous y êtes invité, entrez le nom d'utilisateur (la valeur `cop_user_name` que vous spécifiez dans le fichier `options.cfg`) et appuyez sur Entrée.
3. Lorsque vous y êtes invité, entrez le mot de passe.

Le mot de passe codé s'affiche entre guillemets doubles sur la ligne de commande comme élément de la chaîne renvoyée. Par exemple :

```
C:\Program Files\IBM\SPSS\Modeler\17\bin>pwutil
User name: copuser
Password: Pass1234
copuser, "0Tqb4n.ob0wrs"
```

4. Copiez le mot de passe codé, sans les guillemets doubles, et collez-le entre guillemets doubles qui existent déjà pour la valeur `cop_password` dans le fichier `options.cfg`.

***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.

SSL activé. (`cop_ssl_enabled`) Détermine si SSL est utilisé pour se connecter au coordinateur ou au serveur de processus. Si cette option est utilisée, vous devez importer le fichier de certificat SSL dans l'environnement d'exécution Java SPSS Modeler Server. Pour ce faire, vous devez obtenir le fichier de certificat SSL, ainsi que son nom d'alias et son mot de passe. Exécutez ensuite la commande suivante sur SPSS Modeler Server:

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

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. Une pondération supérieure attire davantage 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.

Chemin de données par défaut. (`cop_service_default_data_path`) Chemin de données par défaut d'une installation IBM SPSS Modeler Server enregistrée par le coordinateur de processus.

Options visibles dans `options.cfg`

La plupart des options de configuration peuvent être modifiées à l'aide de IBM SPSS Modeler Administration Console. Il existe cependant certaines exceptions, comme celles décrites dans cette section. Les options de cette section doivent être modifiées en éditant le fichier `options.cfg`. Pour plus d'informations, voir «Utilisation de IBM SPSS Modeler Administration Console», à la page 38 et «Utilisation du fichier `options.cfg`», à la page 46. Notez qu'il peut exister des paramètres supplémentaires dans `options.cfg` qui ne sont pas répertoriés ici.

Remarque : Ces informations s'appliquent uniquement à un serveur distant IBM SPSS Modeler Server, par exemple).

`administrators`. Spécifiez les noms d'utilisateurs des utilisateurs auxquels vous souhaitez octroyer l'accès en tant qu'administrateurs. Pour plus d'informations, voir la rubrique «Configuration de l'accès avec Modeler Administration Console», à la page 38.

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

`fips_encryption`. Permet un chiffrement FIPS. La valeur par défaut est N.

group_configuration. Lorsqu'il est activé, IBM SPSS Modeler Server vérifie le fichier groups.cfg qui contrôle les utilisateurs qui se connectent au serveur. Pour plus d'informations, voir «Configuration des groupes», à la page 53.

max_transfer_size. Pour une utilisation système interne uniquement. **Ne pas modifier.**

shell. (Serveurs UNIX uniquement) Remplace le paramètre par défaut du shell UNIX, par exemple shell, "/usr/bin/ksh". Par défaut, IBM SPSS Modeler utilise le shell défini dans le profil utilisateur de l'utilisateur se connectant à IBM SPSS Modeler Server.

start_process_as_login_user. Définissez sur Y si vous exécutez SPSS Modeler Server à l'aide d'une base de données de mots de passe privée, en démarrant le service du serveur à partir d'un compte non racine.

use_bigint_for_count. Lorsque le nombre d'enregistrements à compter est supérieur à ce qu'un entier normal ($2^{32}-1$) peut comporter, définissez cette option sur Y. Lorsque cette option est définie sur Y, et qu'un flux est connecté à DB2 ou à SQL Server, une fonction sera utilisée chaque fois qu'un nombre d'enregistrements sera nécessaire (par exemple, le champ **Record_Count** généré par le noeud Agréger).

cop_ssl_enabled. Définissez cette option sur Y si vous utilisez SSL pour vous connecter au coordinateur ou au serveur de processus. Si cette option est utilisée, vous devez importer le fichier de certificat SSL dans l'environnement d'exécution Java SPSS Modeler Server. Pour ce faire, vous devez obtenir le fichier de certificat SSL, ainsi que son nom d'alias et son mot de passe. Exécutez ensuite la commande suivante sur SPSS Modeler Server:

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

cop_service_default_data_path. Vous pouvez utiliser cette option pour définir le chemin de données par défaut d'une installation IBM SPSS Modeler Server enregistrée par le coordinateur de processus.

Par défaut, l'authentification Analytic Server à l'aide de la méthode Kerberos n'est pas activée. Pour autoriser l'authentification Kerberos, utilisez les trois propriétés suivantes :

as_kerberos_auth_mode. Pour activer l'authentification Kerberos, donnez à cette option la valeur Y.

as_kerberos_krb5_conf. Indiquez le chemin du fichier de configuration Kerberos que Analytic Server doit utiliser ; par exemple, c:\windows\krb5.conf.

as_kerberos_krb5_spn. Indiquez le SPN Kerberos Analytic Server ; par exemple, HTTP/ashost.mydomain.com@MYDOMAIN.COM.

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 actualisées à la vitesse indiquée (par défaut, toutes les minutes). Pour actualiser manuellement les données, cliquez sur le bouton **Actualiser**. Pour afficher uniquement les processus SPSS Modeler Server dans la liste, cliquez sur le bouton **Filtrer les processus non-SPSS Modeler**.

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 : la plupart des options de configuration peuvent être modifiées à l'aide de IBM SPSS Modeler Administration Console, plutôt qu'à l'aide de ce fichier de configuration. Il existe cependant quelques exceptions. Pour plus d'informations, voir la rubrique «Options visibles dans options.cfg», à la page 45.

En utilisant IBM SPSS Modeler Administration Console, vous évitez les redémarrages du serveur pour toutes les options, excepté pour le port du serveur. Pour plus d'informations, voir la rubrique «Utilisation de IBM SPSS Modeler Administration Console», à la page 38.

Remarque : Ces informations s'appliquent uniquement à un serveur distant IBM SPSS Modeler Server, par exemple).

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 :

1. Ouvrez le fichier *options.cfg* dans un éditeur de texte.
2. Recherchez les options qui vous intéressent. Pour obtenir la liste complète des options, voir «Configuration de SPSS Modeler Server», à la page 39.
3. 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.
4. Enregistrez le fichier.
5. Arrêtez et redémarrez IBM SPSS Modeler Server pour que les modifications prennent effet. Pour plus d'informations, voir la rubrique «Démarrage et arrêt de IBM SPSS Modeler Server», à la page 27.

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 IBM 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 IBM 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 IBM SPSS Modeler *cache_connection* dans un fichier de configuration de base de données personnalisé. En outre cela peut aussi rendre IBM 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 IBM SPSS Modeler.

Pour fermer les connexions inutilisées à une base de données :

1. Recherchez le répertoire [*chemin d'installation du serveur*]/*config*.
2. Ajoutez le fichier suivant (ou ouvrez-le s'il existe déjà) :
odbc-custom-properties.cfg
3. Ajoutez la ligne suivante au fichier :
cache_connection, N

4. Enregistrez le fichier, puis fermez-le.
5. Redémarrez IBM SPSS Modeler Server pour que les modifications prennent effet.

Utilisation du protocole SSL pour sécuriser le transfert de données

Le protocole SSL (Secure Sockets Layer) permet de chiffrer les données transférées entre deux ordinateurs. Le protocole SSL sécurise la communication entre les ordinateurs. SSL peut chiffrer 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.

1. Lorsqu'un client se connecte à un serveur, le client authentifie ce serveur avec le certificat de clé publique.
2. Le client génère ensuite un chiffre au hasard, code ce chiffre avec la clé publique du serveur et renvoie le message chiffré au serveur.
3. Le serveur décode ce chiffre aléatoire avec sa clé privée.
4. Avec le chiffre aléatoire, le serveur et le client créent les clés de session utilisées pour le chiffrement et le déchiffrement 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 :

1. Installez OpenSSL sur l'ordinateur serveur.
2. Obtenez et installez le certificat et les clés SSL.
3. Activez et configurez SSL dans l'application d'administration du serveur (IBM SPSS Collaboration and Deployment Services Deployment Manager).
4. Si nécessaire, installez le chiffrement de type Unlimited Strength sur les ordinateurs client.
5. Si vous utilisez un certificat signé, copiez-le sur l'ordinateur client.
6. Donnez aux utilisateurs 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, IBM SPSS Statistics Server est le *client*.

Obtention et installation du certificat SSL et des clés

Les premières étapes à suivre pour configurer la prise en charge SSL sont les suivantes :

1. Procurez-vous un certificat SSL et un fichier de clés. Vous disposez pour ce faire de trois méthodes :
 - Achetez-les auprès d'une autorité de certification publique (comme VeriSign, Thawte ou Entrust). L'autorité de certification publique (CA) 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 tierce. Si cette approche est choisie, le certificat racine de l'autorité de certification doit être importé dans les fichiers du magasin de clés client et serveur. Pour plus d'informations, voir la rubrique «Importation d'un certificat d'autorité de certification tierce», à la page 50.

- Générez les fichiers de clé et de certificat par le biais d'une autorité de certification autosignée. Les étapes pour ce faire sont les suivantes :
 - a. Préparez la base de données de clés. Pour plus d'informations, voir la rubrique «Création d'une base de données de clés SSL», à la page 50.
 - b. Créez le certificat autosigné. Pour plus d'informations, voir la rubrique «Création d'un certificat SSL autosigné», à la page 50.
- 2. Copiez les fichiers .kdb et .sth créés à l'étape 1 dans un répertoire auquel IBM SPSS Modeler Server a accès et indiquez le chemin de ce répertoire dans le fichier *options.cfg*.

Remarque : Utilisez les barres obliques comme séparateurs dans le chemin de répertoire.

3. Définissez les paramètres suivants dans le fichier *options.cfg* :
 - `ssl_enabled, Y`
 - `ssl_keystore, "<filename>.kdb"` où `<filename>` est le nom de votre base de données de clés.
 - `ssl_keystore_stash_file, "<filename>.sth"` où `<filename>` est le nom du fichier de mot de passe secret de la base de données de clés.
 - `ssl_keystore_label, <label>` où `<label>` est le libellé de votre certificat.
4. Pour les certificats autosignés ou tiers, installez le certificat sur les systèmes client. Pour les certificats d'autorité de certification achetés, cette étape n'est pas obligatoire. Vérifiez que les droits d'accès sont définis pour refuser toute navigation dans le répertoire qui contient le certificat. Pour plus d'informations, voir la rubrique «Installation d'un certificat SSL autosigné», à la page 50.

Configuration de l'environnement pour l'exécution de GSKit : GSKCapiCmd est un outil de ligne de commande qui n'est pas basé sur Java, et Java ne doit pas nécessairement être installé sur votre système pour que vous puissiez utiliser cet outil. Il se trouve dans le dossier `<répertoire_installation_Modeler>/bin`. Le processus de configuration de votre environnement d'exécution d'IBM Global Security Kit (GSKit) varie en fonction de la plateforme utilisée.

Pour une configuration Linux/Unix, ajoutez le répertoire de bibliothèques partagées `<répertoire_installation_Modeler>/lib` à votre environnement :

```
$export <variable_environnement_chemin_bibliothèque_partagée>=<chemin_installation_serveur_modeler>/bin
$export PATH=$PATH:<chemin_installation_serveur_modeler>/bin
```

Le nom de la variable du chemin de bibliothèque partagée dépend de la plateforme.

- AIX utilise le nom de variable LIBPATH.
- HP-UX utilise le nom de variable SHLIB_PATH.
- Linux et Solaris utilisent le nom de variable LD_LIBRARY_PATH.

Par exemple, pour définir l'environnement sous Linux, utilisez :

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

Accès du compte aux fichiers

Veillez à accorder les droits d'accès appropriés aux comptes qui devront accéder aux fichiers SSL :

1. Pour tous les comptes qui sont utilisés par SPSS Modeler pour la connexion, accordez un droit d'accès en lecture aux fichiers SSL.
2. Pour Windows, il ne suffit pas que les comptes soient dans le groupe Administrateurs et qu'un droit d'accès soit accordé à ce groupe lorsque le contrôle d'accès utilisateur (UAC) est activé. Vous devez en plus exécuter l'une des actions suivantes :
 - Accorder des droits aux comptes de façon distincte.
 - Créer un nouveau groupe, ajouter les comptes à ce groupe et accorder au groupe le droit d'accès aux fichiers SSL.

- Désactiver le contrôle d'accès utilisateur.

Création d'une base de données de clés SSL : Utilisez l'outil GSKCapiCmd pour créer votre base de données de clés. Avant d'utiliser cet outil, vous devez configurer votre environnement. Pour plus d'informations, voir la rubrique «Configuration de l'environnement pour l'exécution de GSKit», à la page 49.

Pour créer la base de données de clés, exécutez GSKit et entrez la commande suivante :

```
gsk<ver>capicmd[_64] -keydb -create -populate -db <nom_fichier>.kdb -pw <mot_de_passe> -stash
```

où <ver> est le numéro de version GSKit, <nom_fichier> est le nom que vous souhaitez utiliser pour le fichier de base de données de clés, et <mot_de_passe> est le mot de passe de la base de données de clés.

L'option *-stash* crée un fichier de dissimulation dans le même chemin que la base de données de clés, avec l'extension de fichier .sth. GSKit utilise le fichier de dissimulation pour se procurer le mot de passe de la base de données de clés pour ne pas avoir à le saisir sur la ligne de commande à chaque fois.

Remarque : Il est conseillé d'utiliser une protection de système de fichiers renforcée pour le fichier .sth.

Création d'un certificat SSL autosigné : Pour générer un certificat autosigné et le stocker dans la base de données de clés, utilisez la commande suivante :

```
gsk<ver>capicmd[_64] -cert -create -db <nom_fichier>.kdb -stashed -dn "CN=myserver,OU=mynetwork,
O=mycompany,C=mycountry" -label <libellé> -expire <validité_certificat_en_nombre_
de_jours> -default_cert yes
```

où <ver> est le numéro de version GSKit, <nom_fichier> est le nom du fichier de la base de données de clés, <validité_certificat_en_nombre_de_jours> est le nombre physique de jours constituant la durée de validité du certificat, et <libellé> est un libellé descriptif vous aidant à identifier le fichier (par exemple, *mon_certificat_autosigné*).

Installation d'un certificat SSL autosigné : Pour les machines client qui se connectent au serveur via SSL, vous devez distribuer la part publique du certificat aux clients pour qu'elle puisse être stockée dans leurs bases de données de clés. Pour ce faire :

1. Extrayez la partie publique dans un fichier via la commande suivante :

```
gsk<ver>capicmd[_64] -cert -extract -db <nom_fichier>.kdb -stashed -label <libellé>
-format ascii -target mycert.arm
```

2. Distribuez *mycert.arm* aux clients. Ce fichier doit être copié dans leur répertoire *jre/bin*.
3. Ajoutez le nouveau certificat à la base de données de clés des clients via la commande suivante :

```
keytool -import -alias <libellé> -keystore cacerts -file mycert.arm
```

Si un message vous le demande, entrez un mot de passe à l'aide de la commande *changeit*. L'outil de clé est situé dans le répertoire <rep_install_Modeler>\jre\bin.

4. Copiez les nouveaux fichiers cacerts générés dans le répertoire *jre\lib\security* pour remplacer la version existante.

Importation d'un certificat d'autorité de certification tierce : Au lieu de vous procurer un certificat auprès d'une autorité de certification (CA) connue ou de créer un certificat autosigné, vous pouvez utiliser une autorité de certificat tierce pour signer vos certificats serveur. Le client et le serveur doivent pouvoir accéder au certificat racine de l'autorité de certification tierce pour vérifier les certificats serveur signés par l'autorité de certification tierce. Pour ce faire :

1. Procurez-vous le certificat racine de l'autorité de certification tierce. La procédure pour ce faire varie en fonction des procédures de l'autorité de certification tierce. Les autorités de certification tierces rendent souvent leurs certificats racine téléchargeables.
2. Ajoutez le certificat à la base de données de clés des serveurs via la commande suivante :

```
gsk<ver>capicmd[_64] -cert -add -db <nom_fichier>.kdb -stashed -label <libellé>
-file <certificat_CA>.crt
-format binary -trust enable
```

3. Ajoutez le certificat à la base de données de clés des clients via la commande suivante :

```
C:> cd <chemin_accès_Modeler_Client>\jre\bin
C:> keytool -import -keystore ..\lib\security\cacerts -file <certificat_CA>.crt -alias <libellé>
```

Si un message vous le demande, entrez un mot de passe à l'aide de la commande *changeit*. L'outil de clé est situé dans le répertoire *<rép_install_Modeler>\jre\bin*.

4. Validez la base de données de clés du serveur avec le certificat CA racine à l'aide de la commande suivante :

```
gsk<ver>capicmd[_64] -cert -validate -db <nom_fichier>.kdb -stashed -label <libellé>
```

La validation aboutit si le message suivant apparaît : OK.

Remarque : Les commandes expliquées ci-dessus utilisent un certificat racine d'autorité de certification tierce au format binaire. Si le certificat est au format ASCII, utilisez l'option *-format ascii*.

Le paramètre *-db* indique le nom de la base de données de clés dans laquelle vous importez le certificat racine d'autorité de certification tierce.

Le paramètre *-label* indique le libellé à utiliser pour le certificat racine d'autorité de certification tierce dans le fichier de la base de données de clés. Vous pouvez indiquer le libellé de votre choix car il n'est pas lié aux libellés utilisés dans le fichier *options.cfg* d'IBM SPSS Modeler.

Le paramètre *-file* indique le fichier qui contient le certificat racine d'autorité de certification tierce.

Activation et configuration de SSL dans IBM SPSS Collaboration and Deployment Services Deployment Manager

1. Si vous installez un certificat SSL autosigné, copiez le fichier *cacerts* que vous avez créé dans le répertoire *<répertoire_installation_Deployment_Manager>\jre\lib\security*. Pour plus d'informations, voir la rubrique «Installation d'un certificat SSL autosigné», à la page 50.
2. Démarrez l'application d'administration du serveur (IBM SPSS Collaboration and Deployment Services Deployment Manager) et connectez-vous au serveur.
3. Sur la page de configuration, définissez **Secure Sockets Layer** sur *Oui* .
4. Dans **fichier de clé publique SSL**, indiquez le chemin d'accès complet au fichier de clé publique.
5. 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**.

6. A partir des menus, sélectionnez :

Fichier > Enregistrer

7. 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.

Installation du chiffrement de type Unlimited Strength

Le chiffrement de type US Export-Strength est activé dans le logiciel Java Runtime Environment livré avec le produit. Pour améliorer la sécurité de vos données, la mise à niveau vers le chiffrement de type Unlimited Strength est recommandée. Cette procédure doit être répétée à la fois pour les installations client et serveur.

Pour installer la fonction de chiffrement de type Unlimited Strength (puissance illimitée)

1. Téléchargez Unrestricted SDK JCE policy files à partir d'IBM.com (sélectionnez les fichiers applicables à Java 7).
2. Extrayez les fichiers de règles sans restriction stockés dans le fichier compressé. Le fichier compressé contient un fichier `US_export_policy.jar` et un fichier `local_policy.jar`.
3. Sauvegardez les copies existantes de `US_export_policy.jar` et `local_policy.jar` dans le répertoire `jre/lib/security`.
4. Remplacez les copies existantes des fichiers `US_export_policy.jar` et `local_policy.jar` par les deux fichiers que vous avez téléchargés et extraits.
5. Redémarrez le client ou le serveur IBM SPSS Modeler, selon le cas.

Indiquer aux utilisateurs d'activer SSL

Lorsque les utilisateurs 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.

Lorsque les utilisateurs se connectent au serveur via un produit client, ils doivent activer SSL dans la boîte de dialogue pour se connecter au serveur.

Connexion SSL Cognos BI

Pour vous connecter au serveur Cognos BI avec HTTPS et un port sécurisé SSL, vous devez d'abord modifier certains des paramètres du répartiteur externe et interne Cognos BI. Pour plus d'informations sur ces modifications, voir le guide d'administration et de configuration de serveur Cognos BI.

Une fois les paramètres du répartiteur modifiés, importez la certification SSL créée dans Cognos BI dans l'environnement JRE SPSS Modeler en exécutant les étapes suivantes :

1. Dans la configuration Cognos, définissez un mot de passe pour le magasin de clés IBM Cognos BI :
 - a. Dans la fenêtre de l'Explorateur, cliquez sur **Cryptographie > Cognos**.
 - b. Dans la fenêtre Propriétés, sous **Paramètres de clé de chiffrement**, définissez le mot de passe du magasin de clés de chiffrement.
 - c. Dans le menu Fichier, sélectionnez Sauvegarder.
 - d. Dans le menu Actions, sélectionnez Redémarrer.
2. Depuis la ligne de commande, accédez au répertoire `c10_location\bin`.
3. Définissez la variable d'environnement `JAVA_HOME` sur l'emplacement JRE (Java™ Runtime Environment) utilisé par le serveur d'applications qui exécute Cognos BI. Par exemple :

```
set JAVA_HOME=c10_location\bin\jre\<version>
```
4. À partir de la ligne de commande, exécutez l'outil de certificat. Par exemple :

```
ThirdPartyCertificateTool.bat -E -T -r ca.cer -k ..\configuration\encryptkeypair\jEncKeystore -p <mot de passe>
```
5. Copiez le fichier `ca.cer` dans l'emplacement SPSS Modeler Server..
6. Ouvrez une ligne de commande et accédez au dossier `<EmplacementInstallationModeler>\jre\bin`.
7. Exécutez la commande pour importer le certificat. Par exemple :

```
.\keytool -import -alias ca -file <répertoire de ca.cer>\ca.cer -keystore "<EmplacementInstallationModeler>\jre\lib\security\cacerts"
```

Vous pouvez ensuite utiliser HTTPS et le répartiteur sécurisé SSL pour vous connecter à Cognos BI. Par exemple :

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

Connexion SSL Cognos TM1

Pour vous connecter à Cognos TM1 avec HTTPS et un port sécurisé SSL, procédez comme suit :

1. Configurez Tomcat SSL. (Par exemple, pour plus d'informations, voir : <http://tomcat.apache.org/tomcat-7.0-doc/ssl-howto.html>).
 - a. Depuis la ligne de commande, accédez au répertoire C:\Program Files\ibm\cognos\tm1_64\bin64\jre\7.0\bin (il s'agit du chemin d'installation par défaut) et exécutez la commande suivante pour générer un fichier appelé .keystore dans votre dossier d'accueil :


```
keytool -genkey -alias tomcat -keyalg RSA
```
 - b. Dans le dossier C:\Program Files\ibm\cognos\tm1_64\tomcat\conf, ajoutez les paramètres de connecteur suivants au fichier server.xml :


```
<Connector SSLEnabled="true" acceptCount="100" clientAuth="false"
  disableUploadTimeout="true" enableLookups="false" maxThreads="25"
  port="8443" keystoreFile="/Users/loiane/.keystore" keystorePass="password"
  protocol="org.apache.coyote.http11.Http11NioProtocol" scheme="https"
  secure="true" sslProtocol="TLS" />
```
 - c. Redémarrez le service IBM Cognos TM1 Application Server.
2. Sur la ligne de commande, lancez la commande suivante pour exporter le fichier de certification du nouveau magasin de clés :


```
keytool -export -alias tomcat -file certfile.cer -keystore C:\Users\Administrator\.keystore
```
3. Sur la ligne de commande, lancez la commande suivante pour importer le fichier de certification dans le JRE utilisé par SPSS Modeler Server :


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

Vous pouvez ensuite utiliser HTTPS et le numéro de port sécurisé SSL pour vous connecter à Cognos TM1.

Configuration des groupes

Un utilisateur authentifié appartient généralement à un ou plusieurs groupes de sécurité et lorsqu'une configuration basée sur des groupes est activée pour SPSS Modeler Server, ces groupes peuvent être utilisés pour autoriser ou refuser la connexion au serveur ou pour personnaliser les paramètres d'option pour la session utilisateur.

La configuration de groupe est prise en charge dans les scénarios suivants :

- Dans une installation *par défaut* où le service SPSS Modeler Server s'exécute sous le compte Système local ou root et où l'utilisateur se connecte avec des données d'identification explicites ou en utilisant la connexion unique (SSO) : dans ce cas, les groupes sont les groupes de sécurité du système d'exploitation de l'utilisateur utilisés pour contrôler l'accès aux fichiers, etc.
- Dans une installation *non-root* où le service SPSS Modeler Server s'exécute sous un compte sans privilèges et où l'utilisateur se connecte via le mode connexion unique (SSO) : dans ce cas, les groupes sont les groupes LDAP associés au principal SSO. Ces groupes sont obtenus auprès du fournisseur de sécurité LDAP dans IBM SPSS Collaboration and Deployment Services ; une configuration supplémentaire est donc requise pour utiliser ce scénario. Pour plus d'informations, voir «Obtention de l'appartenance au groupe de l'utilisateur SSO», à la page 20.

Si aucun de ces scénarios ne s'applique, les groupes de l'utilisateur ne sont pas disponibles et la configuration de groupe n'est pas prise en charge. En particulier, dans une installation *non-root* où le service SPSS Modeler Server s'exécute sous un compte sans privilèges et où l'utilisateur se connecte via un nom d'utilisateur et un mot de passe, les groupes du système d'exploitation ne sont pas disponibles pour le serveur et la configuration de groupe n'est pas prise en charge.

Dans une configuration basée sur les groupes, les paramètres d'option appliqués à une session utilisateur peuvent différer selon l'appartenance au groupe de l'utilisateur. Il s'agit des paramètres côté serveur normalement lus à partir du fichier SPSS Modeler Server `options.cfg` défini de façon identique pour

toutes les sessions. Le fichier `options.cfg` fournit les paramètres par défaut pour toutes les sessions, mais des fichiers de configuration spécifiques au groupe peuvent remplacer un sous-ensemble de paramètres pour des sessions en particulier.

La configuration de groupes permet de contrôler différents paramètres, par exemple :

- Contrôler l'accès à des fichiers ou l'accès DSN
- Contrôler l'utilisation des ressources
- Allouer de l'espace de fichier temporaire

Lorsque l'option `group_configuration` est activée dans `options.cfg`, IBM SPSS Modeler Server vérifie le fichier `groups.cfg` qui contrôle les personnes pouvant se connecter au serveur. La valeur par défaut est `N`. Dans l'exemple de fichier `groups.cfg` suivant, le groupe `Test` se voit refuser l'accès au serveur et le groupe `Fraud` l'obtient avec une configuration spécifiée. L'astérisque octroie l'accès à tous les autres groupes avec la configuration par défaut.

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

Une configuration de groupe spécifique, telle que celle décrite pour `Fraud` ci-dessus, peut restreindre l'accès à des sources de données particulières ou modifier les paramètres de ressource (liés à la répercussion SQL (SQL push back), à l'utilisation de la mémoire, au traitement multitâche, etc.). pour améliorer les performances des membres du groupe.

Contrôle de l'accès DSN par groupe

L'authentification multi-facteur (MFA) exige que les utilisateurs puissent être limités dans l'accès aux DSN ODBC accessibles, en fonction de leur appartenance au groupe.

Pour ce faire, le schéma est similaire au schéma existant de l'accès aux fichiers. Deux paramètres de configuration sont disponibles dans `options.cfg` :

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

Si `sql_data_sources_restricted` a la valeur `Y`, l'utilisateur est limité aux DSN listés dans le chemin associé. Les DSN sont séparés par le séparateur de chemin standard ; (point-virgule) sous Windows et : (deux-points) sous UNIX. Par exemple, sous Windows :

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

Lorsque cette restriction est activée, les résultats obtenus sont les suivants :

- Lorsqu'un utilisateur recherche des sources de données (par exemple, à partir de la boîte de dialogue de connexion ODBC, ou en utilisant l'API PS Session `getServerDataSourceNames`), au lieu de disposer de tous les DSN définis sur le système du serveur, l'utilisateur ne voit que le sous-ensemble de DSN défini dans le chemin `options.cfg`. Notez que le chemin peut contenir des DSN qui ne sont pas définis sur le serveur, et qui sont ignorés. L'utilisateur ne voit pas ces noms.
- Si un utilisateur construit un noeud ODBC (ou tout noeud via une connexion ODBC) qui utilise un script ou une API PS et qu'il spécifie un DSN qui ne figure pas dans le chemin `options.cfg`, le noeud ne s'exécutera pas et l'utilisateur recevra une erreur semblable à **Accès refusé à la source de données : <X>**.

Le chemin de la source de données peut inclure les insertions `PATH`, `GROUP` et `USER` décrites ailleurs pour les chemins d'accès aux fichiers. L'insertion `PATH` permet une construction incrémentielle du chemin en fonction de l'appartenance au groupe de l'utilisateur dans le cas d'une configuration basée sur le groupe. Dans certaines circonstances, il est utile de nommer un DSN en fonction du groupe auquel il appartient.

Suivant l'exemple précédent, si l'accès aux sources de données Fraud est réservé aux membres du groupe Fraud Analysts, le site peut permettre la configuration du groupe et créer une configuration propre au groupe Fraud Analysts contenant au moins la ligne suivante :

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

L'ajout du préfixe PATH dans cet exemple garantit que les membres du groupe Fraud Analysts peuvent toujours accéder à d'autres sources de données accessibles à tous, ou à d'autres groupes dont ils sont membres.

Journal du serveur

IBM SPSS Modeler Server conserve un enregistrement de ces actions importantes dans un fichier appelé `server_logging.log`. Sous UNIX, ce fichier se trouve dans le dossier `log` situé dans le répertoire d'installation. Sous Windows, il se trouve dans : `%ALLUSERSPROFILE%/IBM/SPSS/Modeler Server/<version>/log`.

Les paramètres qui contrôlent la manière dont la journalisation est effectuée dans votre installation figurent dans le fichier `log4cxx.properties`.

Modification de l'emplacement du fichier journal

L'emplacement par défaut du fichier journal est défini dans le fichier `log4cxx.properties` comme suit :

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

Pour modifier l'emplacement du fichier journal, éditez cette entrée.

Activation de la fonction de trace

Dans certains cas, vous pouvez avoir besoin de plus de détails qu'une liste d'informations de base montrant les actions principales. Par exemple, ces détails peuvent être demandés par le personnel d'assistance pour vous aider à identifier un problème. Vous pouvez alors modifier le journal pour qu'il fournisse des informations de trace plus détaillées.

Pour activer la fonction de trace, dans le fichier `log4cxx.properties`, désactivez la ligne `log4j.rootLogger=INFO, MainLog, ConsoleLog` et activez la ligne suivante : `log4j.rootLogger=TRACE, MainLog, TraceLog`

Pour modifier l'emplacement du journal de trace, éditez l'entrée suivante :

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

Modification des options de journalisation

Le fichier `log4cxx.properties` contient les contrôles qui définissent comment différents événements sont journalisés. Ces contrôles sont définis normalement sur **INFO** pour enregistrer les actions dans le fichier journal, ou sur **WARN** pour avertir l'utilisateur d'un problème potentiel. Si vous utilisez le fichier journal pour identifier les erreurs potentielles, vous pouvez également définir certains des contrôles sur **TRACE**.

Contrôle de la taille du fichier journal

Par défaut, la taille du fichier journal continue de croître chaque fois que vous utilisez SPSS Modeler Server. Pour empêcher le journal de devenir trop grand, vous pouvez le paramétrer pour qu'il recommence à zéro chaque jour ou définir une taille limite.

Pour définir le journal de sorte qu'un nouveau journal soit démarré tous les jours, dans le fichier `log4cxx.properties`, utilisez les entrées suivantes :

```
log4j.appender.MainLog=org.apache.log4j.DailyRollingFileAppender  
log4j.appender.MainLog.DatePattern='.'aaaa-MM-jj
```

Sinon, pour définir une taille limite pour le journal (par exemple, 8 Mo), dans le fichier `log4cxx.properties`, utilisez les entrées suivantes :

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

Chapitre 5. 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, la génération et l'optimisation SQL.
- Utilisez un matériel qui répond au minimum aux recommandations données dans la rubrique Chapitre 2, «Recommandations relatives à l'architecture et au matériel», à la page 7.
- Assurez-vous que les paramètres de performance et d'optimisation du client et du serveur sont correctement configurés. Notez que quand SPSS Modeler est connecté à une installation SPSS Modeler Server, les paramètres de performance et d'optimisation remplacent les équivalents du client.
- Concevez les flux afin d'obtenir des performances maximales.

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 Modeler Administration Console fournie avec IBM SPSS Modeler Server. Pour plus d'informations, voir la rubrique «Utilisation de IBM SPSS Modeler Administration Console», à la page 38.

Dans la fenêtre de configuration IBM SPSS Modeler Administration Console, 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 jeux de données et du contenu de vos flux. Pour plus d'informations, voir la rubrique «Performances/Optimisation», à la page 41.

Paramètres de performances et d'optimisation du client

Les paramètres de performance et d'optimisation du client sont disponibles dans l'onglet Options de la boîte de dialogue Propriétés du flux. Pour afficher ces options, choisissez ce qui suit dans le menu du client.

Outils > Propriétés du flux > Options > Optimisation

Vous pouvez utiliser les paramètres d'optimisation pour optimiser la performance du flux. Les paramètres de performances et d'optimisation de IBM SPSS Modeler Server (si ce dernier est utilisé) annulent les paramètres équivalents définis pour le client. Si ces paramètres sont désactivés sur le serveur, le client ne peut pas les activer. Mais s'ils sont activés sur le serveur, le client peut choisir de les désactiver.

Remarque : La modélisation de base de données et l'optimisation SQL requièrent l'activation de la connectivité à IBM SPSS Modeler Server sur l'ordinateur IBM SPSS Modeler. Avec ce paramètre activé, vous pouvez accéder aux algorithmes de la base de données, effectuer le push back de SQL directement depuis IBM SPSS Modeler et accéder à IBM SPSS Modeler Server. Pour vérifier le statut actuel de la licence, choisissez ce qui suit dans le menu IBM SPSS Modeler.

Aide > A propos de > Informations supplémentaires

Si la connectivité est activée, vous voyez l'option **Activation du serveur** dans l'onglet Etat de la licence.

Pour plus d'informations, voir la rubrique «Connexion à IBM SPSS Modeler Server», à la page 15.

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 17.1, consultez 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 IBM SPSS Modeler. Quatre types de réécriture sont disponibles. Vous pouvez sélectionner l'une de ces options ou l'ensemble. 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 IBM SPSS Modeler, mais de telles répercussions contribuent également à réduire le volume des données renvoyées à IBM 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.
- **Optimiser l'expression CLEM** Cette option permet à l'optimiseur de rechercher des expressions CLEM qui peuvent être prétraités avant l'exécution du flux, afin d'accroître la vitesse de traitement. Exemple simple : si vous avez une expression telle que $\log(\text{salary})$, l'optimiseur calcule la valeur de salary et la transmet pour traitement. Cela peut être utilisé pour améliorer les répercussions et les performances IBM SPSS Modeler Server.
- **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 d'exploration de données 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.

- **Utiliser une conversion élargie.** Cette option permet la conversion des données de chaînes à nombres ou de nombres à chaînes, si elles sont stockées dans un format approprié. Par exemple, si les données sont conservées dans la base de données en tant que chaîne, mais contiennent en fait un nombre important, les données peuvent être converties pour être utilisées au moment des répercussions.

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

Utilisation et optimisation de la base de données

Serveur de bases de données Si possible, créez une instance de base de données dédiée à l'exploration de données de façon à ce que le serveur de production ne subisse pas l'impact des requêtes IBM SPSS Modeler. Les instructions SQL générées par IBM 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 d'exploration de données pour leurs produits. Ces extensions permettent d'exécuter des activités d'exploration de données (telles que la création de modèles ou la détermination de scores) sur le serveur de base de données ou sur un serveur dédié distinct. Les fonctions d'exploration des bases de données de IBM SPSS Modeler complètent et étendent ses capacités de génération SQL et offrent ainsi la possibilité de gérer les extensions de base de données propres aux fournisseurs . Dans certains cas, cette approche évite le temps système potentiellement coûteux lié au transfert de données entre IBM 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, voir Chapitre 6, «Optimisation SQL», à la page 61.

Envoi de données basées sur des fichiers

Les données non stockées dans une base de données ne peuvent pas 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 jeux de données intermédiaires issus de la préparation des données ainsi que les résultats du déploiement.

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 IBM 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.

Tableau 2. 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

Chapitre 6. Optimisation SQL

L'une des capacités les plus efficaces de IBM SPSS Modeler réside dans sa capacité à exécuter de nombreuses opérations de préparation des données et d'exploration de données 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 jeux 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 IBM 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 du jeu 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, IBM 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 IBM 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, voir la rubrique «Astuces relatives à l'optimisation de la génération SQL», à la page 65.

Remarque : Gardez à l'esprit les informations suivantes concernant SQL :

- 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 IBM 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 IBM 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 IBM SPSS Modeler, les comparaisons de chaînes des bases de données devraient distinguer les majuscules et les minuscules.
- La modélisation de base de données et l'optimisation SQL requièrent l'activation de la connectivité à IBM SPSS Modeler Server sur l'ordinateur IBM SPSS Modeler. Avec ce paramètre activé, vous pouvez accéder aux algorithmes de la base de données, effectuer la conversion SQL (push back) directement depuis IBM SPSS Modeler et accéder à IBM SPSS Modeler Server. Pour vérifier le statut actuel de la licence, choisissez ce qui suit dans le menu IBM SPSS Modeler.
- Lorsque vous utilisez IBM SPSS Modeler pour générer du code SQL, il est possible que le résultat de la conversion SQL ne soit pas cohérent avec le code IBM SPSS Modeler natif sur certaines plateformes (Linux/zLinux, par exemple) car la virgule flottante est traitée différemment sur les différentes plateformes.

Aide > A propos de > Informations supplémentaires

Si la connectivité est activée, vous voyez l'option **Activation du serveur** dans l'onglet Etat de la licence.

Pour plus d'informations, voir la rubrique «Connexion à IBM SPSS Modeler Server», à la page 15.

Remarque : Lorsque les flux sont exécutés dans une base de données Netezza, les informations relatives à la date et à l'heure sont récupérées dans la base de données. Ces informations peuvent être différentes de l'heure et la date locales ou de l'heure et la date de IBM SPSS Modeler Server, si par exemple, la base de données est hébergée sur un ordinateur situé dans un pays ou un fuseau horaire différent.

Configuration requise pour les bases 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 IBM 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.

Configuration du pilote ODBC

Pour garantir que les informations relatives à l'heure (telles que HH:MM:SS) sont correctement traitées lorsque vous utilisez SQL 2012 sous Windows 8 (systèmes 32 bits), lors de la configuration du pilote de protocole ODBC SQL Server Wire Protocol Driver, sélectionnez les options **Enable Quoted Identifiers** et **Fetch TWFS as Time**.

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 Graphique, 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, voir la rubrique «Astuces relatives à l'optimisation de la génération SQL», à la page 65.

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.** IBM 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.

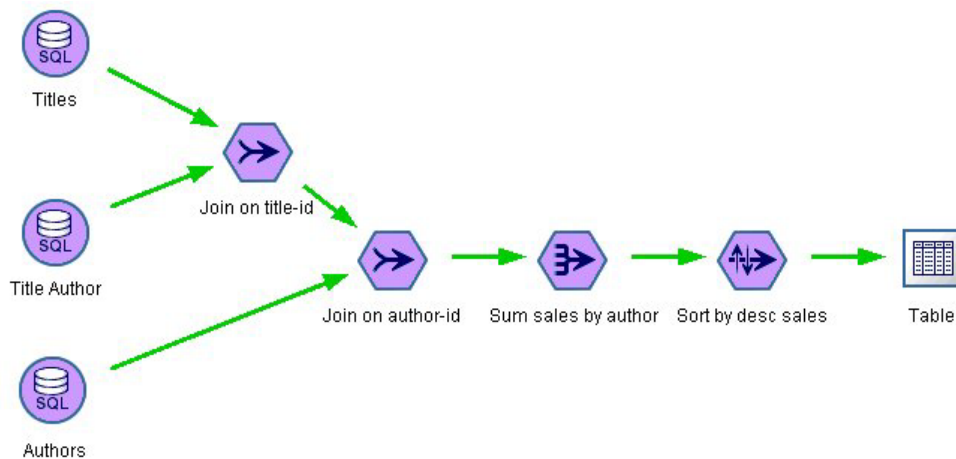


Figure 3. Flux optimisé avec noeuds violets indiquant les répercussions SQL (opérations effectuées dans la base de données)

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 noeud exportation de base de données le flux peut être intégralement exécuté dans la base de données.

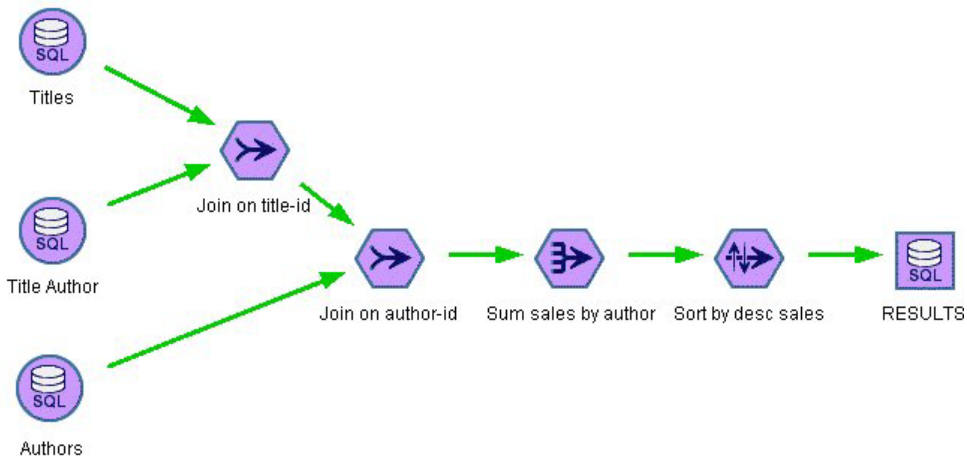


Figure 4. Intégralité du flux exécutée dans la base de données

Configuration de l'optimisation SQL

1. Installez un pilote ODBC et configurez une source de données pour la base de données à utiliser. Pour plus d'informations, voir la rubrique «Accès aux données», à la page 11.
2. Créez un flux utilisant un noeud source pour extraire les données de la base de données.
3. 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

1. Dans le menu Outils, sélectionnez **Propriétés du flux > Options**.
2. 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, voir la rubrique «Paramètres de performances et d'optimisation du client», à la page 57.

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», à la page 41. Si ces paramètres sont désactivés sur le serveur, le client ne peut pas les activer. Mais s'ils sont activés sur le serveur, le client peut choisir de les désactiver.

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.

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.

1. 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, voir la rubrique «Paramètres de performances et d'optimisation du client», à la page 57.
2. Dans l'espace de travail de flux, sélectionnez le noeud ou le flux à prévisualiser.
3. 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.
4. Pour prévisualiser le code SQL généré, choisissez les options suivantes :
Outils > Propriétés du flux > 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, voir la rubrique «Noeuds prenant en charge la génération SQL», à la page 66.

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

1. Sélectionnez l'onglet Paramètres dans le nugget de modèle.
2. Sélectionnez l'option appropriée, **Générer avec/sans prise en charge des valeurs manquantes** ou **Générer SQL pour ce modèle**.
3. Dans le menu du nugget de modèle, sélectionnez :
Fichier > Exporter SQL
4. Enregistrez le fichier.
5. 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 d'exploration de données 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, voir la rubrique «Noeuds prenant en charge la génération SQL», à la page 66.

Expressions CLEM. 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, voir la rubrique «Opérateurs et expressions CLEM prenant en charge la génération SQL», à la page 70.

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.

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, voir la rubrique «Noeuds prenant en charge la génération SQL».

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 IBM SPSS Modeler depuis le gestionnaire de tâches de Windows.

Mise en cache de la 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, voir la rubrique «Configuration de l'optimisation SQL», à la page 64.

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 d'exploration de données 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, voir la rubrique «Prévisualisation du code SQL généré», à la page 65.

Tableau 3. 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, voir la rubrique «Rédaction de requêtes SQL», à la page 73.

Tableau 4. Opérations d'enregistrement

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 à des degrés variables en fonction de la base de données. Voir tableau 5.
Agréger	La prise en charge de la génération SQL pour l'agrégation dépend du type de stockage des données. Voir tableau 6, à la page 68.
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. Prend en charge la génération SQL pour la fusion conditionnelle. 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. Remarque : L'optimisation SQL n'est possible que lorsque les entrées possèdent le même nombre de colonnes.
Distinguer	Un noeud Distinguer (Distinct) avec le mode par défaut Créer un enregistrement composite pour chaque groupe sélectionné ne prend pas en charge l'optimisation SQL.

Tableau 5. Prise en charge de la génération SQL dans le noeud Echantillonner pour l'échantillonnage simple.

Mode	Echantillonner	Taille max	Valeur de départ	DB2 pour z/OS	DB2 pour OS/400	DB2 pour Win/UNIX	Netezza	Oracle	SQL Server	Teradata	
Inclure	Premiers	n/a		Y	Y	Y	Y	Y	Y	Y	
	Tous les	désactivé		Y	Y	Y	Y	Y		Y	
		max		Y	Y	Y	Y	Y		Y	
	% aléatoire	désactivé	désactivé	Y		Y	Y	Y			Y
			activé(e)	Y		Y		Y			
		max	désactivé	Y		Y	Y	Y	Y		Y
activé(e)			Y		Y		Y	Y			

Tableau 5. Prise en charge de la génération SQL dans le noeud Echantillonner pour l'échantillonnage simple (suite).

Mode	Echantillonner	Taille max	Valeur de départ	DB2 pour z/OS	DB2 pour OS/400	DB2 pour Win/UNIX	Netezza	Oracle	SQL Server	Teradata	
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	désactivé	désactivé	Y		Y	Y	Y	Y		Y
			activé(e)	Y		Y		Y			
		max	désactivé	Y		Y	Y	Y	Y		Y
			activé(e)	Y		Y		Y	Y		

Tableau 6. Prise en charge de la génération SQL dans le noeud Agréger.

Stockage	Somme	Moyenne	Min	Max	Ecart-type	Médiane	Comptage	Variance	Centile
Entier	Y	Y	Y	Y	Y		Y	Y	
Réel	Y	Y	Y	Y	Y		Y	Y	
Date			Y	Y			Y		
Heure			Y	Y			Y		
Horodatage			Y	Y			Y		
Chaîne			Y	Y			Y		

Tableau 7. Opérations sur les champs

Noeud prenant en charge la génération SQL	Remarques
Type	Ce noeud prend en charge la génération SQL si le noeud type est instancié, et qu'aucune vérification de 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 indicateurs.
Rechiffrer	
Discrétiser	Prend en charge la génération SQL si la méthode de regroupement par casiers de type « quantile (nombre égal) » est utilisée et que l'option « Lire dans l'onglet Valeurs d'intervalle si disponible » est sélectionnée. Remarque : En raison des différences dans le mode de calcul des limites de casiers (du fait de la nature de la répartition des données dans les champs de casier), vous pouvez constater des écarts dans la sortie du regroupement par casiers lorsque vous comparez l'exécution de flux normale avec les résultats de la conversion SQL. Pour éviter ceci, utilisez la méthode de création de quantiles Nombre d'enregistrements et les quantiles Ajouter au suivant ou Conserver dans l'élément actuel pour obtenir la correspondance la plus proche entre les deux méthodes d'exécution de flux.

Tableau 7. Opérations sur les champs (suite)

Noeud prenant en charge la génération SQL	Remarques
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.
Binariser	
Restructurer	

Tableau 8. Graphiques

Noeud prenant en charge la génération SQL	Remarques
Représentation graphique	La génération SQL est prise en charge pour les types de graphique 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. La génération SQL n'est pas prise en charge pour l'animation dans la Représentation graphique.
Proportion	
Relations	
Evaluation	

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, voir la rubrique «Affichage SQL pour des nuggets de modèle», à la page 65.

Tableau 9. Nuggets de modèle

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 de jeux de données 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 de jeux de données très volumineux.
Réseau de 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 de jeux de données très volumineux.
ACP/Facteur	

Tableau 9. Nuggets de modèle (suite)

Nugget de modèle prenant en charge la génération SQL	Remarques
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 Indicateur.
Ensemble de règles généré	
Discriminant automatique	Si un adaptateur de scoring de type Fonction définie par l'utilisateur (UDF) est activé, ces nuggets prennent en charge la conversion SQL (pushback). En outre, si la Génération SQL pour les cibles continues ou la méthode d'ensemble "La confiance la plus élevée l'emporte" est utilisée, ces nuggets prennent en charge plus de répercussions en aval.
Numérisation automatique	

Tableau 10. Sortie

Noeud prenant en charge la génération SQL	Remarques
Table	Ce noeud 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	
Graphiques	La Répercussion SQL n'est pas prise en charge si l'animation est activée en mode Représentation graphique.
Statistiques	Prend en charge la génération si l'option Corréler n'est pas utilisée.
Rapport	
V. globales	

Tableau 11. 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 de l'exploration de données. Les opérations absentes de ces tableaux ne prennent pas en charge la génération SQL dans la version actuelle.

Tableau 12. Opérateurs

Opérations prenant en charge la génération SQL	Remarques
+	
-	
/	

Tableau 12. Opérateurs (suite)

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

Tableau 13. Opérateurs relationnels

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

Tableau 14. 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	
isaplhacode	
islowercode	
isnumbercode	
isstartstring	
issubstring	
isuppercode	
last	
length	

Tableau 14. Fonctions (suite)

Opérations prenant en charge la génération SQL	Remarques
locchar	
log	
log10	
lowertoupper	
max	
member	
min	
negate	
not	
number	
or	
pi	
real	
rem	
round	
sign	
sin	
sqrt	
string	
strmember	
subscrs	
substring	
substring_between	
uppertolower	
to_string	

Tableau 15. 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 noeud V. globales.
@GLOBAL_SUM	
@GLOBAL_MAX	
@GLOBAL_MEAN	
@GLOBAL_MIN	
@GLOBALSDEV	

Tableau 16. Fonctions Agrégation

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

Utilisation de fonctions SQL dans les expressions CLEM

La fonction @SQLFN peut être utilisée 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 jeux 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 jeu 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;
```

Adaptateur de scoring pour Teradata - Lignes dupliquées

L'Adaptateur de scoring IBM SPSS Modeler Server pour Teradata n'attend pas de lignes identiques dans ses données en entrée. Teradata n'autorise pas l'existence de deux lignes identiques dans une table. Toutefois, les lignes identiques peuvent être créées lors de la jointure de tables, ou lorsque l'utilisateur utilise seulement une partie des champs d'une table comme entrée. Ces lignes dupliquées génèrent un nombre d'enregistrements incorrects après une jointure cartésienne.

Annexe A. 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 IBM 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 IBM SPSS Modeler Server est exécuté avec d'autres paramètres régionaux ou si les paramètres régionaux du shell dans lequel IBM 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 :

- IBM 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 IBM 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 IBM SPSS Modeler Server au début de la connexion à la base de données. IBM SPSS Modeler n'est alors pas en mesure d'optimiser les flux pour Oracle.

Pour vérifier les spécifications des paramètres régionaux et les redéfinir :

1. 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
```

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

```
#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
```

4. Une fois le lien créé, redémarrez IBM SPSS Modeler Server à partir de ce même shell. (IBM 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 IBM SPSS Modeler Server. Comme toute autre variable d'environnement, la variable LANG peut être définie dans le shell, avant le redémarrage de IBM SPSS Modeler Server. Prenons l'exemple de la définition suivante :
`#LANG=en_US.ISO8859-15; export LANG`
- A chaque lancement de IBM 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.

Annexe B. 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 IBM SPSS Modeler Server lorsque l'ordinateur du serveur est redémarré ;
- arrêter et redémarrer manuellement IBM SPSS Modeler Server.
- modifier les droits d'accès sur les fichiers créés par IBM SPSS Modeler Server.
- configurer IBM SPSS Modeler Server de façon à ce qu'il fonctionne avec les pilotes Connect ODBC fournis avec IBM SPSS Modeler Server. Pour plus d'informations, voir la rubrique «IBM SPSS Modeler Server et Data Access Pack», à la page 78.
- Configurer IBM 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 noeud Enterprise View. Pour plus d'informations, voir la rubrique «Configuration d'un pilote pour le noeud Enterprise View», à la page 84.

Scripts

IBM SPSS Modeler Server utilise plusieurs scripts, notamment :

- **modelersrv.sh.** Le script de démarrage manuel de IBM SPSS Modeler Server est situé dans le répertoire d'installation IBM 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 IBM 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.

Tableau 17. Emplacement du script *rc.modeler* en fonction du système d'exploitation.

Système d'exploitation	Emplacement
AIX	<i>/etc/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 IBM SPSS Modeler Server.

Pour configurer le système sur un démarrage et un arrêt automatiques

1. Connectez-vous en tant que racine (root).
2. Passez dans le répertoire d'installation d'IBM SPSS Modeler Server.

3. 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 IBM 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

1. Passez dans le répertoire d'installation d'IBM SPSS Modeler Server.
2. Pour démarrer le serveur, saisissez ce qui suit à l'invite de commande UNIX :

```
./modelersrv.sh start
```
3. 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

1. Arrêter IBM SPSS Modeler Server. (Pour plus d'informations, voir la rubrique «Démarrage et arrêt manuels de IBM SPSS Modeler Server».)
2. Localisez le script approprié. (Pour plus d'informations, voir la rubrique «Scripts», à la page 77.)
3. Ouvrez le script dans un éditeur de texte, apportez vos modifications et enregistrez le fichier.
4. Démarrez IBM 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, voir «Edition de scripts» ci-dessus.) Nous vous recommandons 077 en tant que restriction maximale, de paramétrer *UMASK*. Les paramètres plus restrictifs peuvent engendrer des problèmes en termes de droits d'accès, pour IBM SPSS Modeler Server.

IBM SPSS Modeler Server et Data Access Pack

Pour que vous puissiez utiliser les pilotes ODBC avec IBM SPSS Modeler Server, l'environnement ODBC doit être configuré par *odbc.sh* lors du démarrage du processus IBM 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, voir «Edition de scripts».)

Pour plus d'informations, consultez le site Web de l'assistance technique, à l'adresse <http://www.ibm.com/support>. 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.

Configuration d'ODBC afin qu'il démarre avec IBM SPSS Modeler Server

1. Arrêtez l'hôte IBM SPSS Modeler Server s'il est en cours d'exécution.
2. Téléchargez l'archive TAR compressée correspondant à la plateforme sur laquelle vous avez installé IBM SPSS Modeler Server. Veillez à télécharger les bons pilotes pour la version installée de IBM SPSS Modeler Server. Copiez le fichier à l'emplacement où vous souhaitez installer les pilotes ODBC (par exemple, `/usr/spss/odbc`).
3. Procédez à l'extraction du fichier archive TAR à l'aide de **tar -xvof**.
4. Exécutez le texte du script `setodbcpath.sh` ayant été extrait de l'archive.
5. 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 IBM 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).

6. Enregistrez `odbc.sh`.
7. (Installations IBM 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 fichier, puis de remplacer les chaînes "`LD_LIBRARY_PATH`" dans les nouvelles instructions modifiées `if` et `export` avec `with "LD_LIBRARY_PATH_64"`.

Par exemple, votre fichier `odbc.sh` final sur une installation IBM 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`.

8. 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).
9. Enregistrez le fichier `odbc.ini`.

10. Configurez IBM 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 `SCEMDNAME` :
 `. <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 ; veillez à laisser un espace entre le premier point et le chemin d'accès au fichier.

11. Enregistrez `modelersrv.sh`.

Important : Pour que le pilote SDAP 7.1 fonctionne sur DB2 on z/OS, vous devez accorder l'accès à `SYSIBM.SYSPACKSTMT`.

Pour tester la connexion

1. Redémarrez IBM SPSS Modeler Server.
2. Connectez-vous à IBM SPSS Modeler Server à partir d'un client.
3. Sur le client, ajoutez un noeud source de base de données à l'espace de travail.
4. 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, voir la rubrique «Dépannage de la configuration ODBC», à la page 81.

Configuration d'ODBC afin qu'il démarre avec IBM SPSS Modeler Solution Publisher Runtime

Lorsque vous parvenez à vous connecter à la base de données à partir de IBM 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 démarrage de IBM SPSS Modeler Solution Publisher Runtime.

1. Modifiez le script `modelerrun` dans IBM 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.

2. Enregistrez le fichier du script `modelerrun`.
3. Par défaut, le gestionnaire de pilote DataDirect n'est pas configuré pour que IBM SPSS Modeler Solution Publisher Runtime utilise ODBC sur les systèmes UNIX. Pour configurer le chargement du gestionnaire de pilote DataDirect sous UNIX, saisissez les commandes suivantes (où `sp_install_dir` est le répertoire d'installation de Solution Publisher Runtime) :

```
cd sp_install_dir
rm -f libspssodbc.so
ln -s libspssodbc_datadirect.so libspssodbc.so
```


Configuration d'ODBC afin qu'il démarre avec IBM SPSS Modeler Batch

Aucune configuration du script IBM SPSS Modeler Batch n'est nécessaire pour ODBC, car vous vous connectez à IBM SPSS Modeler Server à partir de IBM SPSS Modeler Batch pour exécuter les flux. Vérifiez que la configuration ODBC de IBM 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

1. Modifiez le fichier `odbc.ini` pour inclure le nouveau nom ou le nom modifié.
2. Testez la connexion comme décrit précédemment dans cette section.

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

Dépannage de la configuration 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.

1. Vérifiez que `$ODBCINI` est défini dans `modelersrv.sh`, soit explicitement dans le script lui-même ou dans le script `odbc.sh` qui est référencé dans `modelersrv.sh`.
2. Dans ce dernier cas, vérifiez que `ODBCINI` pointe vers le chemin d'accès complet du fichier `odbc.ini` que vous avez utilisé pour définir vos sources de données ODBC.
3. Si la spécification du chemin d'accès dans `ODBCINI` est correcte, vérifiez la valeur de `$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 $ODBCINI`
4. Enregistrez puis exécutez `modelersrv.sh`. La valeur de `$ODBCINI` définie dans l'environnement IBM SPSS Modeler Server est écrite dans `stdout` pour la vérification.
5. 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 IBM 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 IBM 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 IBM 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 IBM 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 essaierez 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.

Le pilote indiqué n'a pas pu être chargé

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=
```

1. Vérifiez que l'objet partagé spécifié par le paramètre Pilote existe.
2. Corrigez le chemin d'accès à l'objet partagé s'il est incorrect.
3. 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, voir la rubrique «IBM SPSS Modeler Server et Data Access Pack», à la page 78. Lorsque vous aurez exécuté ce script, la chaîne "ODBCHOME" devrait 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 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 :

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

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

où REPODBC est remplacé par le chemin d'accès à votre répertoire d'installation ODBC.

2. 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).

3. 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:/bigdisk/oracle/product/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 IBM 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
```

4. 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.

5. Corrigez le chemin de bibliothèque dans le script de démarrage IBM SPSS Modeler.
6. Redémarrez le IBM 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.

Tableau 18. Chemin de bibliothèque en fonction du système d'exploitation.

Système d'exploitation	Nom de la variable du chemin de bibliothèque
AIX	LIBPATH
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 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
```

Annexe C. Configuration et exécution de SPSS Modeler Server 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 IBM 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.

Exécution en tant que processus non-racine. L'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. Toutes les sessions utilisateurs démarrées par SPSS Modeler Server utiliseront le même compte UNIX. Ceci signifie que toute donnée lue ou écrite par SPSS Modeler est partagée avec tous les utilisateurs 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. Sans privilèges racine, IBM SPSS Modeler fonctionne de deux manières :

- **Sans base de données de mots de passe privés.** Avec cette méthode, SPSS Modeler utilise la base de donnée de mots de passe UNIX existante, un serveur NIS ou LDAP utilisé généralement pour l'authentification des utilisateurs sur le système UNIX. Pour plus d'informations, voir la rubrique «Configuration en tant qu'utilisateur non racine sans base de données de mots de passe privés».
- **Avec une base de données de mots de passe privés.** Avec cette méthode, SPSS Modeler authentifie les utilisateurs en utilisant une base de donnée de mots de passe privés, distincte de la base de données de mots de passe UNIX, qui est généralement un serveur NIS ou LDAP utilisé pour l'authentification sur le système UNIX. Pour plus d'informations, voir la rubrique «Configuration en tant qu'utilisateur non-racine avec une base de données de mots de passe privés», à la page 86.

Configuration en tant qu'utilisateur non racine sans base de données de mots de passe privés

Pour configurer IBM SPSS Modeler Server en vue d'une exécution sur un compte non-racine sans base de données de mots de passe privés, suivez les étapes ci-dessous :

1. Ouvrez le fichier SPSS Modeler Server *options.cfg* pour l'édition.
2. Affectez la valeur Y à l'option `start_process_as_login_user`.
3. Enregistrez le fichier *options.cfg*, puis fermez-le.

Par défaut, SPSS Modeler Server essaie chaque méthode d'authentification jusqu'à ce que l'une d'elles fonctionne. Toutefois, vous pouvez si vous le souhaitez utiliser l'option `authentication_methods` du fichier *options.cfg* pour configurer le serveur et essayer une seule méthode. Les valeurs possibles pour l'option sont `pasw_modeler`, `gss`, `pam`, `sspi`, `unix` et `windows`.

Notez qu'une exécution en tant que processus non-racine peut nécessiter des mises à jour de configuration. Pour plus d'informations, voir la rubrique «Dépannage des échecs d'authentification de l'utilisateur», à la page 88.

ATTENTION :

N'activez pas le paramètre `start_process_as_login_user`, puis démarrez IBM SPSS Modeler Server en tant que `root`. Ceci permet que, pour tous les utilisateurs connectés au serveur, les processus de serveur s'exécutent en tant que *racine* ; ceci constitue un risque sécuritaire. Notez que le serveur peut s'arrêter automatiquement si vous tentez cette procédure.

Configuration en tant qu'utilisateur non-racine avec une base de données de mots de passe privés

Si vous choisissez d'authentifier les utilisateurs à l'aide d'une base de données de mot de passe privés, toutes les sessions utilisateurs sont démarrées sur le même compte utilisateur non-racine.

Pour configurer IBM 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 IBM SPSS Modeler Server. Ce compte est réservé au processus démon IBM SPSS Modeler Server. Dans cet exemple, appelons-le *modelerserv*.

Lorsque vous créez le compte, notez que :

- Le groupe principal doit être le groupe `<modelerusers>` créé précédemment.
 - Le répertoire personnel peut être le répertoire d'installation de IBM 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 IBM 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, voir la rubrique «Scripts», à la page 77.

- a. 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. Affectez la valeur `Y` à l'option `start_process_as_login_user`.

6. 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 IBM SPSS Modeler. *Remarque* : Ces données propres à IBM 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 IBM 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.

Remarque : Si vous ajoutez de nouveaux utilisateurs à la base de données de mots de passe privée alors que SPSS Modeler Server est en cours d'exécution, vous devez redémarrer SPSS Modeler Server de sorte qu'elle puisse reconnaître les utilisateurs nouvellement définis. Si vous ne procédez pas de la sorte, les connexions échoueront pour chaque nouvel utilisateur ajouté via *pwutil* depuis le dernier démarrage de SPSS Modeler Server.

7. Attribuez, de façon récursive, la propriété du répertoire d'installation de IBM 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 .
```

8. Dans le répertoire de données, créez des sous-répertoires afin que vos utilisateurs IBM 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 SPSS Modeler Server en tant qu'utilisateur non-racine

Pour exécuter SPSS Modeler Server en tant qu'utilisateur non-racine, suivez les étapes ci-dessous :

1. Connectez-vous en utilisant le compte d'utilisateur non-racine créé précédemment.
2. Si vous procédez à l'exécution avec l'option de fichier de configuration *start_process_as_login_user* activée, vous pouvez démarrer, arrêter et vérifier le statut de SPSS Modeler Server. Pour plus d'informations, voir la rubrique «Pour démarrer ou arrêter le serveur sous UNIX, ou vérifier son statut», à la page 27.

Pour accéder à SPSS Modeler Server, les utilisateurs finaux se connectent à partir du logiciel client. Vous devez fournir aux utilisateurs finaux les informations dont ils ont besoin pour se connecter, notamment l'adresse IP ou le nom d'hôte de l'ordinateur serveur.

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

Lors du démarrage de IBM SPSS Modeler Server en tant qu'utilisateur non-racine sous AIX, il est possible que l'erreur *lsattr not found* ressorte sur la console. Cela s'explique par le fait que *lsattr* est situé dans le répertoire */usr/sbin*, mais lorsque IBM 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 IBM SPSS Modeler Server.

Dépannage des échecs d'authentification de l'utilisateur

Selon la configuration du système d'exploitation, il se peut que vous rencontriez des problèmes de connexion à SPSS Modeler Server lors d'une exécution en configuration non-racine. Par exemple, un échec peut se produire si votre système d'exploitation est configuré (à l'aide du fichier */etc/nsswitch.conf* ou d'un fichier similaire) pour vérifier le fichier de mot de passe shadow local, plutôt que d'utiliser un serveur NIS ou LDAP. Ceci se produit car SPSS Modeler Server nécessite un accès en lecture aux fichiers utilisés pour réaliser l'authentification, notamment au fichier */etc/shadow* ou son équivalent (par exemple, */etc/security/passwd* sous AIX), qui stocke les informations de compte utilisateur sécurisées. Toutefois, les autorisations d'accès aux fichiers du système d'exploitation sont généralement définies de telle sorte que le fichier */etc/shadow* est accessible uniquement par l'utilisateur racine. Dans ces circonstances, un processus non-racine ne peut pas lire le fichier */etc/shadow* pour valider les mots de passe utilisateur, ce qui entraîne une erreur d'authentification.

Vous pouvez résoudre ce problème de plusieurs manières :

- Demandez à votre administrateur système de configurer le système d'exploitation de sorte qu'il utilise un serveur NIS ou LDAP pour l'authentification.
- Modifiez les autorisations d'accès aux fichiers protégés, par exemple en octroyant un accès en lecture au fichier */etc/shadow* de sorte que le compte utilisateur local utilisé pour exécuter SPSS Modeler Server puisse accéder au fichier. Alors que cette solution peut sembler inadaptée aux environnements de production, elle peut être temporairement appliquée à un environnement de test afin de vérifier si l'échec de l'authentification est lié à la configuration du système d'exploitation.
- Spécifiez une liste de contrôle d'accès au fichier */etc/shadow* (voir l'exemple donné plus loin dans cette section).
- Exécutez SPSS Modeler Server en tant que racine, afin de permettre aux processus du serveur de lire le fichier */etc/shadow*. *Attention* : Dans ce cas, vérifiez que le fichier *options.cfg* de SPSS Modeler Server contient l'option *start_process_as_login_user*, N afin d'éviter le problème de sécurité expliqué précédemment.

Exemple de liste de contrôle d'accès (ACL)

Sur les systèmes d'exploitation Solaris, vous pouvez utiliser la commande *setfacl* pour créer une liste de contrôle d'accès (consultez votre administrateur système afin de vous assurer que cette commande est disponible sur votre système). L'exemple suivant vous montre comment créer une telle liste de contrôle d'accès pour l'utilisateur *modserv* :

```
# vi /tmp/shadow.acl
user::r--
user:modserv:r--
group:--
mask:r--
other:--

# setfacl -f /tmp/shadow.acl /etc/shadow

# getfacl /etc/shadow

# file: /etc/shadow
# owner: root
# group: sys
user::r--
user:modserv:r--      #effective:r--
group:---             #effective:---
mask:r--
other:---
```

Annexe D. Configuration et exécution de SPSS Modeler Server avec un fichier de mots de passe privés sous Windows

Introduction

Ces instructions fournissent des informations sur l'exécution de IBM SPSS Modeler Server avec un fichier de mots de passe privés sur les systèmes Windows. Avec cette méthode, IBM SPSS Modeler authentifie les utilisateurs en utilisant une base de données de mots de passe privés, distincte de l'authentification système sous Windows.

Configuration d'une base de données de mots de passe privés

Si vous choisissez d'authentifier les utilisateurs à l'aide d'une base de données de mots de passe privés, toutes les sessions utilisateurs sont démarrées sur le même compte utilisateur.

Pour configurer SPSS Modeler Server de cette manière, suivez les étapes ci-dessous :

1. Créez le compte utilisateur sur lequel exécuter SPSS Modeler Server. Ce compte est réservé au processus démon SPSS Modeler Server. Vous devez démarrer le processus démon avec le compte utilisateur figurant dans l'onglet Connexion du service SPSS Modeler Server 17.1. Dans cet exemple, appelons-le *modelerserv*.
2. Editez le fichier d'options du serveur (config/options.cfg) pour définir l'option `start_process_as_login_user` sur Y et 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"`
3. 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. Notez que ces combinaisons sont propres à SPSS Modeler et qu'elles n'ont aucun lien avec les noms d'utilisateur et les mots de passe utilisés pour la connexion sous Windows. 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, vous devez utiliser 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 appelé `modeler` avec le mot de passe `data mining`, vous devez utiliser une invite DOS pour accéder au répertoire d'installation de SPSS Modeler Server, puis taper :

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

Remarque : Vérifiez qu'il n'y a qu'une seule instance de chaque utilisateur dans le fichier ; les doubles empêchent le démarrage de SPSS Modeler Server.

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.

Si vous ajoutez de nouveaux utilisateurs à la base de données de mots de passe privés alors que SPSS Modeler Server est en cours d'exécution, vous devez redémarrer SPSS Modeler Server de sorte qu'il puisse reconnaître les utilisateurs nouvellement définis. Si vous ne procédez pas de la sorte, les connexions échoueront pour chaque nouvel utilisateur ajouté via pwutil depuis le dernier démarrage de SPSS Modeler Server.

4. Donnez à l'utilisateur créé à l'étape 1 le contrôle complet sur le fichier d'options de serveur config\options.cfg et sur le répertoire %ALLUSERSPROFILE%\IBM\SPSS.
5. Dans les services système, arrêtez le service **IBM SPSS Modeler Server** et modifiez **Log on** défini sur **Local System Account** en remplaçant sa valeur par le compte utilisateur créé à l'étape 1, puis redémarrez le service.

Annexe E. Equilibrage de charge avec clusters 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 capacités 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.

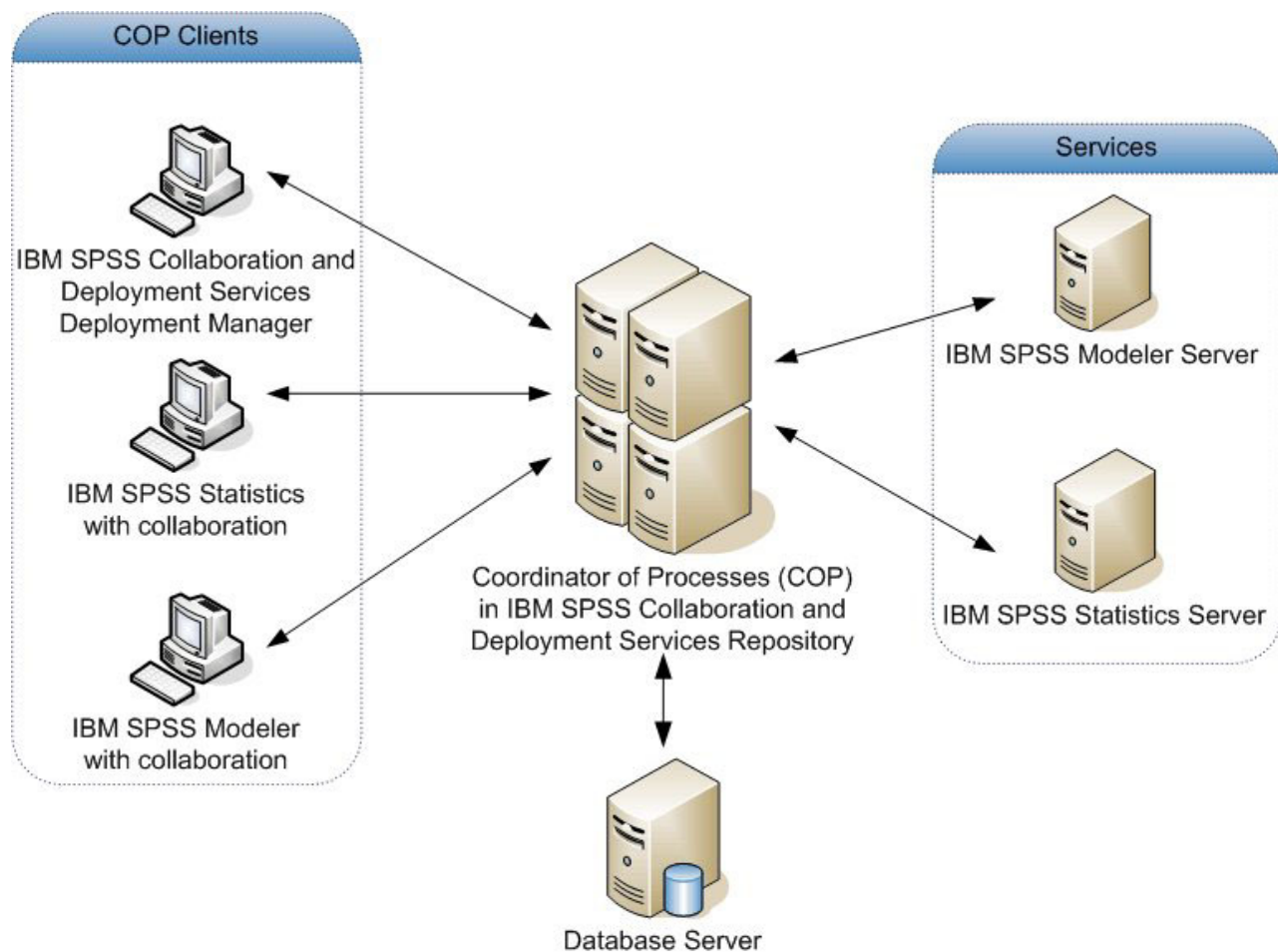


Figure 5. Architecture du Coordinateur de processus

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 cluster de serveurs.

Un cluster de serveurs contient plusieurs serveurs qui sont interchangeable 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 dans la suite de documentation IBM SPSS Collaboration and Deployment Services.

A chaque fois que vous vous connectez à un serveur ou à un cluster de serveurs IBM SPSS Modeler, vous pouvez saisir un serveur manuellement ou rechercher un serveur ou un cluster de serveurs utilisant le Coordinateur de processus. Pour plus d'informations, voir la rubrique «Connexion à IBM SPSS Modeler Server», à la page 15.

Annexe F. Authentification LDAP

Ces instructions fournissent des conseils de base sur la configuration de SPSS Modeler Server sous UNIX pour utiliser l'authentification LDAP, où les identités des utilisateurs qui se connectent au serveur sont stockées dans un annuaire LDAP.

Remarque : Comme condition préalable, le logiciel client LDAP doit être correctement configuré sur le système d'exploitation hôte. Pour plus d'informations, consultez la documentation d'origine du fournisseur.

Habituellement, aucune configuration supplémentaire n'est nécessaire et l'utilisation de LDAP est invisible pour le serveur. Exemples de situations ne nécessitant pas de modifications supplémentaires :

- Les logiciels client LDAP et serveur sont configurés selon RFC 2307.
- L'accès à la base de données passwd (et, le cas échéant, shadow) est redirigé vers LDAP, par exemple dans nsswitch.conf.
- Chaque utilisateur valide de SPSS Modeler Server possède une entrée passwd (et shadow) stockée dans l'annuaire LDAP.
- Le service SPSS Modeler Server est démarré à l'aide du compte superutilisateur (root).

Deux ensembles de situations peuvent exiger la configuration de SPSS Modeler Server spécifiquement pour LDAP :

- Lorsque le service est démarré à l'aide d'un compte autre que root, il ne possède pas nécessairement les droits pour effectuer une authentification à l'aide de la méthode par défaut. Généralement, l'accès à la base de données shadow est restreint.
- Lorsque les utilisateurs ne disposent pas d'entrée passwd (ou shadow) stockée dans le répertoire, c'est-à-dire que leur identité d'utilisateur n'est pas valide pour se connecter au système hôte.

La procédure d'authentification LDAP utilise le sous-système PAM et nécessite qu'un module LDAP PAM existe et soit correctement configuré pour le système d'exploitation hôte. Pour plus d'informations, consultez la documentation d'origine du fournisseur.

Effectuez la procédure suivante pour configurer SPSS Modeler Server afin d'utiliser l'authentification LDAP exclusivement.

Remarque : Il s'agit de la procédure de configuration de base attendue pour ce travail. D'autres options ou paramètres peuvent être nécessaires en fonction du système d'exploitation et de la règle de sécurité locale. Pour plus d'informations, consultez la documentation d'origine.

1. Editez le fichier de configuration de service (`options.cfg`) et ajoutez (ou éditez) la ligne suivante : `authentication_methods, pam`. Cette ligne indique au serveur qu'il doit privilégier l'authentification PAM par rapport à l'authentification par défaut.
2. Indiquez une configuration PAM pour le service SPSS Modeler Server, lequel requiert souvent des privilèges de superutilisateur (root). Le service est identifié par le nom `modelerserver`.
3. Sur un système de type Linux/Unix qui utilise `/etc/pam.d`, créez un fichier dans ce répertoire sous le nom `modelerserver` et ajoutez un contenu similaire à l'exemple suivant :

```
# IBM SPSS Modeler Server
auth required pam_ldap.so
account required pam_ldap.so
password required pam_deny.so
session required pam_deny.so
```

4. Pour les systèmes plus anciens, qui utilisent `/etc/pam.conf`, éditez ce fichier et ajoutez des lignes similaires à l'exemple suivant :

```
# IBM SPSS Modeler Server
modelerserver auth required pam_ldap.so
modelerserver account required pam_ldap.so
modelerserver password required pam_deny.so
modelerserver session required pam_deny.so
```

5. Les noms des modules PAM référencés varient en fonction du système d'exploitation. Confirmez les modules requis pour votre système d'exploitation hôte.

Remarque : Les lignes des étapes 3 et 4 indiquent que SPSS Modeler Server doit faire référence au module LDAP PAM pour l'authentification et la gestion des comptes. Cependant, le changement de mot de passe et la gestion de session n'étant pas pris en charge, ces actions ne sont pas autorisées. Si la gestion des comptes n'est pas requise ou n'est pas appropriée, modifiez la ligne concernée pour autoriser toutes les demandes, comme dans l'exemple suivant :

```
# IBM SPSS Modeler Server
auth required pam_ldap.so
account required pam_permit.so
password required pam_deny.so
session required pam_deny.so
```

Remarques

Ces informations ont été développées pour les produits et services offerts dans le monde.

Le présent document peut contenir des informations ou des références concernant certains produits, logiciels ou services IBM non annoncés dans ce pays. Pour plus de détails, référez-vous aux documents d'annonce disponibles dans votre pays, ou adressez-vous à votre partenaire commercial IBM. Toute référence à un produit, programme ou service IBM n'implique pas que seul ce produit, programme ou service IBM puisse être utilisé. Tout produit, programme ou service fonctionnellement équivalent peut être utilisé s'il n'enfreint aucun droit de propriété intellectuelle d'IBM. Cependant l'utilisateur doit évaluer et vérifier l'utilisation d'un produit, programme ou service non IBM.

IBM peut détenir des brevets ou des demandes de brevet couvrant les produits mentionnés dans le présent document. L'octroi de ce document n'équivaut aucunement à celui d'une licence pour ces brevets. Vous pouvez envoyer par écrit des questions concernant la licence à :

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

Pour le Canada, veuillez adresser votre courrier à :

IBM Director of Commercial Relations
IBM Canada Ltd
3600 Steeles Avenue East
Markham, Ontario
L3R 9Z7 Canada

Pour toute demande au sujet des licences concernant les jeux de caractères codés sur deux octets (DBCS), contactez le service Propriété intellectuelle IBM de votre pays ou adressez vos questions par écrit à :

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

Le paragraphe suivant ne s'applique ni au Royaume-Uni, ni dans aucun pays dans lequel il serait contraire aux lois locales. LE PRESENT DOCUMENT EST LIVRE EN L'ETAT SANS AUCUNE GARANTIE EXPLICITE OU IMPLICITE. IBM DECLINE NOTAMMENT TOUTE RESPONSABILITE RELATIVE A CES INFORMATIONS EN CAS DE CONTREFACON AINSI QU'EN CAS DE DEFAUT D'APTITUDE A L'EXECUTION D'UN TRAVAIL DONNE. Certains états n'autorisent pas l'exclusion de garanties explicites ou implicites lors de certaines transactions, par conséquent, il est possible que cet énoncé ne vous concerne pas.

Ces informations peuvent contenir des erreurs techniques ou des erreurs typographiques. Ce document est mis à jour périodiquement. Chaque nouvelle édition inclut les mises à jour. IBM peut, à tout moment et sans préavis, modifier les produits et logiciels décrits dans ce document.

Toute référence dans ces informations à des sites Web autres qu'IBM est fournie dans un but pratique uniquement et ne sert en aucun cas de recommandation pour ces sites Web. Les éléments figurant sur ces sites Web ne font pas partie des éléments du présent produit IBM et l'utilisation de ces sites relève de votre seule responsabilité.

IBM pourra utiliser ou diffuser, de toute manière qu'elle jugera appropriée et sans aucune obligation à votre égard, tout ou partie des informations qui lui seront fournies.

Les licenciés souhaitant obtenir des informations permettant : (i) l'échange des données entre des logiciels créés de façon indépendante et d'autres logiciels (dont celui-ci), et (ii) l'utilisation mutuelle des données ainsi échangées, doivent adresser leur demande à :

IBM Software Group
ATTN: Licensing
200 W. Madison St.
Chicago, IL; 60606
U.S.A.

Ces informations peuvent être disponibles, soumises à des conditions générales, et dans certains cas payantes.

Le programme sous licence décrit dans le présent document et tous les éléments sous licence disponibles s'y rapportant sont fournis par IBM conformément aux dispositions du Livret Contractuel IBM, des Conditions internationales d'utilisation des Logiciels IBM ou de tout autre contrat équivalent.

Toutes les données sur les performances contenues dans le présent document ont été obtenues dans un environnement contrôlé. Par conséquent, les résultats obtenus dans d'autres environnements d'exploitation peuvent varier de manière significative. Certaines mesures peuvent avoir été effectuées sur des systèmes en cours de développement et il est impossible de garantir que ces mesures seront les mêmes sur les systèmes commercialisés. De plus, certaines mesures peuvent avoir été estimées par extrapolation. Les résultats réels peuvent être différents. Les utilisateurs de ce document doivent vérifier les données applicables à leur environnement spécifique.

les informations concernant les produits autres qu'IBM ont été obtenues auprès des fabricants de ces produits, leurs annonces publiques ou d'autres sources publiques disponibles. IBM n'a pas testé ces produits et ne peut confirmer l'exactitude de leurs performances ni leur compatibilité. Aucune réclamation relative à des produits non IBM ne pourra être reçue par IBM. Les questions sur les capacités de produits autres qu'IBM doivent être adressées aux fabricants de ces produits.

Toutes les déclarations concernant la direction ou les intentions futures d'IBM peuvent être modifiées ou retirées sans avertissement préalable et représentent uniquement des buts et des objectifs.

Ces informations contiennent des exemples de données et de rapports utilisés au cours d'opérations quotidiennes standard. Pour les illustrer le mieux possible, ces exemples contiennent des noms d'individus, d'entreprises, de marques et de produits. Tous ces noms sont fictifs et toute ressemblance avec des noms et des adresses utilisés par une entreprise réelle ne serait que pure coïncidence.

Si vous consultez la version papier de ces informations, il est possible que certaines photographies et illustrations en couleurs n'apparaissent pas.

Marques

IBM, le logo IBM et ibm.com sont des marques d'International Business Machines dans de nombreux pays. Les autres noms de produits et de services peuvent être des marques d'IBM ou d'autres sociétés. La liste actualisée de toutes les marques d'IBM est disponible sur la page Web "Copyright and trademark information" à l'adresse www.ibm.com/legal/copytrade.shtml.

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 des marques déposées de Intel Corporation ou de ses filiales aux Etats-Unis et dans d'autres pays.

Linux est une marque déposée de Linus Torvalds aux Etats-Unis et/ou dans d'autres pays.

Microsoft, Windows, Windows NT et le logo Windows sont des marques commerciales de Microsoft Corporation aux Etats-Unis et/ou dans d'autres pays.

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

Java ainsi que tous les logos et toutes les marques incluant Java sont des marques d'Oracle et/ou de ses sociétés affiliées.

Les autres noms de produits et de services peuvent être des marques d'IBM ou d'autres sociétés.

Index

A

- accès administrateur
 - contrôle d'accès utilisateur 38
 - pour IBM SPSS Modeler Server 38
- accès aux données 11
- administration
 - d'IBM SPSS Modeler Server 37
- ajout de connexions IBM SPSS Modeler Server 21, 22
- allow_modelling_memory_override
 - fichier options.cfg 41
- aperçu
 - génération SQL 65
- architecture
 - composants 7
- arrêt d'IBM SPSS Modeler Server
 - sous UNIX 27
 - sous Windows 27
- authentification 23
- authentification de l'utilisateur 23
- auto.sh (UNIX)
 - emplacement 77

B

- bases de données
 - accès 11

C

- cache_compression
 - fichier options.cfg 41
- chemins 12
- chemsrv.sh (UNIX)
 - emplacement 77
- chiffrement
 - FIPS 45
 - SSL 48
- chiffrement de données SSL
 - activation de IBM SPSS Modeler Server 43
- chiffrement FIPS 45
- client
 - connexion unique 19
- client IBM SPSS Modeler
 - connexion unique 19
- compression du cache 41
- comptes utilisateur
 - droits 24
 - IBM SPSS Modeler Server 23
- configuration du coordinateur de processus
 - pour IBM SPSS Modeler Server 44
- configuration du COP
 - pour IBM SPSS Modeler Server 44
- configuration du pilote ODBC 61
- connexion à IBM SPSS Modeler Server 15
- connexion SSL Cognos BI 52
- Connexion SSL Cognos TM1 52

- connexion unique 15
- connexions
 - à IBM SPSS Modeler Server 15, 21, 22
 - cluster de serveurs 22
- connexions à la base de données
 - fermeture 47
- coordinateur de processus
 - clusters de serveurs 91
 - équilibre de charge 91
- Coordinateur de processus 22
- COP 22
 - clusters de serveurs 91
 - équilibre de charge 91
- cop_enabled
 - fichier options.cfg 44
- cop_host
 - fichier options.cfg 44
- cop_password
 - fichier options.cfg 44
- cop_port_number
 - fichier options.cfg 44
- cop_service_description
 - fichier options.cfg 44
- cop_service_host
 - fichier options.cfg 44
- cop_service_name
 - fichier options.cfg 44
- cop_service_weight
 - fichier options.cfg 44
- cop_update_interval
 - fichier options.cfg 44
- cop_user_name
 - fichier options.cfg 44

D

- Data Access Pack
 - chemins de bibliothèque UNIX 84
 - configuration d'UNIX 78
 - dépannage de ODBC sous UNIX 81
 - ODBC, configuration sous UNIX 78
- data_file_path
 - fichier options.cfg 40
- data_files_restricted
 - fichier options.cfg 40
- DB2
 - optimisation SQL 61, 62
- démarrage automatique du serveur
 - configuration sous UNIX 77
- démarrage d'IBM SPSS Modeler Server
 - sous UNIX 27
 - sous Windows 27
- disques durs 10
- documentation 3
- droits 24
- droits d'accès aux fichiers
 - configuration sous UNIX 78
 - sur IBM SPSS Modeler Server 24

E

- erreur d'exécution de flux 41
- espace disque 10
- Espace disque
 - calcul 10
- exécution de flux multiples 41
- exemples
 - Guide des applications 3
 - présentation 5
- exemples d'application 3
- exploration des bases de données 59
- expressions CLEM
 - génération SQL 70

F

- fichier options.cfg 46
- fichiers de données
 - IBM SPSS Statistics 13
 - importation et exportation 13
- fichiers de données IBM SPSS Statistics
 - importation et exportation 13
- fichiers journaux
 - affichage du code SQL généré 65
 - pour IBM SPSS Modeler Server 55
- fichiers temporaires 10
 - droits d'accès (IBM SPSS Modeler Server) 24
- fonction @SQLFN 73

G

- génération SQL 62
 - activation 64
 - expressions CLEM 70
- génération SQL 61, 66, 74
 - activation pour IBM SPSS Modeler Server 43
 - affichage pour des nuggets de modèle 65
 - astuces 65
 - consignation 65
 - expressions CLEM 65
 - prévisualisation 65
 - réécriture des flux 65
- gestion de la mémoire
 - options d'administration 41
- group_configuration 45

I

- IBM SPSS Analytic Server
 - options de configuration 40
- IBM SPSS Modeler 1
 - documentation 3
- IBM SPSS Modeler Administration Console 38
 - accès administrateur 38

- IBM SPSS Modeler Administration
 - Console (*suite*)
 - accès avec contrôle d'accès utilisateur 38
- IBM SPSS Modeler Server 1
 - accès administrateur 38
 - accès avec contrôle d'accès utilisateur 38
 - administration 38
 - authentification de l'utilisateur 23
 - comptes utilisateur 23
 - configuration du coordinateur de processus 44
 - configuration du COP 44
 - connexion unique 16, 18
 - connexion unique pour sources de données 21
 - création de fichiers 24
 - droits 24
 - ID utilisateur 15
 - informations pour les utilisateurs finaux 23
 - mot de passe 15
 - nom d'hôte 15, 21
 - nom de domaine (Windows) 15
 - numéro de port 15, 21, 40
 - options d'administration 37
 - options de configuration 39
 - processus qui ne répondent pas 28
 - processus serveur 46
 - répertoire temporaire 40
 - résultats différents de ceux de Clementine Client 25
 - surveillance de l'utilisation 46
- ID utilisateur
 - IBM SPSS Modeler Server 15
- io_buffer_size
 - fichier options.cfg 41

K

- Kerberos 45

L

- LDAP
 - authentification 93
- limites du noyau sous UNIX 27
- limites du noyau UNIX 27
- Linux
 - connexion unique 18

M

- max_file_size
 - fichier options.cfg 40
- max_login_attempts
 - fichier options.cfg 40
- max_parallelism
 - fichier options.cfg 41
- max_sessions
 - fichier options.cfg 40
- max_sql_string_length
 - fichier options.cfg 43
- Mémoire 11

- memory_usage
 - fichier options.cfg 41
- messages
 - affichage du code SQL généré 65
- Microsoft SQL Server
 - optimisation SQL 61, 62
- mise en cache de base de données
 - contrôle à partir du fichier options.cfg 46
 - génération SQL 65
- mise en mémoire cache, bases de données 46
- mise en mémoire cache des bases de données 46
- mise en mémoire cache des noeuds
 - écriture dans la base de données 65
 - génération SQL 65
- modélisation
 - gestion de la mémoire 41
- modelling_memory_limit_percentage
 - fichier options.cfg 41

N

- noeuds
 - prenant en charge la génération SQL 66, 74
- noeuds violets
 - optimisation SQL 63
- nom d'hôte
 - IBM SPSS Modeler Server 15, 21
- nom de domaine (Windows)
 - IBM SPSS Modeler Server 15
- noms de fichier
 - Fenêtres 12
 - UNIX 12
- noms de fichier UNC 12
- nuggets de modèle
 - affichage SQL pour 65
- numéro de port
 - IBM SPSS Modeler Server 15, 21, 40

O

- ODBC
 - configuration sous UNIX 78
- opérateurs
 - génération SQL 70
- optimisation
 - génération SQL 61, 62, 64
- option cache_connection 47
- options.cfg 45
- options de configuration
 - Accès aux fichiers de données 40
 - chiffrement de données SSL 43
 - connexions et sessions 40
 - coordinateur de processus 44
 - COP 44
 - d'IBM SPSS Modeler Server 37
 - génération SQL automatique 43
 - gestion de la mémoire 41
 - longueur de chaîne SQL 43
 - numéro de port 40
 - performances et optimisation 41
 - présentation 39, 40
 - réécriture des flux 41

- options de configuration (*suite*)
 - répertoire temporaire 40
 - tentatives de connexion 40
 - traitement parallèle 41
- Oracle
 - optimisation SQL 61, 62, 75

P

- PAM
 - authentification 93
- paramètres de port
 - fichier options.cfg 41
- Paramètres de port du serveur
 - fichier options.cfg 41
- paramètres du pare-feu
 - fichier options.cfg 41
- password
 - IBM SPSS Modeler Server 15
- Performances
 - d'IBM SPSS Modeler Server 57
- port_number
 - fichier options.cfg 40
- processeurs 9
 - Multiple 41
- processus, ne répondant pas 28
- processus zombi, IBM SPSS Modeler Server 28
- program_file_path
 - fichier options.cfg 40
- program_files_restricted
 - fichier options.cfg 40

R

- RAM 11
- rc.modeler (UNIX)
 - emplacement 77
- recherche de connexions dans COP 22
- recommandations matérielles
 - pour IBM SPSS Modeler Server 9
- réécriture des flux 65
- répercussions 61, 62, 64, 66, 74
 - expressions CLEM 70
- répercussions SQL. Consultez également Génération SQL 61
- répertoire temporaire
 - pour IBM SPSS Modeler Server 40
- résultats
 - arrondi 25
 - différences entre Clementine Client et Clementine Server 25
 - ordre des enregistrements 25

S

- Scripts UNIX
 - auto.sh 77
 - modelersrv.sh 77
 - modification 78
 - rc.modeler 77
- sécurité
 - configuration de la création de fichier sous UNIX 78
 - création de fichiers 24
 - SSL 48

- server_logging.log 55
- serveur
 - ajout de connexions 21
 - connexion 15
 - connexion unique 16, 18
 - recherche de serveurs dans COP 22
- serveurs de base de données 59
- shell UNIX 45
- sources de données
 - connexion unique 21
- sources de données ODBC
 - et UNIX 78
 - scripts ODBC et UNIX 78
- SQL
 - affichage pour des nuggets de modèle 65
 - noms de colonne en double 73
 - optimisation d'Oracle 75
 - prévisualisation générée 65
 - requêtes 73
- sql_generation_enabled
 - fichier options.cfg 43
- SQL Server
 - optimisation SQL 61, 62
- SSL 48
 - connexion Cognos BI 52
 - Connexion SSL Cognos TM1 52
 - présentation 48
 - sécurisation des communications 48
- SSL (Secure Sockets Layer) 48
- ssl_certificate_file
 - fichier options.cfg 43
- ssl_enabled
 - fichier options.cfg 43
- ssl_private_key_file
 - fichier options.cfg 43
- ssl_private_key_password
 - fichier options.cfg 43
- statut
 - d'IBM SPSS Modeler Server sous UNIX 27
 - d'IBM SPSS Modeler Server sous Windows 27
- stream_rewriting_enabled
 - fichier options.cfg 41
- systèmes d'exploitation
 - 64 bits 9
- systèmes d'exploitation 64 bits 9

T

- technologie d'accès aux données IBM
 - SPSS Statistics 11
- temp_directory
 - fichier options.cfg 40
- traitement parallèle
 - contrôle 41

U

- UNIX
 - authentification de l'utilisateur 23
 - chemins de bibliothèque 84
 - configuration des droits d'accès aux fichiers 78
 - connexion unique 18

