

Guide sommaire de IBM SPSS Statistics 19



Note: Before using this information and the product it supports, read the general information under Notices sur p. 167.

This document contains proprietary information of SPSS Inc, an IBM Company. It is provided under a license agreement and is protected by copyright law. The information contained in this publication does not include any product warranties, and any statements provided in this manual should not be interpreted as such.

When you send information to IBM or SPSS, you grant IBM and SPSS a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

© Copyright SPSS Inc. 1989, 2010.

Préface

Le manuel *Guide sommaire de IBM SPSS Statistics 19* propose un ensemble de didacticiels conçus pour vous familiariser avec les différents composants de IBM® SPSS® Statistics. Ce guide est exploitable sous toutes les versions de système d'exploitation du logiciel, notamment : Windows, Macintosh et Linux. Vous pouvez suivre les exercices dans l'ordre ou consulter les rubriques pour lesquelles vous souhaitez obtenir des informations supplémentaires. Vous pouvez consulter ce guide en complément ou en remplacement du didacticiel en ligne fourni avec le système central SPSS Statistics ou ignorer le didacticiel en ligne et commencer directement avec les didacticiels fournis ici.

IBM SPSS Statistics 19

IBM® SPSS® Statistics 19 est un système complet d'analyse de données. SPSS Statistics peut utiliser les données de presque tout type de fichier pour générer des rapports mis en tableau, des diagrammes de distributions et de tendances, des statistiques descriptives et des analyses statistiques complexes.

Grâce à SPSS Statistics, l'analyse statistique est plus accessible aux débutants et plus pratique pour les utilisateurs chevronnés. Grâce à des menus et des sélections de boîte de dialogue simples, vous pouvez exécuter des analyses complexes sans saisir une seule ligne de syntaxe de commande. L'éditeur de données offre une fonctionnalité de type tableur simple et efficace pour la saisie des données et la lecture du fichier de travail.

Ressources Internet

Le site Web de SPSS Inc. (<http://www.spss.com>) fournit des réponses aux questions fréquemment posées, un accès aux fichiers de données et d'autres informations utiles.

En outre, le groupe de discussion SPSS USENET (non patronné par SPSS Inc.) est ouvert à toutes les personnes intéressées. L'adresse du groupe USENET est *comp.soft-sys.stat.spss*.

Vous pouvez également vous inscrire sur une liste de messages électroniques reliée par passerelle au groupe USENET. Pour vous inscrire, envoyez un message électronique à l'adresse *listserv@uga.cc.uga.edu*. Le texte du message électronique doit être : *subscribe SPSSX-L prénom nom*. Vous pouvez ensuite envoyer des messages vers la liste. Pour ce faire, envoyez vos messages à l'adresse *listserv@uga.cc.uga.edu*.

Documents supplémentaires

Le *module supplémentaire de procédures statistiques IBM SPSS Statistics*, créé par Marija Norušis, a été publié par Prentice Hall. Il contient des présentations des procédures du système IBM® SPSS® Statistics de base, ainsi que des options de régression logistique et des modèles linéaires généraux. Le *module supplémentaire de procédures statistiques avancées IBM SPSS Statistics a*

également été publié par Prentice Hall. Il inclut les présentations des procédures dans les modules des modèles avancés et de régression.

Options IBM SPSS Statistics

Les options suivantes sont fournies comme améliorations complémentaires du système de base complet IBM® SPSS® Statistics (pas de la version Student) :

Statistics Base vous offre une large gamme de procédures statistiques permettant d'effectuer des analyses et des rapports de base, comprenant des effectifs, des tableaux croisés et des statistiques descriptives, les cubes OLAP et des rapports du livre de code. Il fournit aussi une grande variété de techniques de réduction de dimension, classification et segmentation telles que les analyses factorielles, les classifications, les analyses du voisin le plus proche et de fonction discriminante. En outre, SPSS Statistics Base offre une large variété d'algorithmes pour la comparaison des moyennes et des techniques prédictives telles que le test-t, l'analyse de la variance, la régression linéaire et la régression ordinale.

Le module **Statistiques avancées** décrit les techniques souvent utilisées dans la recherche biomédicale et expérimentale avancée. Il inclut des procédures pour les modèles linéaires généraux (GLM), les modèles mixtes linéaires, l'analyse des composantes de variance, l'analyse log-linéaire, la régression ordinale, la durée de vie actuarielle, l'analyse de survie de Kaplan-Meier, et la régression de Cox de base et étendue.

L'**amorce** est une méthode consistant à dériver des estimations robustes des erreurs standard et des intervalles de confiance pour des estimations telles que la moyenne, la médiane, le calcul de la proportion, l'odds ratio, le coefficient de corrélation ou de régression.

Le module **Modalités** exécute des procédures de codage optimal comme l'analyse des correspondances.

Le module **Echantillons complexes** permet aux chercheurs chargés d'effectuer des enquêtes (notamment d'opinion), des études de marché, ou des études dans le domaine de la santé, ainsi qu'aux spécialistes des sciences sociales qui utilisent une méthodologie d'étude fondée sur les échantillons, d'incorporer leurs propres plans d'échantillonnage complexes dans l'analyse des données.

Conjoint offre une manière réaliste de mesurer la façon dont les attributs du produit individuel affectent les préférences des consommateurs et des citoyens. Avec Conjoint, il est possible de mesurer facilement l'effet de compromis de chaque attribut de produits dans le contexte d'un ensemble d'attributs de produits—comme le font certains consommateurs lorsqu'ils décident de ce qu'ils vont acheter.

Le module **Tableaux personnalisés** crée toute une gamme de rapports en tableau de qualité présentation, y compris des tableaux croisés complexes et les affichages de données de réponses multiples.

Préparation des données fournit un cliché visuel rapide de vos données. Il permet d'appliquer des règles de validation qui identifient les valeurs de données non valides. Vous pouvez créer des règles qui repèrent les valeurs hors plage, les valeurs manquantes et les valeurs vides. Vous pouvez également enregistrer des variables qui enregistrent les violations de règles individuelles et

le nombre total de violations de règles par observation. Un ensemble limité de règles prédéfinies que vous pouvez copier ou modifier vous est fourni.

Arbre de décision crée un modèle d'arbre de segmentation. Elle classe les observations en groupes ou estime les valeurs d'une variable (cible) dépendante à partir des valeurs de variables (explicatives) indépendantes. Cette procédure fournit des outils de validation pour les analyses de classification d'exploration et de confirmation.

Direct Marketing permet aux organisations de rendre leurs programmes de marketing aussi efficaces que possible, grâce à des techniques conçues spécialement pour le marketing direct.

Le module **Tests exacts (Exact Tests™)** calcule les valeurs p exactes pour les tests statistiques lorsque de petits échantillons ou des échantillons distribués de façon très inégale risquent de fausser la précision des tests habituels. Cette option n'est disponible que sous les systèmes d'exploitation Windows.

Le module **Prévision** effectue des prévisions et des analyses de séries chronologiques complètes avec plusieurs modèles d'ajustement aux courbes, des modèles de lissage et des méthodes d'estimation des fonctions autorégressives.

Le module **Valeurs manquantes** décrit les types des données manquantes, évalue les moyennes et d'autres statistiques, et affecte des valeurs aux observations manquantes.

Le module **Réseaux neuronaux (Neural Networks)** permet de prendre des décisions commerciales en prévoyant la demande d'un produit en fonction du prix et d'autres variables, ou en catégorisant les clients en fonction des habitudes d'achat et des caractéristiques démographiques. Les réseaux neuronaux sont des outils de modélisation de données non linéaires. Ils permettent de modéliser les relations complexes entre les entrées et les résultats, ou de rechercher des modèles dans les données.

Le module **Régression** fournit des techniques d'analyse des données non adaptées aux modèles statistiques linéaires classiques. Il contient des procédures pour les modèles de choix binaire, la régression logistique, la pondération estimée, la régression par les doubles moindres carrés et la régression non linéaire générale.

Amos™ (analyse de structures de moments) utilise la modélisation d'équation structurelle pour confirmer et expliquer des modèles conceptuels qui impliquent l'attitude, les perceptions et d'autres facteurs qui entraînent un comportement.

Séminaires de formation

SPSS Inc. propose des séminaires de formation, publics et sur site pour IBM® SPSS® Statistics. Tous les séminaires font appel à des ateliers de travaux pratiques et auront lieu régulièrement dans les grandes villes des Etats-Unis et d'Europe. Pour plus d'informations sur ces séminaires, contactez votre bureau local dont les coordonnées sont indiquées sur le site Web de SPSS Inc. à l'adresse suivante : <http://www.spss.com/worldwide>.

Support technique

Un support 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 SPSS Inc. ou l'installation dans l'un des environnements matériels pris

en charge. Pour contacter l'assistance technique, consultez le site Web SPSS Inc. à l'adresse <http://support.spss.com>, ou recherchez votre représentant local à la page <http://support.spss.com/default.asp?refpage=contactus.asp> Votre nom, celui de votre société, ainsi que votre contrat d'assistance vous seront demandés.

Version Student de IBM SPSS Statistics 19

La version Student de IBM® SPSS® Statistics 19 est une version performante, bien que limitée, de SPSS Statistics.

Capacité

La version Student contient de nombreux outils importants d'analyse des données de IBM® SPSS® Statistics, notamment :

- Editeur de données de type tableur pour la saisie, la modification et l'affichage des fichiers de données.
- Procédures statistiques, notamment tests t , analyse des variances et tabulations croisées.
- Diagrammes interactifs qui vous permettent de modifier ou d'ajouter des éléments et des variables de manière dynamique. Les modifications apparaissent dès que vous les effectuez.
- Diagrammes standard haute résolution permettant d'obtenir une gamme complète de diagrammes et de tableaux de présentation et analytiques.

Limites

Créée à des fins pédagogiques, la version Student ne doit être utilisée que par des étudiants et des instructeurs dans un contexte de formation. Les restrictions suivantes s'appliquent à la version Student de IBM® SPSS® Statistics 19 :

- Les fichiers de données ne peuvent pas contenir plus de 50 variables.
- Les fichiers de données ne peuvent pas contenir plus de 1 500 observations. Les modules SPSS Statistics complémentaires (tels que les modèles de régression ou les statistiques avancées) ne peuvent pas être utilisés avec la version Student.
- La syntaxe de commande SPSS Statistics n'est pas disponible. En effet, contrairement à la version complète de IBM® SPSS® Statistics, la version Student ne permet pas de répéter une analyse en enregistrant une série de commandes dans une syntaxe ou un fichier de tâche.
- L'utilisateur ne peut ni écrire des scripts, ni automatiser des tâches. En effet, contrairement à la version complète de SPSS Statistics, la version Student ne permet pas de créer des scripts afin d'automatiser les tâches répétitives.

Support technique pour les étudiants

Pour bénéficier du support technique, les étudiants peuvent s'adresser à leurs instructeurs ou à l'équipe du support technique local dont les coordonnées leur seront fournies par leurs instructeurs. Le support technique concernant la version Student de IBM® SPSS® Statistics 19 n'est fourni qu'aux *instructeurs utilisant le système dans leurs classes*.

Avant de demander de l'aide à votre instructeur, notez les informations mentionnées ci-après. Sans ces informations, votre instructeur sera dans l'impossibilité de vous aider :

- Type d'ordinateur utilisé, ainsi que quantité de RAM et d'espace disque disponible.
- Système d'exploitation de votre ordinateur.
- Description précise de l'incident et de la tâche que vous effectuiez au moment où il est survenu. Si possible, essayez de reproduire le problème avec l'un des fichiers d'échantillons fournis avec le programme.
- Intitulé exact des messages d'erreur ou d'avertissement qui apparaissent.
- Solutions que vous avez personnellement tentées pour résoudre le problème.

Support technique pour les instructeurs

Les instructeurs qui utilisent la version Student lors de leurs cours peuvent contacter le support technique afin d'obtenir de l'aide. Aux Etats-Unis et au Canada, contactez le support technique au (312) 651-3410, ou rendez-vous sur <http://support.spss.com>.

Dans les autres pays, les instructeurs doivent contacter le bureau local dont les coordonnées sont indiquées sur le site Web à l'adresse <http://www.spss.com/worldwide>.

Contenu

1	<i>Introduction</i>	1
	Fichiers d'exemple	1
	Ouverture d'un fichier de données	2
	Exécution d'une analyse	3
	Affichage des résultats	7
	Création de diagrammes	8
2	<i>Lecture des données</i>	11
	Structure de base d'un fichier de données IBM SPSS Statistics	11
	Lecture des fichiers de données IBM SPSS Statistics	12
	Lecture de données à partir de feuilles de calcul	12
	Lecture de données à partir d'une base de données	14
	Lecture de données à partir d'un fichier texte	20
3	<i>Utilisation de l'éditeur de données</i>	28
	Saisie de données numériques	28
	Saisie de données chaîne	31
	Définition de données	33
	Ajout des étiquettes de variables	33
	Modification du type et du format de variable	34
	Ajout d'étiquettes de valeur pour les variables numériques	34
	Ajout d'étiquettes de valeur pour les variables chaîne	36
	Utilisation des étiquettes de valeurs pour la saisie de données	37
	Gestion des données manquantes	38
	Valeurs manquantes des variables numériques	39
	Valeurs manquantes d'une variable chaîne	41
	Copier-coller des attributs de variable	42
	Définition des propriétés des variables qualitatives	45

4	<i>Utilisation des sources de données multiples</i>	51
	Manipulation de base de plusieurs sources de données	52
	Utilisation de plusieurs ensembles de données dans une syntaxe de commande	54
	Copie et collage d'informations entre les ensembles de données	54
	Attribution d'un nouveau nom aux ensembles de données	55
	Suppression de plusieurs ensembles de données	55
5	<i>Examen des statistiques récapitulatives pour chaque variable</i>	56
	Niveau de mesure.	56
	Mesures récapitulatives pour données qualitatives	56
	Diagrammes pour données qualitatives	58
	Mesures récapitulatives pour variables d'échelle	60
	Histogrammes pour variables d'échelle	62
6	<i>Création et modification de diagrammes</i>	64
	Création de diagrammes - Notions de base.	64
	Utilisation de la galerie Générateur de diagrammes.	65
	Définition des variables et des statistiques	66
	Ajout de texte	69
	Création du diagramme	70
	Modification de diagrammes – Notions de base	70
	Sélection d'éléments de diagramme	71
	Utilisation de la fenêtre Propriétés.	72
	Changement de la couleur des bâtons.	73
	Formatage des valeurs des étiquettes de graduation	74
	Modification de texte.	76
	Affichage des étiquettes des valeurs de données	77
	Utilisation de modèles	78
	Définition des options de diagramme	83
7	<i>Utilisation du résultat</i>	87
	Utilisation du Viewer.	87

Utilisation de l'éditeur de tableau pivotant	89
Accès aux définitions de résultats	89
Tableaux pivotants	90
Création et affichage de strates	92
Modification de tableaux	94
Masquage de lignes et de colonnes	95
Modification des formats d'affichage des données	95
Modèles de tableaux	97
Utilisation de formats prédéfinis	97
Personnalisation des styles Modèles de tableau	98
Modification des formats par défaut d'un tableau	101
Personnalisation des paramètres de l'affichage initial	102
Affichage d'étiquettes de variable et de valeur	104
Utilisation de résultats dans d'autres applications	105
Collage des résultats sous forme de tableaux Word	106
Collage des résultats sous forme de texte	107
Exportation des résultats vers un fichier Microsoft Word, PowerPoint ou Excel	108
Exportation des résultats vers PDF	116
Exportation des résultats vers HTML	119

8 Utilisation de la syntaxe 120

Collage d'une syntaxe	120
Modification d'une syntaxe	122
Ouverture et exécution d'un fichier de syntaxe	124
Présentation du Panneau d'erreur	124
Utilisation des points d'arrêt	125

9 Modification des valeurs de données 127

Création d'une variable qualitative à partir d'une variable d'échelle	127
Calcul de nouvelles variables	133
Utilisation de fonctions dans les expressions	135
Utilisation d'expressions conditionnelles	137
Utilisation des dates et heures	139
Calcul du temps écoulé entre deux dates	140
Ajout d'une durée à une date	143

10 Tri et sélection de données 147

Tri des données	147
Traitement d'un fichier scindé	148
Tri d'observations pour le traitement d'un fichier scindé	149
Activation et désactivation du traitement d'un fichier scindé	150
Sélection de sous-groupes d'observations	150
Sélectionner des sous-ensembles d'observations à l'aide d'une expression conditionnelle	151
Sélectionner un échantillon aléatoire	152
Sélection d'un intervalle de temps ou d'observations	153
Traitement des observations exclues	154
Etat de la sélection d'une observation	155

Annexes

A Fichiers d'exemple	156
-----------------------------	------------

B Notices	167
------------------	------------

Index	169
--------------	------------

Introduction

Ce guide met à votre disposition un ensemble de didacticiels vous permettant d'effectuer des analyses utiles de vos données. Vous pouvez suivre les exercices dans l'ordre ou consulter les rubriques pour lesquelles vous souhaitez obtenir des informations supplémentaires.

Ce chapitre vous présente les fonctions de base et décrit une session normale. Nous allons récupérer un fichier de données IBM® SPSS® Statistics préalablement défini, puis créer un récapitulatif statistique simple et un diagramme.

Des instructions plus détaillées sur de nombreux sujets abordés dans ce chapitre seront données dans les chapitres suivants. Nous espérons vous offrir une structure essentielle pour une bonne compréhension des derniers didacticiels.

Fichiers d'exemple

La plupart des exemples présentés ici utilisent le fichier de données *demo.sav*. Ce fichier de données est une enquête factice menée auprès de plusieurs milliers de personnes, contenant des informations démographiques et relatives à la consommation.

Si vous utilisez la version Student, votre version du fichier *demo.sav* est un échantillon représentatif du fichier de données d'origine, qui a été réduit pour respecter la limite de 1 500 observations. Les résultats obtenus avec ce fichier de données seront différents des résultats présentés ici.

Les fichiers d'exemple installés avec le produit figurent dans le sous-répertoire *Echantillons* du répertoire d'installation. Il existe un dossier distinct au sein du sous-répertoire *Echantillons* pour chacune des langues suivantes : Anglais, Français, Allemand, Italien, Japonais, Coréen, Polonais, Russe, Chinois simplifié, Espagnol et Chinois traditionnel.

Seuls quelques fichiers d'exemples sont disponibles dans toutes les langues. Si un fichier d'exemple n'est pas disponible dans une langue, le dossier de langue contient la version anglaise du fichier d'exemple.

Ouverture d'un fichier de données

Ouvrir un fichier de données :

- ▶ A partir des menus, sélectionnez :
Fichier > Ouvrir > Données

Vous pouvez également utiliser le bouton Ouvrir fichier de la barre d'outils.

Figure 1-1

Bouton Ouvrir fichier de la barre d'outils



Une boîte de dialogue d'ouverture des fichiers apparaît.

Par défaut, les fichiers de données IBM® SPSS® Statistics (extension *.sav*) apparaissent.

Cet exemple utilise le fichier *demo.sav*.

Figure 1-2

Fichier *demo.sav* dans Data Editor

	age	marital	address	income	inccat	car
13	31	Marital status	9	40,00	2,00	21,30
14	42	0	8	137,00	4,00	68,90
15	35	0	8	70,00	3,00	34,10
16	52	1	24	159,00	4,00	78,90
17	21	1	1	37,00	2,00	18,60
18	32	0	0	28,00	2,00	13,70
19	42	0	9	109,00	4,00	54,70
20	40	1	12	117,00	4,00	58,30
21	30	0	3	23,00	1,00	11,80
22	48	0	14	21,00	1,00	9,50

Le fichier de données apparaît dans l'Editeur de données. Dans Data Editor, si vous placez le curseur de la souris sur le nom d'une variable (en-tête de colonne), une étiquette de variable plus détaillée apparaît, si elle a été définie pour cette variable.

Par défaut, les valeurs de données réelles sont affichées. Pour afficher les étiquettes :

- ▶ A partir des menus, sélectionnez :
Affichage > Etiquettes de valeurs

Vous pouvez également utiliser le bouton Etiquettes de valeurs dans la barre d'outils.

Figure 1-3
Bouton Etiquettes de valeurs



Des étiquettes de valeurs descriptives s'affichent pour faciliter l'interprétation des réponses.

Figure 1-4
Étiquettes de valeurs affichées dans l'éditeur de données

The screenshot shows the 'Data Editor' window for a file named 'demo.sav'. The 'age' variable is selected in the top-left corner, and its value '40' is entered in the adjacent field. The main table displays data for 11 rows across several variables: 'age', 'marital', 'address', 'income', 'inccat', 'car', and 'ca'. The 'marital' variable has descriptive labels: 'Marrié(e)', 'Non marri', and 'Inf à \$25'. The 'inccat' variable has labels like '\$50 - \$7', '\$75 - \$1', and 'Inf à \$25'. The 'ca' variable has labels 'Econ' and 'St'.

	age	marital	address	income	inccat	car	ca
1	55	Marrié(e)	12	72,00	\$50 - \$7	36,20	
2	56	Non marri	29	153,00	\$75 - \$1	76,90	
3	28	Marrié(e)	9	28,00	\$25 - \$4	13,70	Econ
4	24	Marrié(e)	4	26,00	\$25 - \$4	12,50	Econ
5	25	Non marri	2	23,00	Inf à \$25	11,30	Econ
6	45	Marrié(e)	9	76,00	\$75 - \$1	37,20	
7	42	Non marri	19	40,00	\$25 - \$4	19,80	St
8	35	Non marri	15	57,00	\$50 - \$7	28,20	St
9	46	Non marri	26	24,00	Inf à \$25	12,20	Econ
10	34	Marrié(e)	0	89,00	\$75 - \$1	46,10	
11	55	Marrié(e)	17	72,00	\$50 - \$7	35,50	

Exécution d'une analyse

Si vous disposez de modules complémentaires, le menu Analyse contient une liste de modalités d'analyse de rapport et d'analyse statistique.

Nous commencerons par créer un simple tableau des effectifs. Cette exemple nécessite l'option Statistiques de base.

- ▶ A partir des menus, sélectionnez :
Analyse > Statistiques descriptives > Effectifs

La boîte de dialogue Effectifs s’affiche.

Figure 1-5
Boîte de dialogue Effectifs

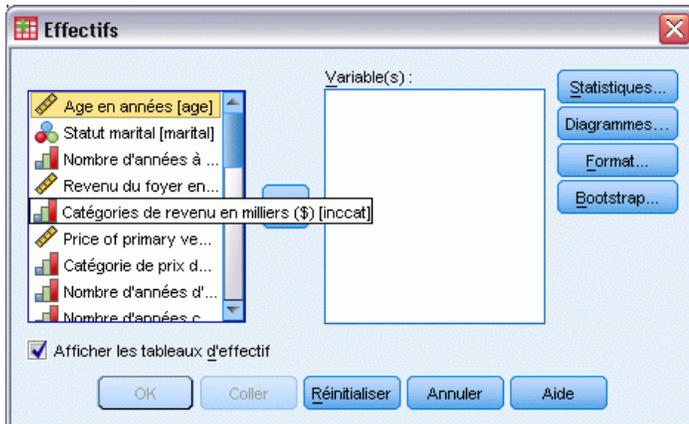


Une icône à côté de chaque variable fournit des informations sur le type de données et le niveau de mesure.

Niveau de mesure	Le type de données			
	Numérique	Chaîne	Date	Heure
Echelle (continue).		n/a		
Ordinales				
Nominales				

- Cliquez sur la variable *Catégories de revenu en milliers [inccat]*.

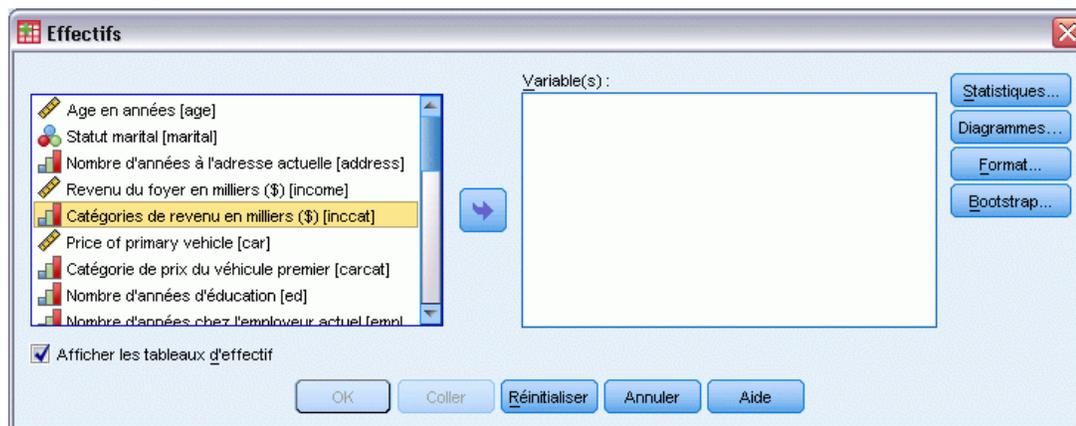
Figure 1-6
Noms et étiquettes de variable dans la boîte de dialogue Effectifs



Si l'étiquette et/ou le nom de la variable apparaît tronqué dans la liste, l'étiquette ou le nom complet est affiché lorsque vous placez le curseur dessus. Le nom de variable *inccat* est affiché entre crochets à la suite de l'étiquette de variable descriptive. *Catégories de revenu en milliers* est l'étiquette de variable. S'il n'y avait pas d'étiquette de variable, seul le nom de la variable apparaîtrait dans la zone de liste.

A l'image des fenêtres, vous pouvez redimensionner les boîtes de dialogue en cliquant sur les bords externes ou sur les angles, et en les faisant glisser. Par exemple, si vous élargissez la boîte de dialogue, vous élargissez également les listes de variables.

Figure 1-7
Boîte de dialogue redimensionnée

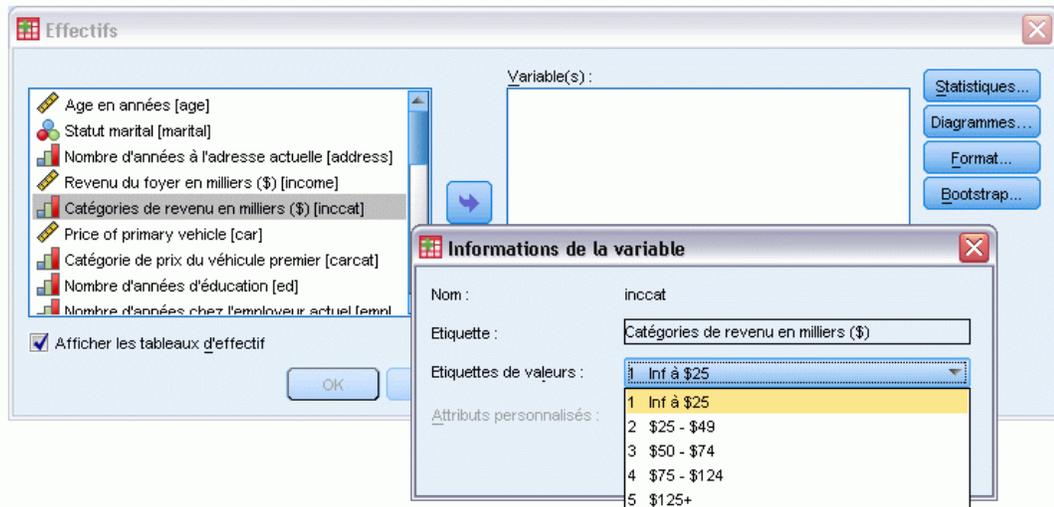


Dans la boîte de dialogue, choisissez les variables à analyser dans la liste source à gauche et faites-les glisser dans la liste Variable(s) à droite. Le bouton OK, qui exécute l'analyse, est désactivé jusqu'à ce qu'une variable soit placée dans la liste Variable(s).

Vous pouvez obtenir des informations supplémentaires en cliquant avec le bouton droit sur tout nom de variable dans la liste.

- ▶ Cliquez avec le bouton droit de la souris sur *Catégories de revenu en milliers [inccat]* et sélectionnez Informations de la variable.
- ▶ Cliquez sur la flèche vers le bas dans la liste déroulante Etiquettes de valeurs.

Figure 1-8
Étiquettes définies pour les variables de revenus



Toutes les étiquettes de valeur définies pour la variable apparaissent.

- Cliquez sur *Sexe [gender]* dans la liste des variables source, puis faites glisser la variable dans la liste cible Variable(s).
- Cliquez sur la variable *Modalité de revenus en milliers [rev_dis]* dans la liste source, puis faites-la glisser vers la liste cible.

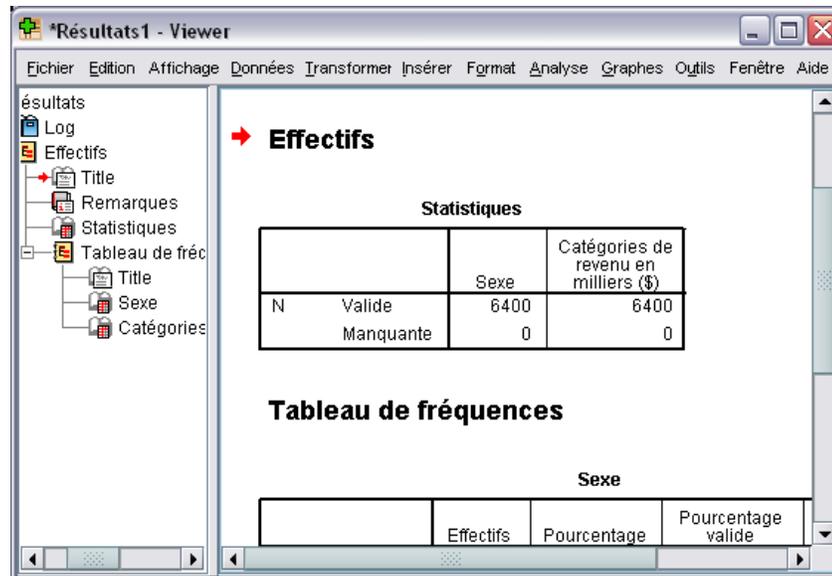
Figure 1-9
Variables sélectionnées pour l'analyse



- Cliquez sur OK pour exécuter la procédure.

Affichage des résultats

Figure 1-10
Fenêtre du Viewer

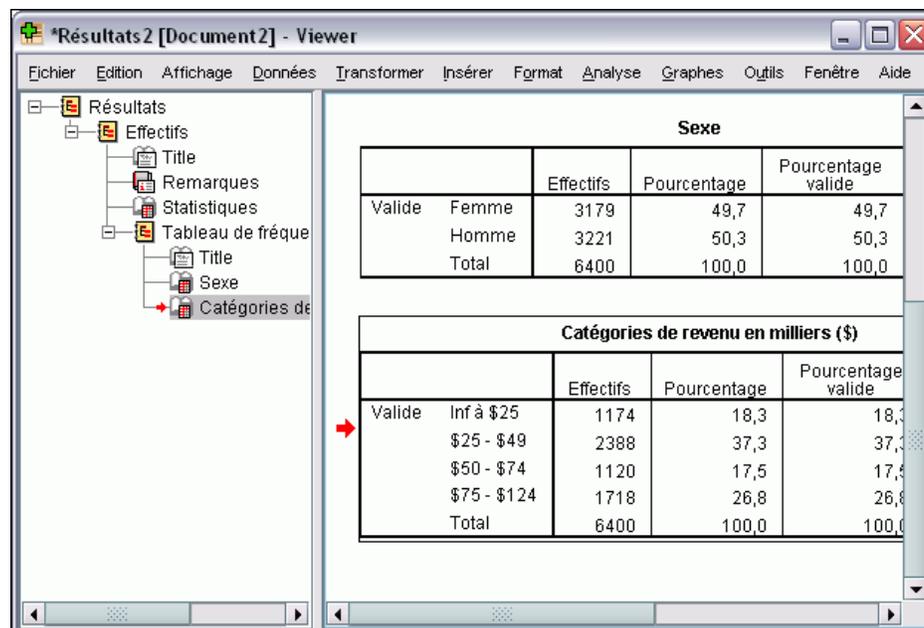


Les résultats sont affichés dans la fenêtre du Viewer.

Vous pouvez rapidement accéder à n'importe quel élément du Viewer en le sélectionnant dans le panneau de légende.

- Cliquez sur Catégories de revenu en milliers [inccat].

Figure 1-11
Tableau d'effectifs des modalités de revenus



Le tableau d'effectifs des modalités de revenus s'affiche. Ce tableau d'effectifs indique le nombre et le pourcentage de personnes dans chaque catégorie de revenus.

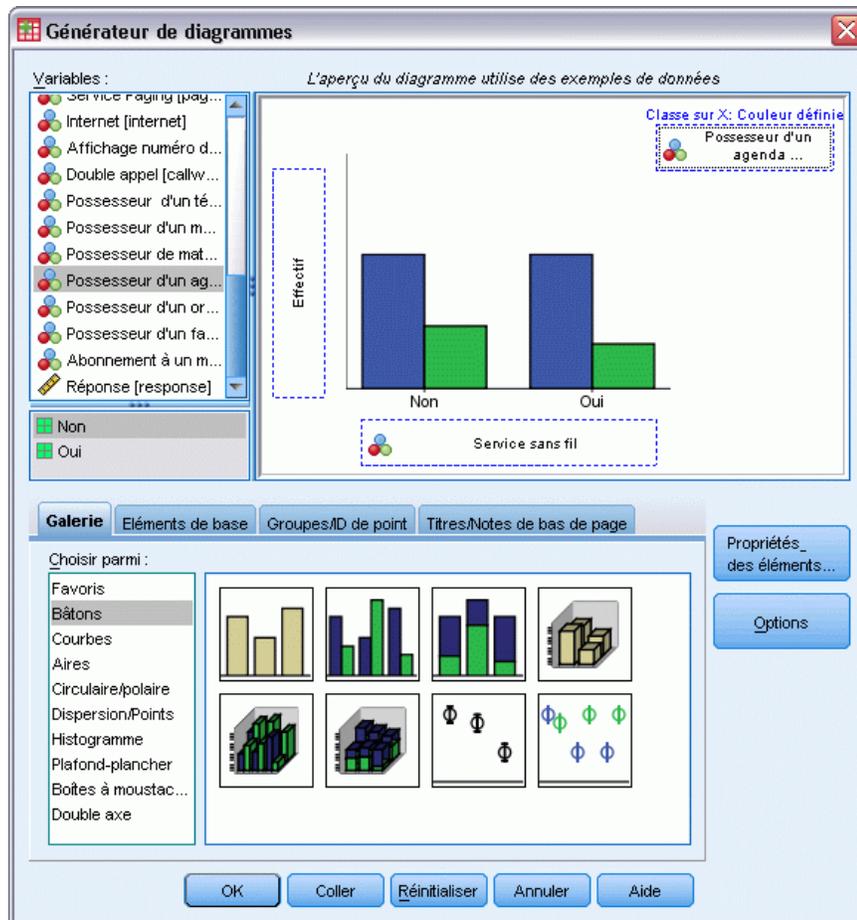
Création de diagrammes

Bien que certaines procédures statistiques puissent créer des diagrammes, vous pouvez également utiliser le menu Graphes pour la création de diagrammes.

Vous pouvez, par exemple, créer un diagramme montrant la relation entre l'utilisation de services de téléphonie sans fil et la propriété d'un agenda électronique.

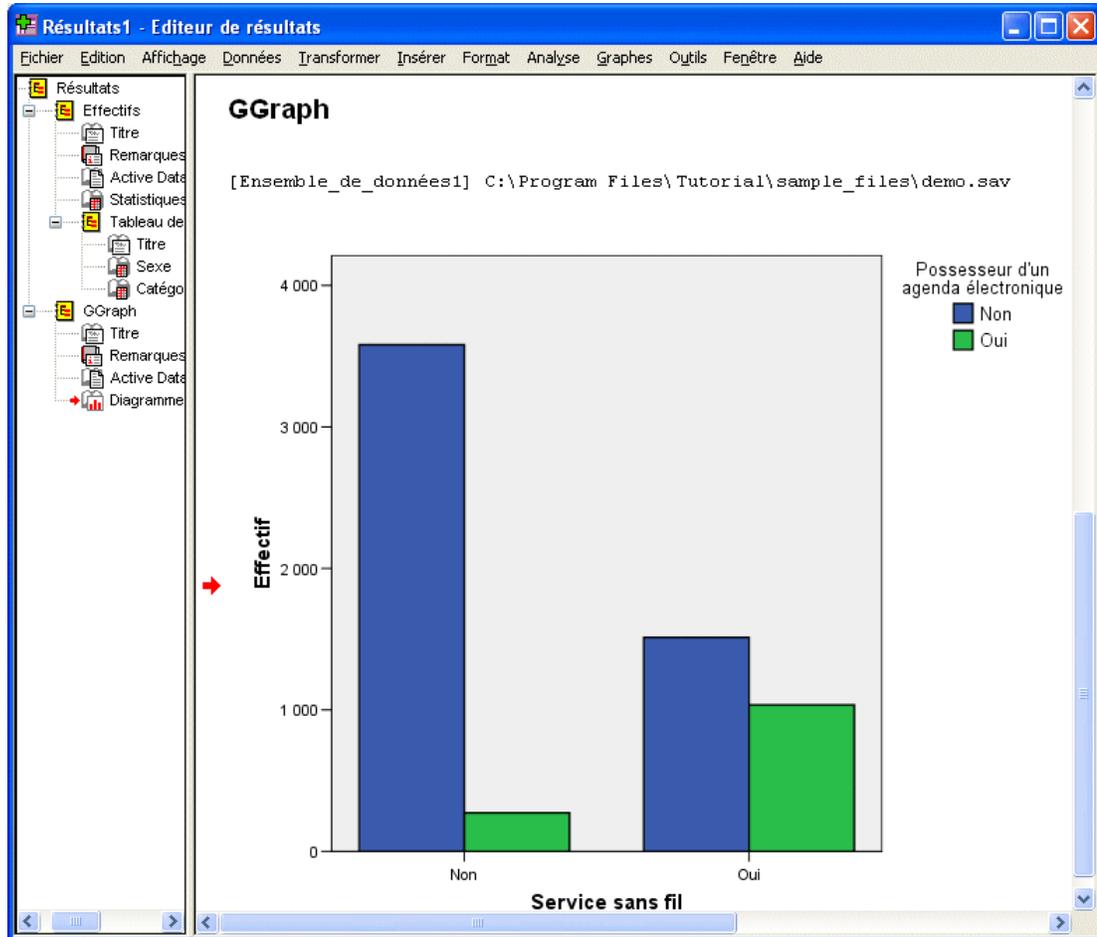
- ▶ A partir des menus, sélectionnez :
Graphes > Générateur de diagrammes...
- ▶ Cliquez sur l'onglet Galerie (s'il n'est pas sélectionné).
- ▶ Cliquez sur Bâton (s'il n'est pas sélectionné).
- ▶ Faites glisser l'icône Bâtons juxtaposés sur le canevas, qui est la zone étendue au-dessus de la galerie.

Figure 1-12
Boîte de dialogue Générateur de diagrammes



- ▶ Faites défiler la liste Variables et cliquez avec le bouton droit de la souris sur *Service sans fil [WiFi]* et sélectionnez Nominales comme niveau de mesure.
- ▶ Faites glisser la variable *Service sans fil [WiFi]* dans l'axe *x*.
- ▶ Cliquez avec le bouton droit de la souris sur *Possède un agenda électronique [pda]* et sélectionnez Nominales comme niveau de mesure.
- ▶ Faites glisser la variable *Possède un agenda électronique [pda]* dans la zone de déplacement de juxtaposition dans le coin supérieur droit du canevas.
- ▶ Cliquez sur OK pour créer le diagramme.

Figure 1-13
Diagramme en bâtons affiché dans la fenêtre du Viewer



Le diagramme en bâtons est affiché dans le Viewer. Le diagramme indique que la proportion de personnes possédant un agenda électronique est plus importante chez les personnes utilisant un service de téléphonie sans fil.

Vous pouvez modifier les diagrammes et les tableaux en double-cliquant dessus dans le panneau de contenu de la fenêtre de l'Editeur de résultats et vous pouvez copier-coller vos résultats dans d'autres applications. Ces sujets seront traités ultérieurement.

Lecture des données

Les données peuvent être saisies directement ou importées à partir de sources différentes. Ce chapitre traite des processus de lecture de données stockées dans les fichiers de données IBM® SPSS® Statistics, dans des applications de feuilles de calcul telles que Microsoft Excel, dans les applications de base de données telles que Microsoft Access et dans des fichiers texte.

Structure de base d'un fichier de données IBM SPSS Statistics

Figure 2-1
Editeur de données

The screenshot shows the 'demo.sav [Ensemble_de_données1] - Editeur de données' window. The menu bar includes 'Fichier', 'Edition', 'Affichage', 'Données', 'Transforme', 'Analyse', 'Graphes', 'Outils', 'Mgdules complémentaires', 'Fenêtre', and 'Aide'. The toolbar contains various icons for file operations and data manipulation. The main area displays a data table with the following columns: 'age', 'situatio', 'address', 'revenu', 'inccat', and 'car'. The row for 'age' 40 is selected. The status bar at the bottom indicates 'Affichage des données' and 'Processeur prêt'.

	age	situatio	address	revenu	inccat	car
13	31	1	12	40,00	2,00	21,30
14	42	0	29	137,00	4,00	68,90
15	35	0	8	70,00	3,00	34,10
16	52	1	24	159,00	4,00	78,90
17	21	1	1	37,00	2,00	18,60
18	32	0	0	28,00	2,00	13,70
19	42	0	9	109,00	4,00	54,70
20	40	1	12	117,00	4,00	58,30
21	30	0	3	23,00	1,00	11,80
22	48	0	14	21,00	1,00	9,50

Les fichiers de données IBM® SPSS® Statistics sont organisés par observations (lignes) et par variables (colonnes). Dans ce fichier de données, les observations représentent des répondants dans le cadre d'une enquête. Les variables représentent des réponses à chaque question posée au cours du sondage.

Lecture des fichiers de données IBM SPSS Statistics

Les fichiers de données IBM® SPSS® Statistics, portant l'extension de fichier *.sav*, contiennent les données que vous avez enregistrées. Pour ouvrir *demo.sav*, un fichier d'exemple installé avec le produit, procédez comme suit :

- ▶ A partir des menus, sélectionnez :
Fichier > Ouvrir > Données
- ▶ Recherchez le fichier *demo.sav* et ouvrez-le. Pour plus d'informations, reportez-vous à la section [Fichiers d'exemple dans l'annexe A sur p. 156](#).

Les données sont affichées dans l'éditeur de données.

Figure 2-2
Fichier de données ouvert

	age	situatio	address	revenu	inccat	car
13	31	1	12	40,00	2,00	21,30
14	42	0	29	137,00	4,00	68,90
15	35	0	8	70,00	3,00	34,10
16	52	1	24	159,00	4,00	78,90
17	21	1	1	37,00	2,00	18,60
18	32	0	0	28,00	2,00	13,70
19	42	0	9	109,00	4,00	54,70
20	40	1	12	117,00	4,00	58,30
21	30	0	3	23,00	1,00	11,80
22	48	0	14	21,00	1,00	9,50

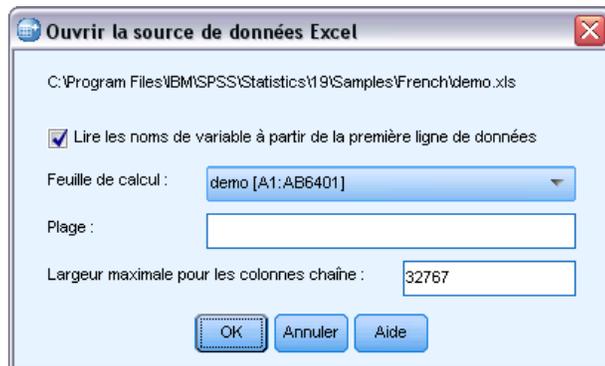
Lecture de données à partir de feuilles de calcul

Au lieu de saisir toutes vos données directement dans l'éditeur de données, vous pouvez les lire à partir d'applications telles que Microsoft Excel. Vous pouvez également lire des en-têtes de colonne comme noms de variable.

- ▶ A partir des menus, sélectionnez :
Fichier > Ouvrir > Données
- ▶ Sélectionnez Excel (*.xls) comme type de fichier à afficher.
- ▶ Ouvrez *demo.xls*. Pour plus d'informations, reportez-vous à la section [Fichiers d'exemple dans l'annexe A sur p. 156](#).

La boîte de dialogue Ouvrir la source de données Excel apparaît ; elle vous permet de choisir si les noms de variable doivent être inclus dans la feuille de calcul, et d'indiquer les cellules à importer. Dans Excel 95 ou version supérieure, vous pouvez également indiquer les feuilles de calcul à importer.

Figure 2-3
Boîte de dialogue Ouvrir la source de données Excel



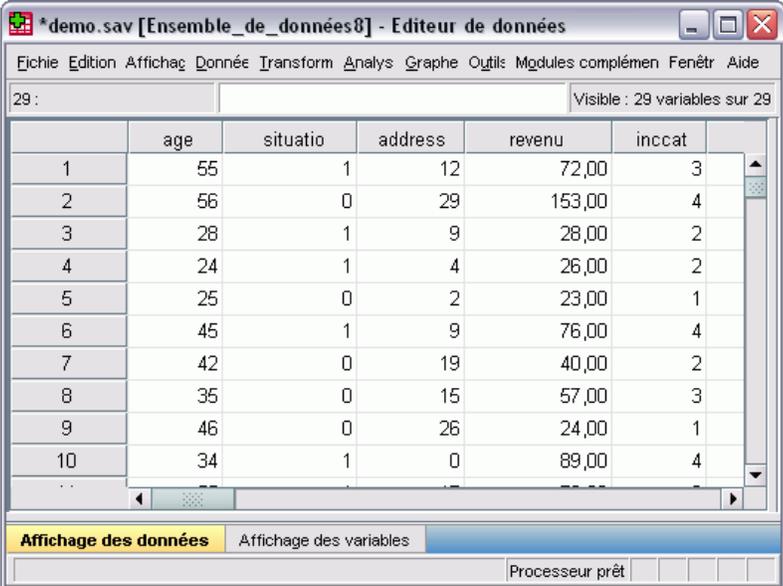
- Vérifiez que la case Lire le nom des variables à partir de la première ligne de données est sélectionnée. Cette option lit les en-têtes de colonne comme des noms de variable.

Si les en-têtes de colonne ne sont pas conformes aux règles de dénomination de variable IBM® SPSS® Statistics, ils sont convertis en noms de variable valides et les en-têtes de colonne d'origine sont enregistrés comme étiquettes de variable. Pour n'importer qu'une partie de la feuille de calcul, indiquez l'intervalle de cellules à importer dans la zone de texte Intervalle.

- Cliquez sur OK pour lire le fichier Excel.

Les données apparaissent à présent dans l'Editeur de données, les en-têtes de colonne étant utilisés comme noms de variable. Puisque les noms de variables ne peuvent pas contenir d'espaces, les espaces des en-têtes de colonne d'origine ont été supprimés. Par exemple, *Situation familiale* dans le fichier Excel devient la variable *Situatio*. L'en-tête de colonne d'origine est conservé comme étiquette de variable.

Figure 2-4
Données Excel importées



	age	situatio	address	revenu	inccat
1	55	1	12	72,00	3
2	56	0	29	153,00	4
3	28	1	9	28,00	2
4	24	1	4	26,00	2
5	25	0	2	23,00	1
6	45	1	9	76,00	4
7	42	0	19	40,00	2
8	35	0	15	57,00	3
9	46	0	26	24,00	1
10	34	1	0	89,00	4

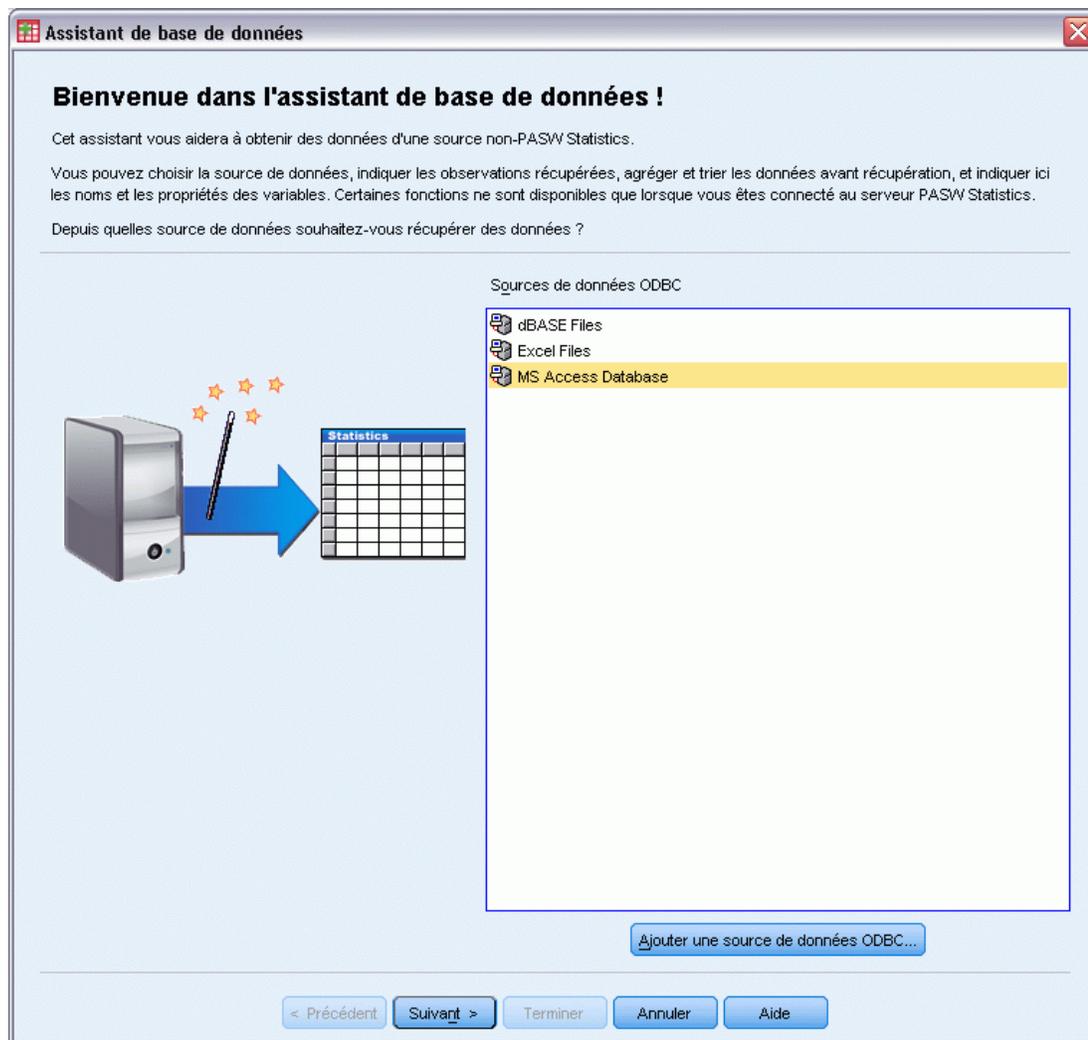
Lecture de données à partir d'une base de données

Les données issues de sources de bases de données peuvent facilement être importées à l'aide de l'Assistant de base de données. Toute base de données utilisant des pilotes ODBC (Open Database Connectivity) peut être lue directement une fois ces pilotes installés. Les pilotes ODBC de nombreux formats de base de données sont fournis sur le CD d'installation. Vous pouvez obtenir des pilotes supplémentaires auprès de distributeurs indépendants. L'une des applications de base de données les plus courantes, Microsoft Access, est traitée dans cet exemple.

Remarque : Cet exemple est propre à Microsoft Windows et nécessite un pilote ODBC pour Access. Les étapes sont semblables sur les autres plateformes mais peuvent nécessiter un pilote ODBC tiers pour Access.

- ▶ A partir des menus, sélectionnez :
Fichier > Ouvrir la base de données > Nouvelle requête...

Figure 2-5
Boîte de dialogue Bienvenue dans l'assistant de base de données



- Sélectionnez Base de données MS Access dans la liste des sources de données et cliquez sur Suivant.

Remarque : En fonction de votre installation, vous pouvez voir une liste des sources de données OLEDB à gauche de l'assistant (systèmes d'exploitation Windows uniquement), mais cet exemple utilise la liste de sources de données ODBC affichée sur la droite.

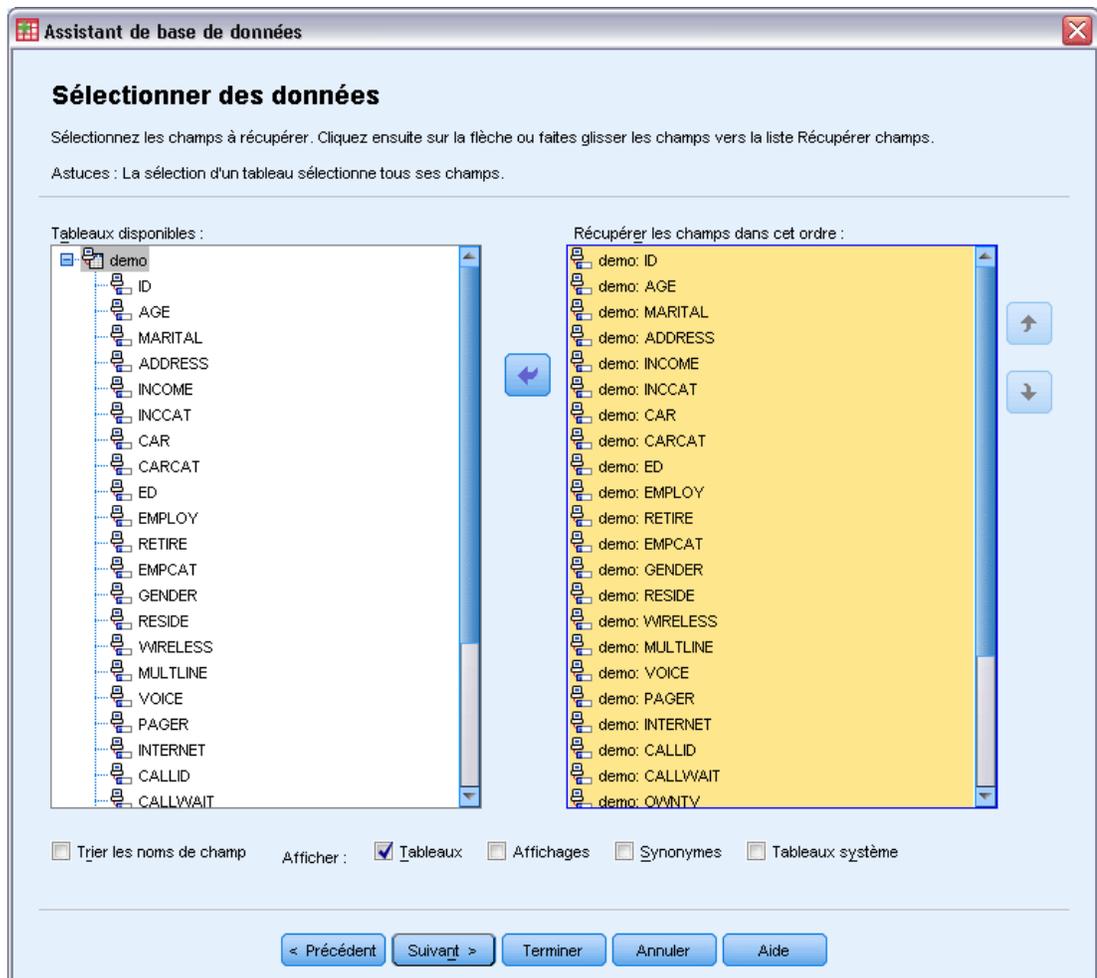
Figure 2-6
Boîte de dialogue Connexion au pilote ODBC



- ▶ Cliquez sur Parcourir pour accéder au fichier de base de données Access à ouvrir.
- ▶ Ouvrez *demo.mdb*. Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.
- ▶ Cliquez sur OK dans la boîte de dialogue de connexion.

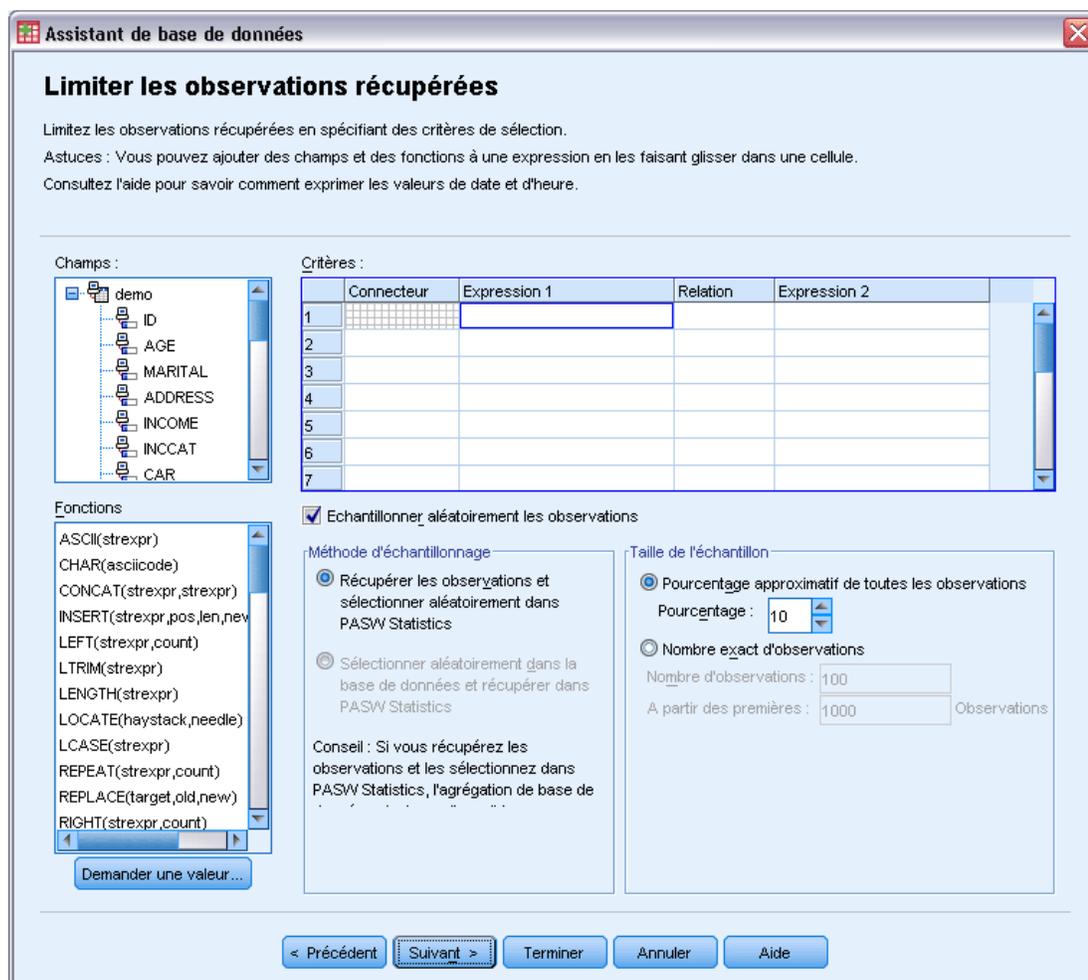
A l'étape suivante, vous pourrez indiquer les tables et les variables que vous souhaitez importer.

Figure 2-7
Etape Sélectionner des données



- Faites glisser le tableau demo vers la liste Rétablir les champs dans cet ordre.
 - Cliquez sur Suivant.
- A l'étape 4, sélectionnez les enregistrements (observations) à importer.

Figure 2-8
Etape Limiter les observations récupérées



Si vous ne souhaitez pas importer toutes les observations, vous pouvez sélectionner un sous-ensemble d'observations (par exemple, les hommes de plus de 30 ans) ou un échantillon aléatoire d'observations dans la source de données. Pour les sources de données volumineuses, vous pouvez limiter le nombre d'observations à un échantillon restreint et représentatif afin de réduire la durée du traitement.

- Cliquez sur Suivant pour continuer.

Les noms de champ permettent de créer des noms de variable. Si nécessaire, ils sont convertis en noms de variable valides. Les noms de champ d'origine sont conservés en tant qu'étiquettes de variable. Vous pouvez également modifier les noms de variable avant d'importer la base de données.

Figure 2-9
Etape Définir variables

Assistant de base de données

Définir des variables

Modifier les noms et les propriétés des variables.

Des noms seront générés automatiquement si vous n'en indiquez aucun.

Pour convertir une variable chaîne au format numérique en utilisant les valeurs d'origine comme étiquettes de valeur, sélectionnez Recoder en numérique.

	Nom de la variable de résultat	Le type de don...	Recoder en numérique
demo: ID	ID	Numérique	
demo: AGE	AGE	Numérique	
demo: MARITAL	MARITAL	Numérique	
demo: ADDRESS	ADDRESS	Numérique	
demo: INCOME	INCOME	Numérique	
demo: INCCAT	INCCAT	Numérique	
demo: CAR	CAR	Numérique	
demo: CARCAT	CARCAT	Numérique	
demo: ED	ED	Numérique	
demo: EMPLOY	EMPLOY	Numérique	
demo: RETIRE	RETIRE	Numérique	
demo: EMPCAT	EMPCAT	Numérique	
demo: GENDER	GENDER	Chaîne	<input checked="" type="checkbox"/>
demo: RESIDE	RESIDE	Numérique	
demo: WIRELESS	WIRELESS	Numérique	
demo: MULTLINE	MULTLINE	Numérique	
demo: VOICE	VOICE	Numérique	
demo: PAGER	PAGER	Numérique	
demo: INTERNET	INTERNET	Numérique	
demo: CALLID	CALLID	Numérique	
demo: CALLWAIT	CALLWAIT	Numérique	
demo: OWNTV	OWNTV	Numérique	

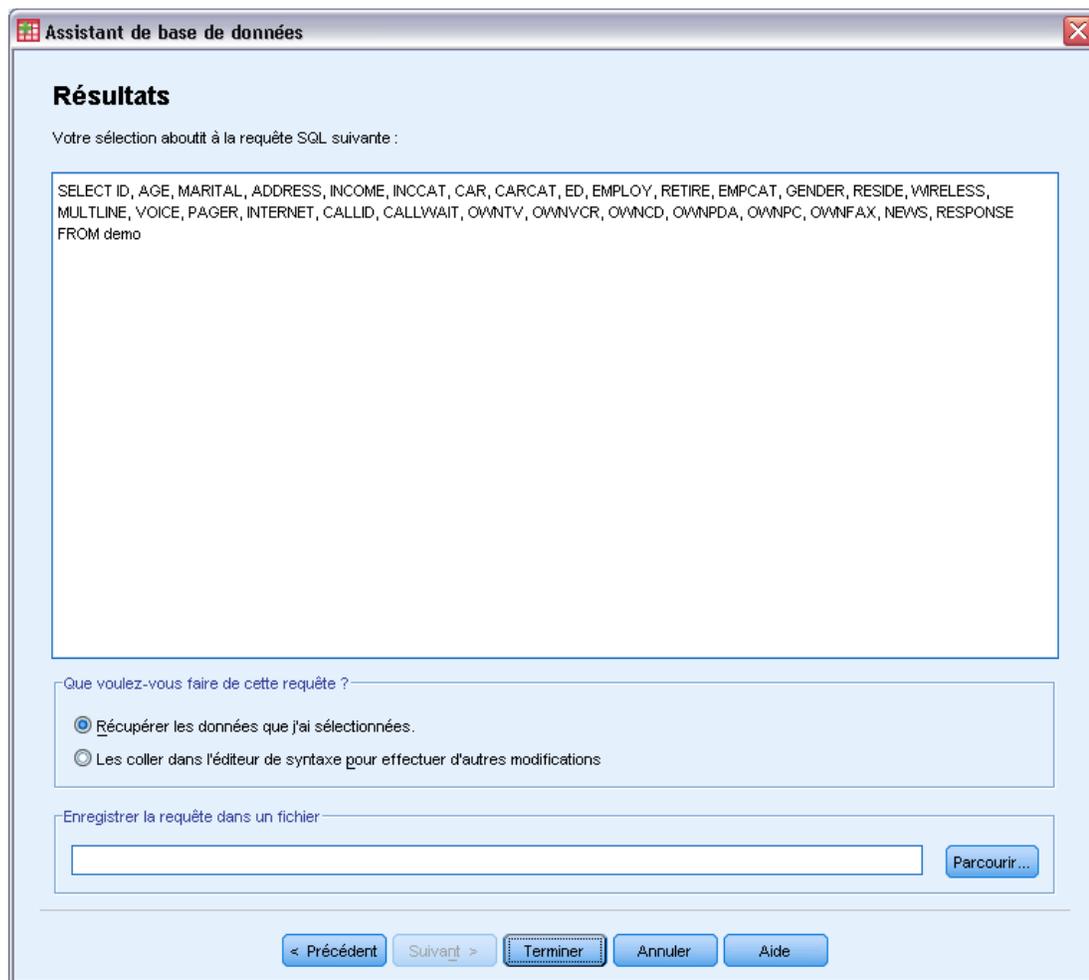
Largeur des champs de chaîne à largeur variable : Réduire les largeurs de chaîne en fonction des valeurs observées

< Précédent **Suivant** > Terminer Annuler Aide

- ▶ Cliquez sur la cellule Recoder en numérique dans le champ Sexe. Cette option convertit les variables chaîne en variables numériques (entiers) et conserve la valeur d'origine comme étiquette de valeur de la nouvelle variable.
- ▶ Cliquez sur Suivant pour continuer.

L'instruction SQL créée à partir de vos sélections dans l'Assistant de base de données apparaît à l'étape Résultats. Cette instruction peut être exécutée immédiatement ou enregistrée dans un fichier pour une utilisation ultérieure.

Figure 2-10
Etape Résultats



- Cliquez sur Terminer pour importer les données.

Toutes les données de la base de données Access sélectionnées pour l'importation sont à présent disponibles dans l'éditeur de données.

Figure 2-11

Données importées à partir d'une base de données Access

	age	situatio	address	revenu	inccat
1	55	1	12	72,00	3
2	56	0	29	153,00	4
3	28	1	9	28,00	2
4	24	1	4	26,00	2
5	25	0	2	23,00	1
6	45	1	9	76,00	4
7	42	0	19	40,00	2
8	35	0	15	57,00	3
9	46	0	26	24,00	1
10	34	1	0	89,00	4

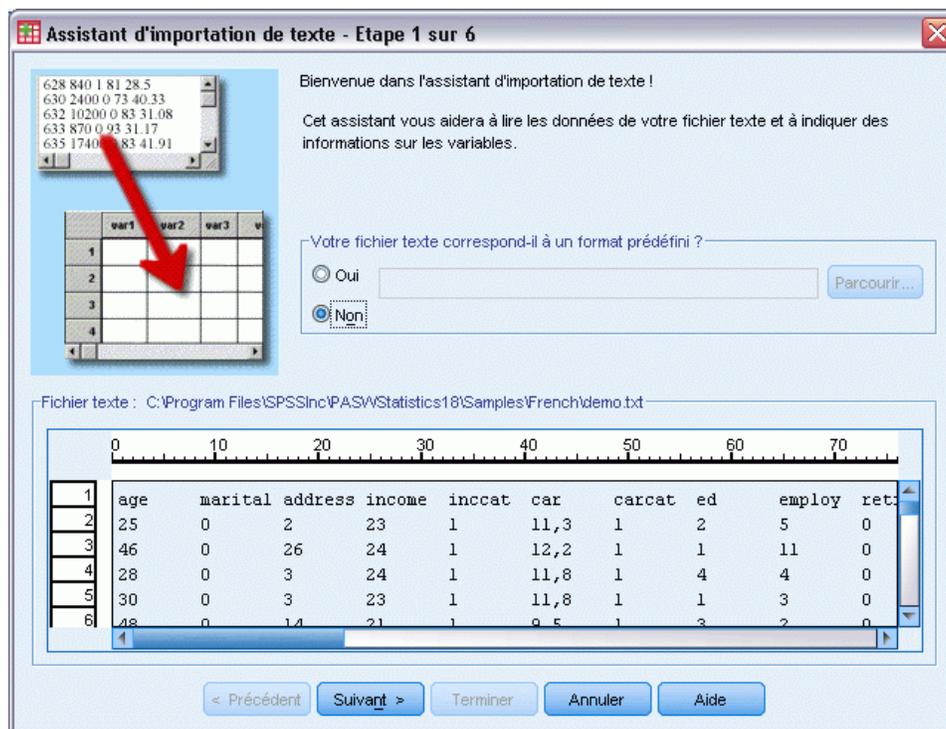
Lecture de données à partir d'un fichier texte

Les fichiers texte sont une autre source de données courante. De nombreux tableurs et bases de données peuvent enregistrer leur contenu dans l'un des nombreux formats de fichier texte. Les fichiers délimités par une virgule ou une tabulation se rapportent aux lignes de données utilisant des virgules ou des tabulations pour indiquer chaque variable. Dans cet exemple, les données sont délimitées par des tabulations.

- ▶ A partir des menus, sélectionnez :
Fichier > Lire les données texte...
- ▶ Sélectionnez Texte (*.txt) comme type de fichier à afficher.
- ▶ Ouvrez *demo.txt*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.](#)

L'Assistant d'importation de texte vous guide tout au long du processus permettant de définir le mode d'impression du fichier texte indiqué.

Figure 2-12
Assistant d'importation de texte : Etape 1 sur 6

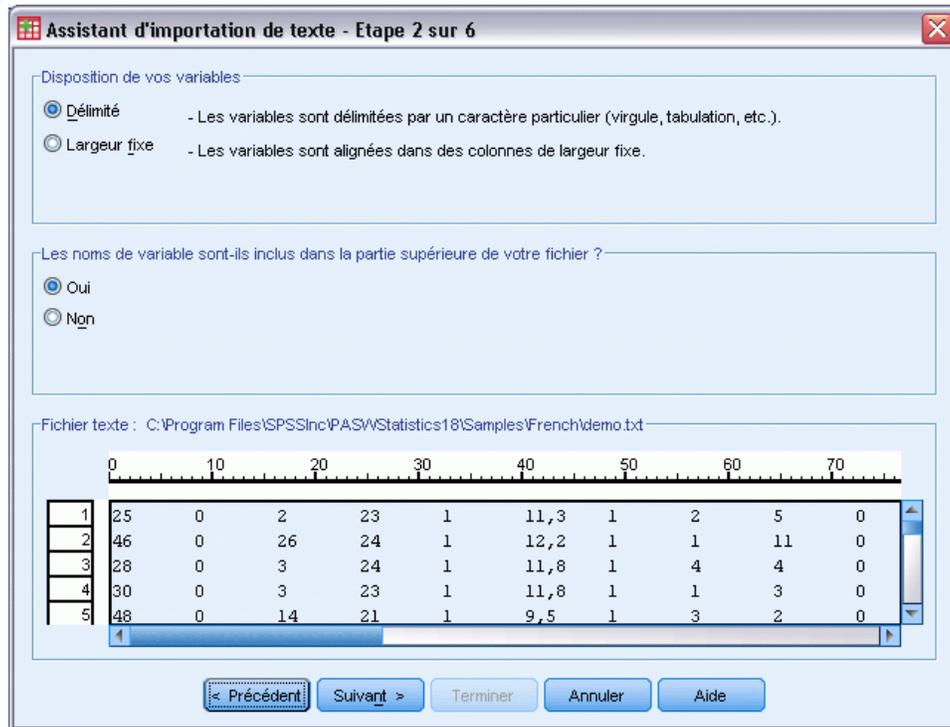


- ▶ A l'étape 1, vous pourrez sélectionner un format prédéfini ou créer un format dans l'Assistant. Sélectionnez Non pour indiquer qu'un nouveau format doit être créé.
- ▶ Cliquez sur Suivant pour continuer.

Comme indiqué précédemment, ce fichier utilise un format délimité par des tabulations. En outre, les noms de variable sont définis sur la première ligne de ce fichier.

Figure 2-13

Assistant d'importation de texte : Etape 2 sur 6



- ▶ Sélectionnez Délimité pour indiquer que les données utilisent une structure de format délimité.
- ▶ Sélectionnez Oui pour indiquer que les noms de variable doivent être lus à partir du début du fichier.
- ▶ Cliquez sur Suivant pour continuer.

- Saisissez 2 dans la section supérieure de la boîte de dialogue suivante pour indiquer que la première ligne de données commence sur la deuxième ligne du fichier texte.

Figure 2-14

Assistant d'importation de texte : Etape 3 sur 6

Assistant d'importation de texte - Délimité, étape 3 sur 6

La première observation commence à la ligne

Représentation de vos observations:

Chaque ligne représente une observation

Un nombre spécifique de variables représente une observation :

Combien d'observations souhaitez-vous importer ?

Toutes les observations

Les premières observations.

Un pourcentage aléatoire des observations (approximatif) : %

Aperçu des données:

	0	10	20	30	40	50	60	70		
1	25	0	2	23	1	11,3	1	2	5	0
2	46	0	26	24	1	12,2	1	1	11	0
3	28	0	3	24	1	11,8	1	4	4	0
4	30	0	3	23	1	11,8	1	1	3	0

< Précédent Suivant > Terminer Annuler Aide

- Conservez les valeurs par défaut des autres champs de cette boîte de dialogue et cliquez sur Suivant pour continuer.

L'aperçu des données de l'étape 4 vous offre un moyen rapide de vérifier que vos données ont été lues correctement.

Figure 2-15

Assistant d'importation de texte : Etape 4 sur 6

Assistant d'importation de texte - Délimité, étape 4 sur 6

Quels séparateurs s'affichent entre les variables ?

Tab Barre d'espace
 Virgule Point-virgule
 Autre :

Qu'est-ce qu'un qualificateur de texte ?

Aucun
 Apostrophe
 Guillemet
 Autre :

Aperçu des données

age	marital	address	income	inccat	car	carcat	ed
25	0	2	23	1	11,3	1	2
46	0	26	24	1	12,2	1	1
28	0	3	24	1	11,8	1	4
30	0	3	23	1	11,8	1	1
48	0	14	21	1	9,5	1	3
39	1	17	17	1	8,5	1	4
34	0	10	20	1	10	1	3
47	1	6	22	1	11,4	1	3
30	1	7	21	1	10,5	1	4

< Précédent Suivant > Terminer Annuler Aide

- ▶ Sélectionnez Tab et désélectionnez les autres options.
- ▶ Cliquez sur Suivant pour continuer.

Etant donné que des noms de variable peuvent avoir été tronqués pour des raisons de formatage, cette boîte de dialogue vous permet de modifier les noms superflus.

Figure 2-16
Assistant d'importation de texte : Etape 5 sur 6

Assistant d'importation de texte - Etape 5 sur 6

Spécifications pour les variables sélectionnées dans l'aperçu de données

Nom de variable : age
Nom d'origine : age
Format de données : Numérique

Aperçu des données

age	marital	address	income	inccat	car	carcat	ed
25	0	2	23	1	11,3	1	2
46	0	26	24	1	12,2	1	1
28	0	3	24	1	11,8	1	4
30	0	3	23	1	11,8	1	1
48	0	14	21	1	9,5	1	3

< Précédent Suivant > Terminer Annuler Aide

Vous pouvez également définir les types de données dans cette boîte de dialogue. Par exemple, nous pouvons supposer que la variable de revenus doit contenir une certaine somme en dollars.

Pour modifier un type de données :

- Sous l'aperçu de données, sélectionnez la variable à modifier, c'est-à-dire *Revenu* dans cet exemple.

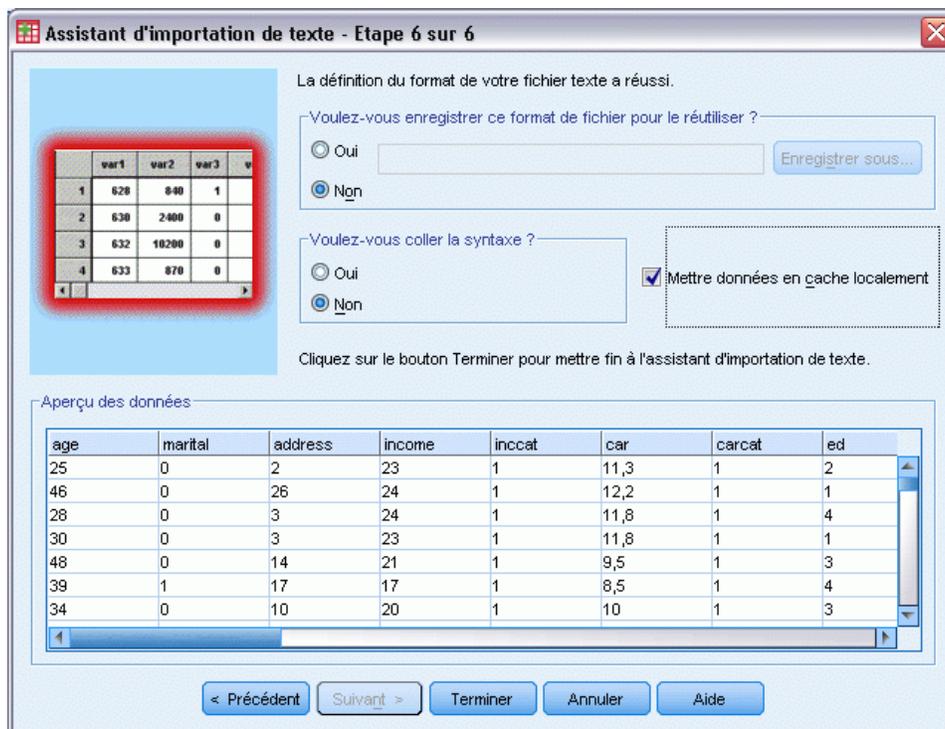
- Sélectionnez Dollar dans la liste déroulante Format des données.

Figure 2-17
Sélectionnez le type de données.



- Cliquez sur Suivant pour continuer.

Figure 2-18
Assistant d'importation de texte : Etape 6 sur 6



- Conservez les sélections par défaut dans cette boîte de dialogue et cliquez sur Terminer pour importer les données.

Utilisation de l'éditeur de données

Cette fenêtre affiche le contenu du fichier de données actif. Les informations contenues dans l'éditeur de données se composent de variables et d'observations.

- Dans l'affichage des données, les colonnes représentent des variables et les lignes des observations.
- Dans l'affichage des variables, chaque ligne est une variable et chaque colonne est un attribut associé à cette variable.

Les variables sont utilisées pour représenter les différents types de données que vous avez rassemblées. Une analogie courante est celle de l'enquête. La réponse à chaque question d'une enquête est équivalente à une variable. Il existe de nombreux types de variable : nombres, caractères, monnaies et dates.

Saisie de données numériques

Les données peuvent être saisies dans l'éditeur de données, qui peut s'avérer utile pour traiter les fichiers de données peu volumineux ou pour apporter de légères modifications à des fichiers de données plus volumineux.

- ▶ Cliquez sur l'onglet *Affichage des variables* en bas de la fenêtre de l'éditeur de données.

Vous devez définir les variables qui seront utilisées. Dans cet exemple, seules trois variables sont nécessaires : *age*, *situatio* et *revenu*.

Figure 3-1
Noms de variable dans l'affichage des variables

	Nom	Type	Largeur	Décimales	Etiquette	Valeurs	Manq
1	age	Numérique	8	2		Aucun	Aucun
2	situatio	Numérique	8	2		Aucun	Aucun
3	revenu	Numérique	8	2		Aucun	Aucun
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							

Affichage des données **Affichage des variables**

- ▶ Dans la première ligne de la première colonne, saisissez age.
- ▶ Dans la deuxième ligne, saisissez situatio.
- ▶ Dans la troisième ligne, saisissez revenu.

Un type de données numérique est automatiquement attribué aux nouvelles variables.

Si vous ne saisissez pas de noms de variable, des noms uniques sont automatiquement créés. Cependant, ces noms ne sont pas descriptifs et ne sont pas recommandés pour les fichiers de données volumineux.

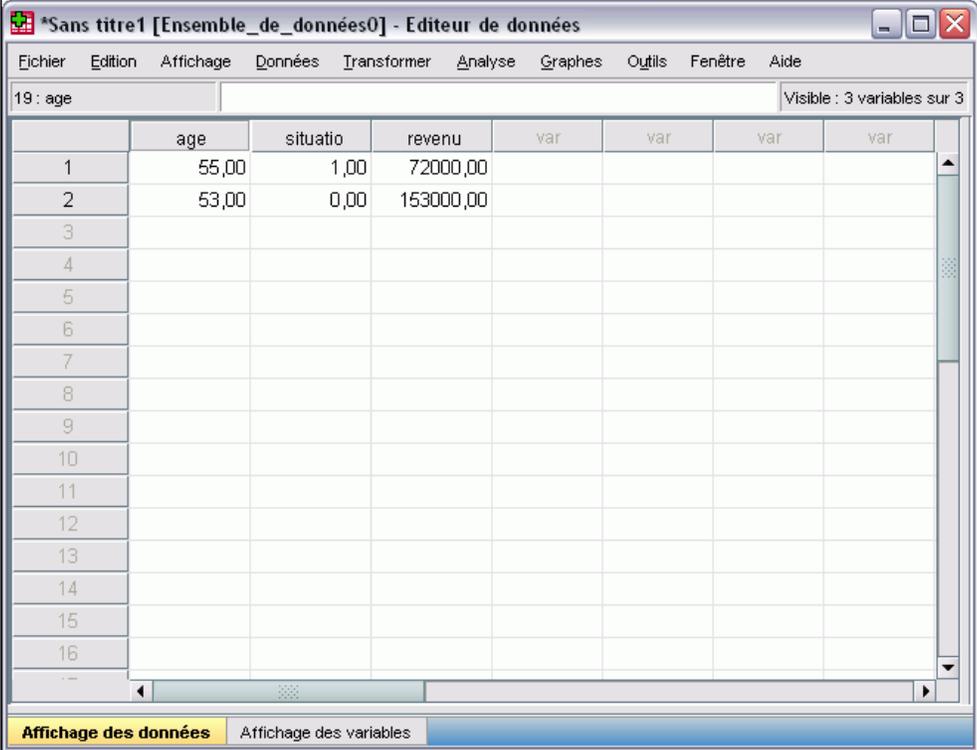
- ▶ Cliquez sur l'onglet Affichage des données pour continuer à saisir des données.

Les noms saisis dans l'affichage des variables sont à présent les en-têtes des trois premières colonnes de l'affichage des données.

Commencez à saisir des données dans la première ligne, en commençant par la première colonne.

Figure 3-2

Valeurs entrées dans l'affichage des données



	age	situatio	revenu	var	var	var	var
1	55,00	1,00	72000,00				
2	53,00	0,00	153000,00				
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							

- ▶ Dans la colonne *age*, saisissez 55.
- ▶ Dans la colonne *situatio*, saisissez 1.
- ▶ Dans la colonne *revenu*, saisissez 72000.
- ▶ Placez le curseur dans la seconde ligne de la première colonne pour ajouter les données du sujet suivant.
- ▶ Dans la colonne *age*, saisissez 53.
- ▶ Dans la colonne *situatio*, saisissez 0.
- ▶ Dans la colonne *revenu*, saisissez 153 000.

Les colonnes *age* et *situatio* affichent actuellement un séparateur décimal alors que les valeurs sont des entiers. Pour masquer le séparateur décimal de ces variables :

- ▶ Cliquez sur l'onglet *Affichage des variables* en bas de la fenêtre de l'éditeur de données.
- ▶ Dans la colonne *Décimales* de la ligne *âge*, saisissez 0 pour masquer la décimale.

- Dans la colonne *Décimales* de la ligne *situatio*, saisissez 0 pour masquer la décimale.

Figure 3-3

Propriété *Décimales* des variables *age* et *situatio* mise à jour

	Nom	Type	Largeur	Décimales	Etiquette	Valeurs	Manq
1	age	Numérique	8	0		Aucun	Aucun
2	situatio	Numérique	8	0		Aucun	Aucun
3	revenu	Numérique	8	2		Aucun	Aucun
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							

Affichage des données **Affichage des variables**

Saisie de données chaîne

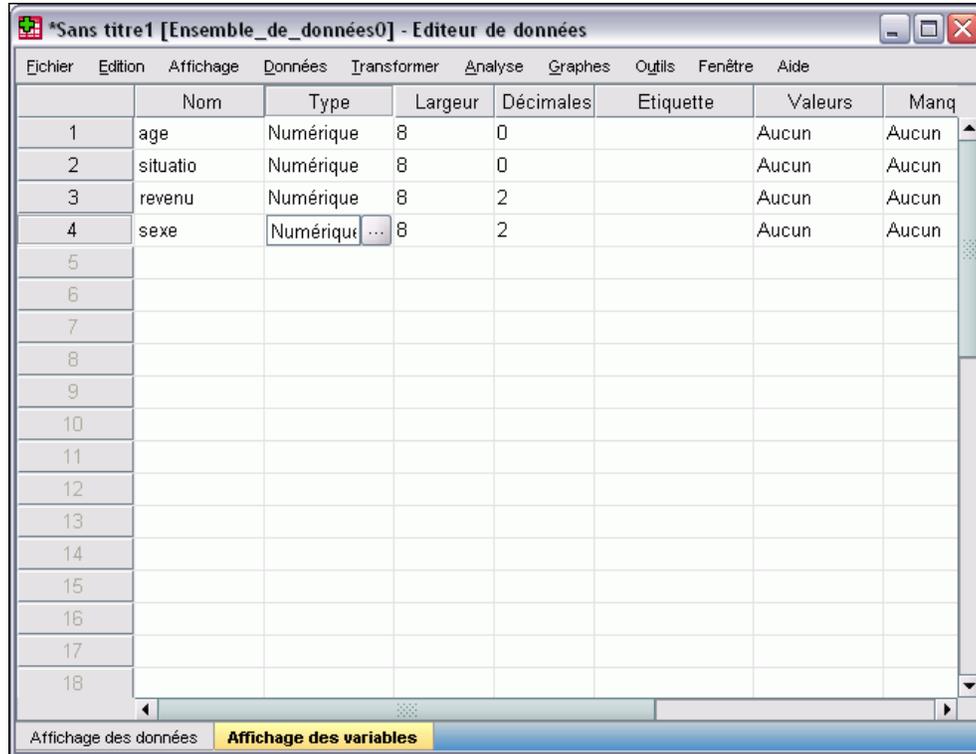
Des données non numériques, telles que des chaînes de texte, peuvent également être saisies dans l'éditeur de données.

- Cliquez sur l'onglet *Affichage des variables* en bas de la fenêtre de l'éditeur de données.
- Dans la première cellule de la première ligne vide, saisissez *sexe* comme nom de variable.
- Cliquez sur la cellule *Type* à côté de votre entrée.

- Cliquez sur le bouton à droite de la cellule *Type* pour ouvrir la boîte de dialogue Type de variable.

Figure 3-4

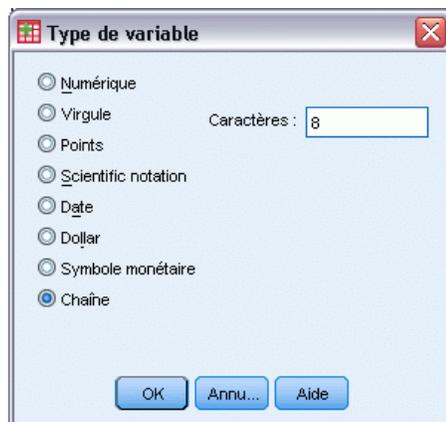
Bouton présent dans la cellule *Type* de la variable *sexe*



- Sélectionnez Chaîne pour indiquer le type de variable.
- Cliquez sur OK pour enregistrer votre sélection et revenir dans Data Editor.

Figure 3-5

Boîte de dialogue Type de variable



Définition de données

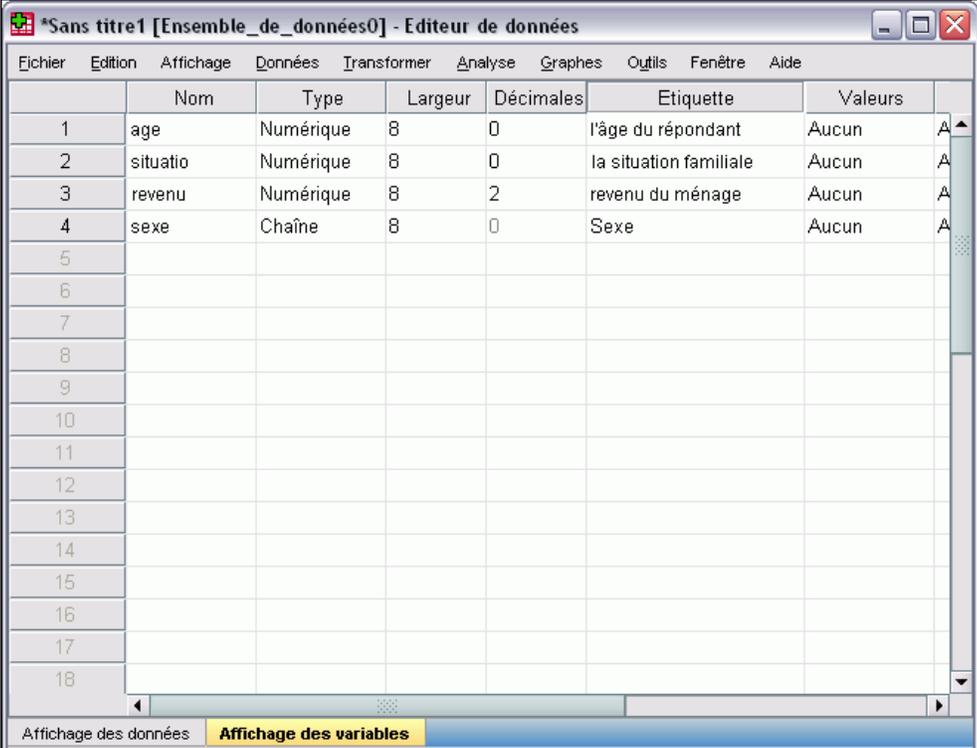
Outre les données, vous pouvez aussi définir des étiquettes de variables descriptives et des étiquettes de valeurs pour les noms de variables et les valeurs de données. Ces étiquettes descriptives sont utilisées dans les rapports statistiques et les diagrammes.

Ajout des étiquettes de variables

Les étiquettes fournissent une description des variables. Ces descriptions correspondent souvent à une version plus longue des noms de variable. Les étiquettes peuvent s'élever jusqu'à 255 octets. Ces étiquettes sont utilisées dans les résultats pour identifier les variables différentes.

- ▶ Cliquez sur l'onglet Affichage des variables en bas de la fenêtre de l'éditeur de données.
- ▶ Dans la colonne *Etiquette* de la ligne *age*, saisissez l'âge du répondant.
- ▶ Dans la colonne *Etiquette* de la ligne *situatio*, saisissez la situation familiale.
- ▶ Dans la colonne *Etiquette* de la ligne *revenu*, saisissez revenu du ménage.
- ▶ Dans la colonne *Etiquette* de la ligne *sexe*, saisissez le sexe.

Figure 3-6
Etiquettes de variable entrées dans l'affichage des variables



	Nom	Type	Largeur	Décimales	Etiquette	Valeurs	
1	age	Numérique	8	0	l'âge du répondant	Aucun	A ▲
2	situatio	Numérique	8	0	la situation familiale	Aucun	A
3	revenu	Numérique	8	2	revenu du ménage	Aucun	A
4	sexe	Chaîne	8	0	Sexe	Aucun	A
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							

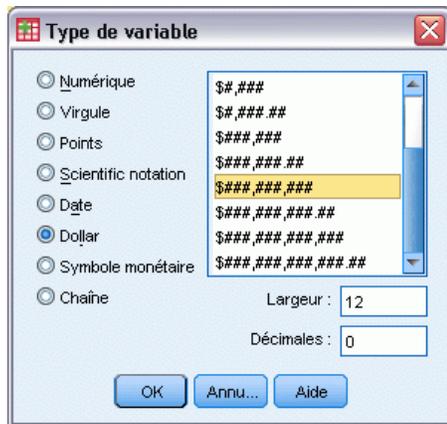
Affichage des données **Affichage des variables**

Modification du type et du format de variable

La colonne *Type* affiche le type de données actuel de chaque variable. Les types de données les plus courants sont les formats numériques et chaînes, mais de nombreux autres formats sont pris en charge. Dans le fichier de données actuel, la variable *revenu* est définie comme étant un type numérique.

- ▶ Cliquez sur la cellule *Type* de la ligne *revenu*, puis sur le bouton à droite de la cellule pour ouvrir la boîte de dialogue Type de variable.
- ▶ Sélectionnez Dollar.

Figure 3-7
Boîte de dialogue Type de variable



Les options de format du type de données sélectionné apparaissent.

- ▶ Dans cet exemple, sélectionnez `$###,###,###` comme format de devise.
- ▶ Cliquez sur OK pour appliquer vos modifications.

Ajout d'étiquettes de valeur pour les variables numériques

Les étiquettes de valeur correspondent à une façon d'associer aux valeurs de variable des étiquettes de chaîne. Dans cet exemple, il y a deux valeurs possibles pour la variable *situatio*. La valeur 0 signifie que le sujet est célibataire et la valeur 1 qu'il est marié.

- ▶ Cliquez sur la cellule *Valeurs* de la ligne *situatio*, puis sur le bouton à droite de la cellule pour ouvrir la boîte de dialogue Etiquettes de valeur.

Le champ **Valeur** correspond à la valeur numérique réelle.

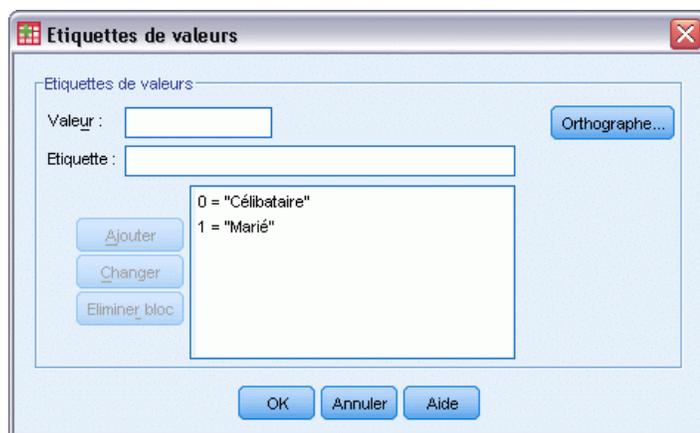
Le champ **Étiquette de valeur** correspond à l'étiquette de chaîne appliquée à la valeur numérique indiquée.

- ▶ Saisissez 0 dans le champ Valeur.
- ▶ Saisissez Célibataire dans le champ Etiquette.

- ▶ Cliquez sur Ajouter pour ajouter l'étiquette à la liste.

Figure 3-8

Boîte de dialogue Etiquettes de valeurs



- ▶ Saisissez 1 dans le champ Valeur, puis Marié dans le champ Etiquette.
- ▶ Cliquez sur Ajouter, puis sur OK pour enregistrer vos modifications et revenir dans l'éditeur de données.

Ces étiquettes peuvent également apparaître dans l'affichage des données pour faciliter la lecture de vos données.

- ▶ Cliquez sur l'onglet Affichage des données en bas de la fenêtre de l'éditeur de données.
- ▶ A partir des menus, sélectionnez :
Affichage > Etiquettes de valeur

Les étiquettes figurent à présent dans une liste lorsque vous entrez des valeurs dans Data Editor. Cette configuration a l'avantage de proposer des réponses valides et plus descriptives.

Si l'élément de menu *Étiquettes de valeurs* est déjà activé (lorsqu'il est indiqué par une coche), sélectionner à nouveau *Étiquettes de valeur désactive* l'affichage des étiquettes de valeurs.

Figure 3-9
Étiquettes de valeur dans l'affichage des données

The screenshot shows a window titled '*Sans titre1 [Ensemble_de_données0] - Editeur de données'. The menu bar includes 'Fichier', 'Edition', 'Affichage', 'Données', 'Transformer', 'Analyse', 'Graphes', 'Outils', 'Fenêtre', and 'Aide'. Below the menu bar, there is a status bar indicating '22:' and 'Visible : 4 variables sur 4'. The main area contains a table with the following data:

	age	situatio	revenu	sexe	var	var	var
1	55	Marié	\$72,000.00				
2	53	Célibataire	\$153,000.00				
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							

At the bottom of the window, there are two tabs: 'Affichage des données' (selected) and 'Affichage des variables'.

Ajout d'étiquettes de valeur pour les variables chaîne

Les variables chaîne peuvent également nécessiter des étiquettes de valeur. Par exemple, les lettres *M* ou *F* sont utilisées pour identifier le sexe du sujet. Les étiquettes de valeur peuvent être utilisées pour préciser que *M* signifie *Masculin* et *F* *Féminin*.

- ▶ Cliquez sur l'onglet *Affichage des variables* en bas de la fenêtre de l'éditeur de données.
- ▶ Cliquez sur la cellule *Valeurs* de la ligne *sexe*, puis sur le bouton à droite de la cellule pour ouvrir la boîte de dialogue *Étiquettes de valeur*.
- ▶ Saisissez *F* dans le champ *Valeur*, puis *Féminin* dans le champ *Étiquette*.

- ▶ Cliquez sur Ajouter pour ajouter l'étiquette au fichier de données.

Figure 3-10
Boîte de dialogue Etiquettes de valeurs



- ▶ Saisissez M dans le champ Valeur, puis Masculin dans le champ Etiquette.
- ▶ Cliquez sur Ajouter, puis sur OK pour enregistrer vos modifications et revenir dans l'éditeur de données.

Les valeurs de chaîne font la distinction entre les majuscules et les minuscules, vous devez donc être cohérent. Le *m* minuscule est différent du *M* majuscule.

Utilisation des étiquettes de valeurs pour la saisie de données

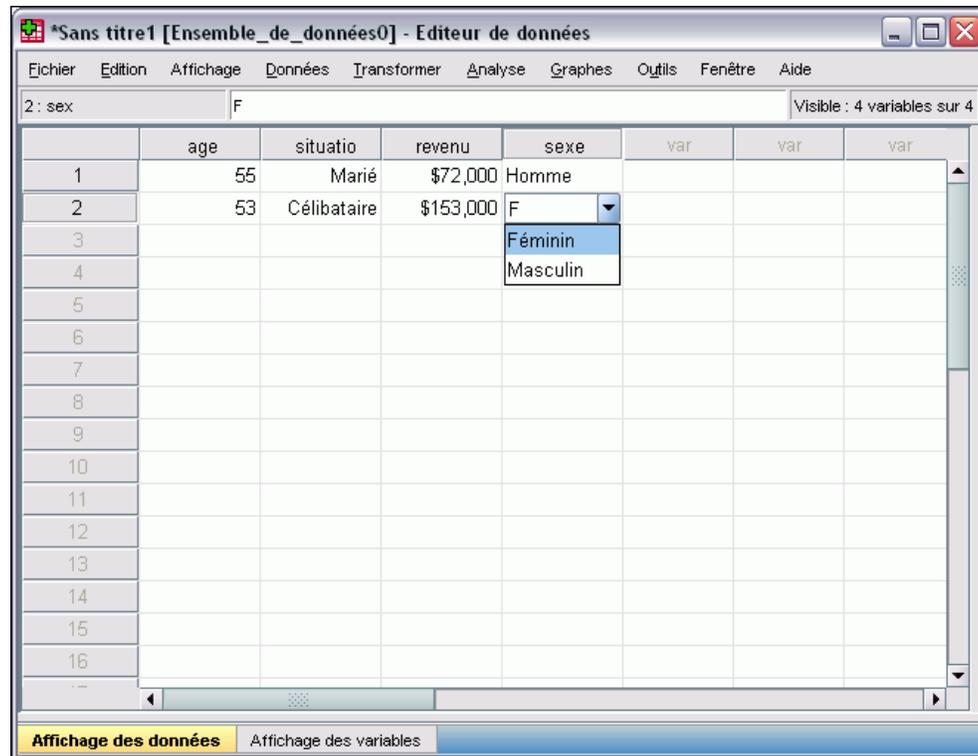
Vous pouvez utiliser des étiquettes de valeurs pour la saisie de données.

- ▶ Cliquez sur l'onglet Affichage des données en bas de la fenêtre de l'éditeur de données.
- ▶ Dans la première ligne, sélectionnez la cellule pour *sexe*.
- ▶ Cliquez sur le bouton à droite de la cellule et choisissez Masculin dans la liste déroulante.
- ▶ Dans la seconde ligne, sélectionnez la cellule pour *sexe*.

- Cliquez sur le bouton à droite de la cellule et choisissez Féminin dans la liste déroulante.

Figure 3-11

Utilisation d'étiquettes de variable pour sélectionner des valeurs



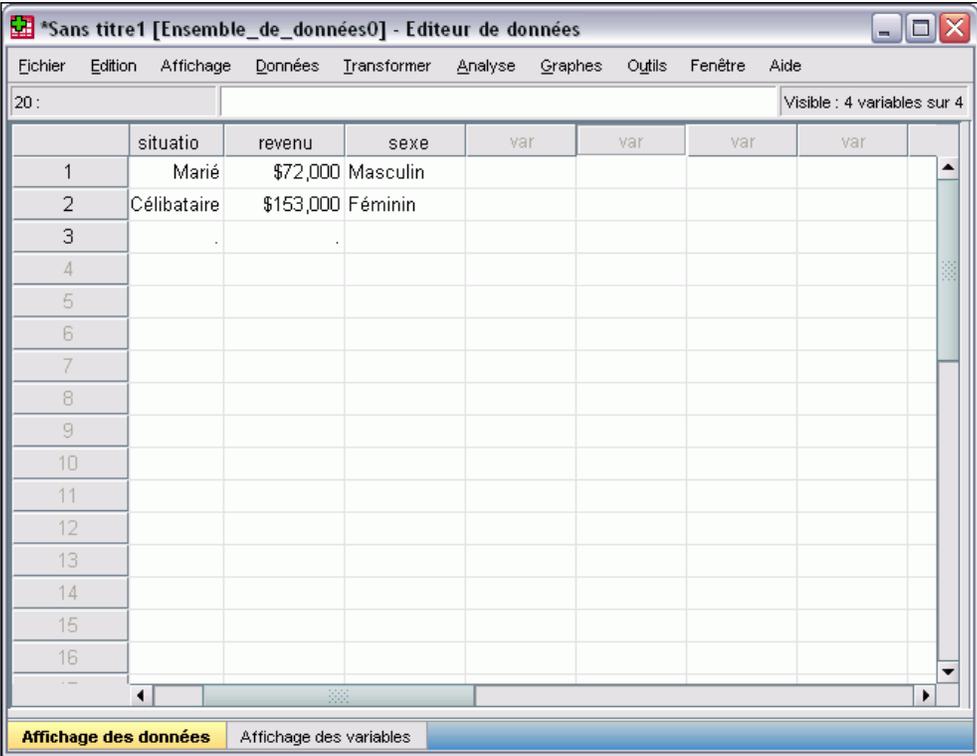
Seules les valeurs définies sont répertoriées ; ainsi, vous êtes sûr que les données saisies sont au format souhaité.

Gestion des données manquantes

Les données manquantes ou non valides sont trop fréquentes pour être ignorées. Les personnes interrogées peuvent refuser de répondre à certaines questions, ne pas connaître la réponse ou donner une réponse dont le format est inattendu. Si vous ne filtrez ou n'identifiez pas ces données, les résultats de votre analyse risquent d'être imprécis.

En ce qui concerne les données numériques, les champs de données vides ou contenant des entrées non valides sont convertis en données manquantes par défaut, signalées par une virgule (Windows français) ou un point (Windows anglais).

Figure 3-12
Valeurs manquantes affichées sous la forme de virgules



The screenshot shows a software window titled '*Sans titre1 [Ensemble_de données0] - Editeur de données'. The window has a menu bar with 'Fichier', 'Edition', 'Affichage', 'Données', 'Transformer', 'Analyse', 'Graphes', 'Outils', 'Fenêtre', and 'Aide'. Below the menu bar, there is a status bar that says '20:' and 'Visible : 4 variables sur 4'. The main area is a data table with the following columns: 'situatio', 'revenu', 'sexe', and four 'var' columns. The first two rows contain data, while the remaining rows have missing values represented by commas. The 'Affichage des données' tab is selected at the bottom.

	situatio	revenu	sexe	var	var	var	var
1	Marié	\$72,000	Masculin				
2	Célibataire	\$153,000	Féminin				
3	.	.	.				
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
--							

La raison pour laquelle une valeur est manquante peut être importante pour votre analyse. Par exemple, vous pouvez juger utile de distinguer les personnes qui ont refusé de répondre à une question de celles qui n'ont pas répondu car cette question ne les concernait pas.

Valeurs manquantes des variables numériques

- ▶ Cliquez sur l'onglet **Affichage des variables** en bas de la fenêtre de l'éditeur de données.
- ▶ Cliquez sur la cellule *Manquante* de la ligne *âge*, puis sur le bouton à droite de la cellule pour ouvrir la boîte de dialogue **Valeurs manquantes**.

Dans cette boîte de dialogue, vous pouvez indiquer jusqu'à trois valeurs manquantes ou un intervalle de valeurs et une valeur discrète supplémentaire.

Figure 3-13
Boîte de dialogue Valeurs manquantes



- ▶ Sélectionnez Valeurs manquantes discrètes.
- ▶ Saisissez 999 dans la première zone de texte. Ne remplissez pas les deux autres zones de texte.
- ▶ Cliquez sur OK pour enregistrer vos modifications et revenir dans l'éditeur de données.

La valeur des données manquantes ayant été ajoutée, vous pouvez appliquer une étiquette à cette valeur.

- ▶ Cliquez sur la cellule *Valeurs* de la ligne *âge*, puis sur le bouton à droite de la cellule pour ouvrir la boîte de dialogue Etiquettes de valeur.
- ▶ Saisissez 999 dans le champ Valeur.
- ▶ Saisissez Non répondu dans le champ Etiquette.

Figure 3-14
Boîte de dialogue Etiquettes de valeurs



- ▶ Cliquez sur Ajouter pour ajouter l'étiquette au fichier de données.
- ▶ Cliquez sur OK pour enregistrer vos modifications et revenir dans l'éditeur de données.

Valeurs manquantes d'une variable chaîne

Les valeurs manquantes des variables de chaîne sont gérées de la même façon que les valeurs manquantes des variables numériques. Cependant, contrairement aux valeurs numériques, les champs vides dans les variables de chaîne ne sont pas désignés comme données manquantes par défaut. Ils sont interprétés comme des chaînes de caractères vides.

- ▶ Cliquez sur l'onglet Affichage des variables en bas de la fenêtre de l'éditeur de données.
- ▶ Cliquez sur la cellule *Manquante* de la ligne *sexe*, puis sur le bouton à droite de la cellule pour ouvrir la boîte de dialogue Valeurs manquantes.
- ▶ Sélectionnez Valeurs manquantes discrètes.
- ▶ Saisissez NR dans la première zone de texte.

Les valeurs manquantes des variables chaîne distinguent les majuscules des minuscules. Par conséquent, la valeur *nr* n'est pas traitée comme une valeur manquante.

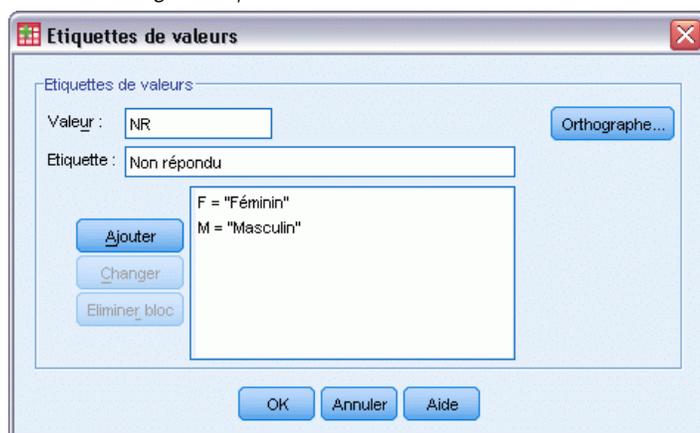
- ▶ Cliquez sur OK pour enregistrer vos modifications et revenir dans l'éditeur de données.

Vous pouvez à présent ajouter une étiquette pour la valeur manquante.

- ▶ Cliquez sur la cellule *Valeurs* de la ligne *sexe*, puis sur le bouton à droite de la cellule pour ouvrir la boîte de dialogue Etiquettes de valeur.
- ▶ Saisissez NR dans le champ Valeur.
- ▶ Saisissez Non répondu dans le champ Etiquette.

Figure 3-15

Boîte de dialogue Etiquettes de valeurs



- ▶ Cliquez sur Ajouter pour ajouter l'étiquette au projet.
- ▶ Cliquez sur OK pour enregistrer vos modifications et revenir dans l'éditeur de données.

Copier-coller des attributs de variable

Une fois les attributs d'une variable définis, vous pouvez les copier et les appliquer à d'autres variables.

- ▶ Dans l'affichage des variables, saisissez *agemar* dans la première cellule de la première ligne vide.

Figure 3-16
Variable *agemar* dans l'affichage des variables

The screenshot shows a window titled '*Sans titre1 [Ensemble_de_données0] - Editeur de données'. The window contains a table with the following columns: 'Nom', 'Type', 'Largeur', 'Décimales', 'Etiquette', and 'Valeurs'. The table has 18 rows. The first five rows are populated with data:

	Nom	Type	Largeur	Décimales	Etiquette	Valeurs
1	age	Numérique	8	0	l'âge du répondant	{999, Pas d...
2	situatio	Numérique	8	0	la situation familiale	{0, Célibatai...
3	revenu	Dollar	12	0	revenu du ménage	Aucun
4	sexe	Chaîne	8	0	Sexe	{F, Femme}...
5	agemar	Numérique	8	2	Age mariage	Aucun
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						

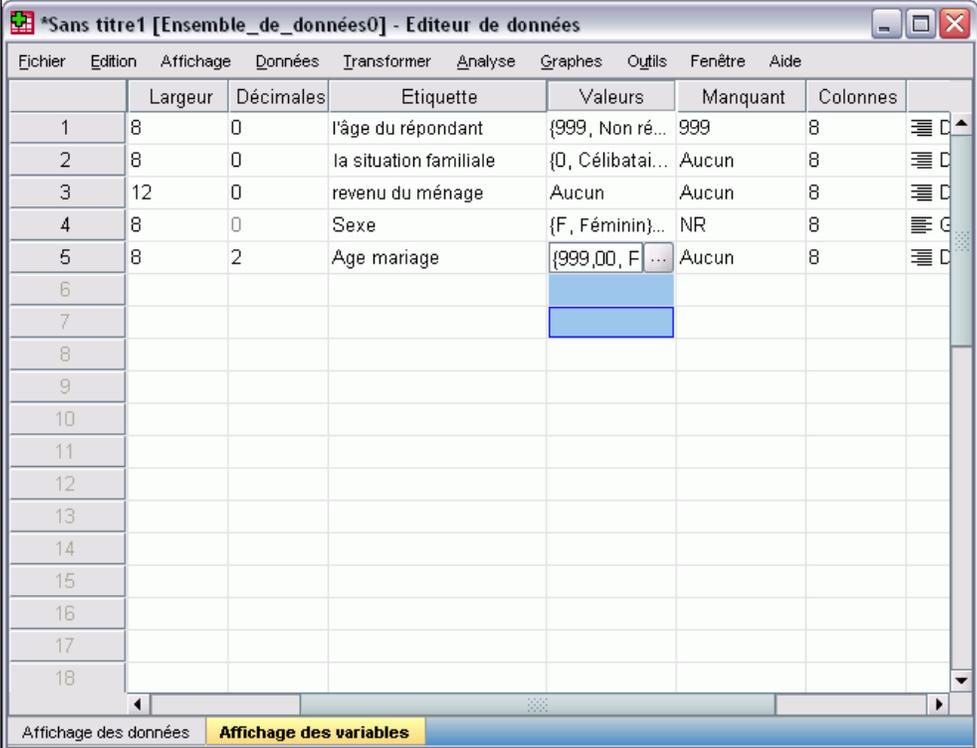
At the bottom of the window, there are two tabs: 'Affichage des données' and 'Affichage des variables', with the latter being the active tab.

- ▶ Dans la colonne *Etiquette*, saisissez Age mariage.
- ▶ Cliquez sur la cellule *Valeurs* de la ligne *age*.
- ▶ A partir des menus, sélectionnez :
Edition > Copier
- ▶ Cliquez sur la cellule *Valeurs* de la ligne *agemar*.
- ▶ A partir des menus, sélectionnez :
Edition > Coller

Les valeurs définies de la variable *age* sont à présent appliquées à la variable *agemar*.

Pour appliquer l'attribut à plusieurs variables, il suffit de sélectionner plusieurs cellules cible (cliquez et faites glisser vers le bas la colonne).

Figure 3-17
Plusieurs cellules sélectionnées



The screenshot shows a window titled '*Sans titre1 [Ensemble_de_données0] - Editeur de données'. The window contains a table with the following columns: 'Largueur', 'Décimales', 'Etiquette', 'Valeurs', 'Manquant', and 'Colonnes'. The rows are numbered 1 to 18. The 'Valeurs' column for rows 6 and 7 is highlighted in blue, indicating selection.

	Largueur	Décimales	Etiquette	Valeurs	Manquant	Colonnes
1	8	0	l'âge du répondant	{999, Non ré...	999	8
2	8	0	la situation familiale	{0, Célibatai...	Aucun	8
3	12	0	revenu du ménage	Aucun	Aucun	8
4	8	0	Sexe	{F, Féminin}...	NR	8
5	8	2	Age mariage	{999,00, F ...	Aucun	8
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						

At the bottom of the window, there are two tabs: 'Affichage des données' and 'Affichage des variables', with the latter being the active tab.

Lorsque vous collez l'attribut, il s'applique à toutes les cellules sélectionnées.

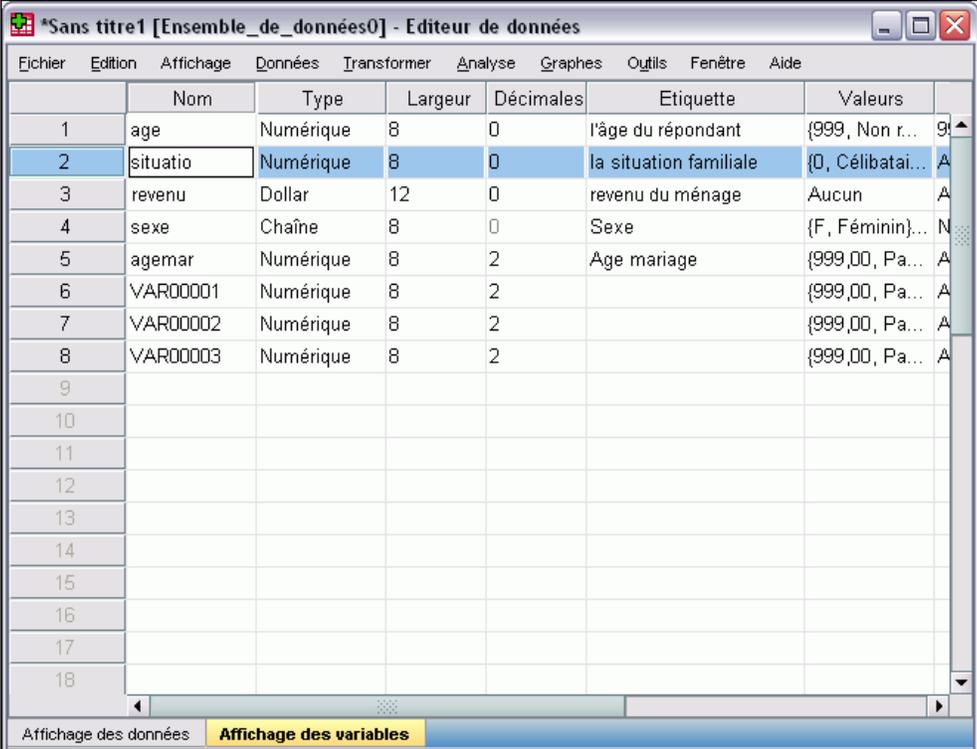
De nouvelles variables sont automatiquement créées si vous collez les valeurs dans des lignes vides.

Pour copier tous les attributs d'une variable vers une autre variable :

- Cliquez sur le numéro de la ligne dans la ligne *situatio*.

Figure 3-18

Ligne sélectionnée



The screenshot shows a window titled '*Sans titre1 [Ensemble_de_données0] - Editeur de données'. The window contains a table with the following data:

	Nom	Type	Largeur	Décimales	Etiquette	Valeurs	
1	age	Numérique	8	0	l'âge du répondant	{999, Non r...	9
2	situatio	Numérique	8	0	la situation familiale	{0, Célibatai...	A
3	revenu	Dollar	12	0	revenu du ménage	Aucun	A
4	sexe	Chaîne	8	0	Sexe	{F, Féminin}...	N
5	agemar	Numérique	8	2	Age mariage	{999,00, Pa...	A
6	VAR00001	Numérique	8	2		{999,00, Pa...	A
7	VAR00002	Numérique	8	2		{999,00, Pa...	A
8	VAR00003	Numérique	8	2		{999,00, Pa...	A
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							

At the bottom of the window, there are two tabs: 'Affichage des données' and 'Affichages des variables', with the latter being active.

- A partir des menus, sélectionnez :
Edition > Copier
- Cliquez sur le numéro de ligne de la première ligne vide.
- A partir des menus, sélectionnez :
Edition > Coller

Tous les attributs de la variable *situatio* sont appliqués à la nouvelle variable.

Figure 3-19

Toutes les valeurs collées dans une ligne

	Nom	Type	Largeur	Décimales	Etiquette	Valeurs
1	age	Numérique	8	0	l'âge du répondant	{999, Non r...}
2	situatio	Numérique	8	0	la situation familiale	{0, Célibatai...}
3	revenu	Dollar	12	0	revenu du ménage	Aucun
4	sexe	Chaîne	8	0	Sexe	{F, Féminin}...
5	agemar	Numérique	8	2	Age mariage	{999,00, Pa...}
6	VAR00001	Numérique	8	2		{999,00, Pa...}
7	VAR00002	Numérique	8	2		{999,00, Pa...}
8	VAR00003	Numérique	8	2		{999,00, Pa...}
9	VAR00004	Numérique	8	0	la situation familiale	{0, Célibatai...}
10						
11						
12						
13						
14						
15						
16						
17						
18						

Définition des propriétés des variables qualitatives

Pour les données qualitatives (nominales, ordinales), vous pouvez utiliser la boîte de dialogue Définir les propriétés de variable afin de définir des étiquettes de valeur et d'autres propriétés de variable. Le processus Définir les propriétés de variable :

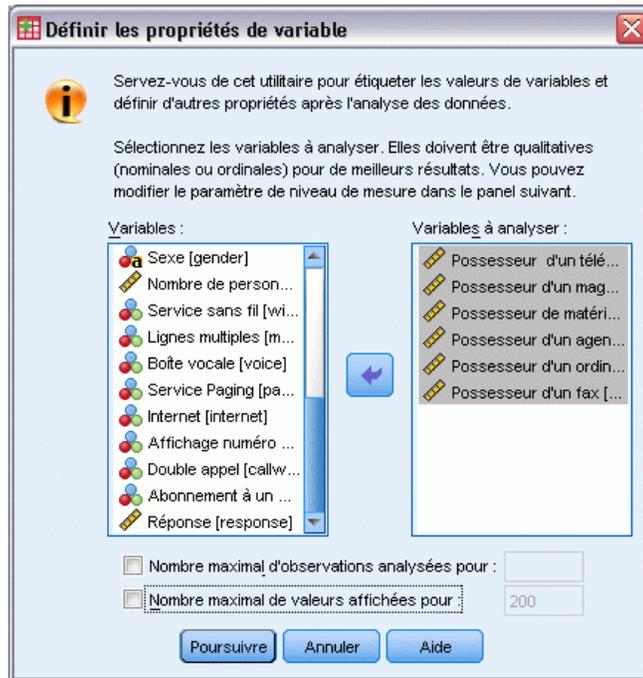
- Analyse les valeurs réelles des données et répertorie toutes les valeurs de données uniques pour chaque variable sélectionnée.
- Identifie les valeurs non étiquetées et possède une fonction d'” étiquetage automatique “.
- Permet de copier des étiquettes de valeurs définies d'une autre variable vers la variable sélectionnée ou de la variable sélectionnée vers d'autres variables.

Cet exemple utilise le fichier de données *demo.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156](#). Ce fichier de données comporte déjà des étiquettes de valeur définies, saisissons donc une valeur à laquelle aucune étiquette de valeur ne correspond.

- Dans l'affichage des données de Data Editor, cliquez sur la première cellule de données de la variable *ownpc* (vous pouvez être amené à faire défiler l'écran vers la droite) et saisissez 99.

- A partir des menus, sélectionnez :
Données > Définir les propriétés de variables

Figure 3-20
Boîte de dialogue initiale Définir les propriétés de variable



Dans la boîte de dialogue initiale Définir les propriétés de variable, sélectionnez les variables nominales ou ordinales pour lesquelles définir des étiquettes de valeur et/ou d'autres propriétés.

- Faites glisser les variables *Possède un ordinateur [ownpc]* à *Possède un magnétoscope [ownvcr]* vers la liste Variables à analyser.

Vous aurez sûrement remarqué que les icônes de niveau de mesure de toutes les variables sélectionnées indiquent qu'il s'agit de variables d'échelle, et non de variables qualitatives. Toutes les variables sélectionnées dans cet exemple sont de véritables variables qualitatives utilisant les valeurs numériques 0 et 1 pour représenter respectivement *Non* et *Oui*. Dans la boîte de dialogue Définir les propriétés de variable, nous allons notamment modifier la propriété de variable du niveau de mesure.

- Cliquez sur Poursuivre.

Figure 3-21
Définir les propriétés de variable, boîte de dialogue principale

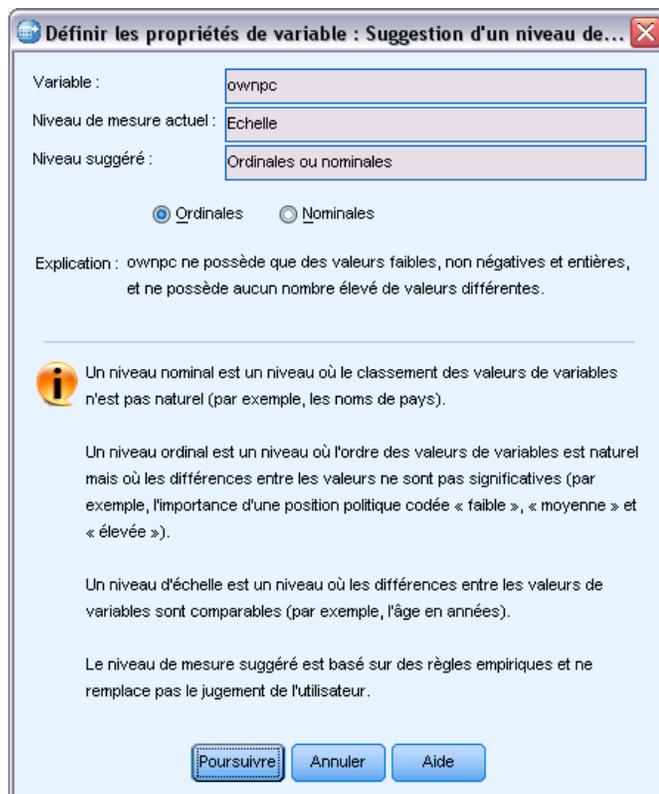
Liste des variables analysées: Variable actuelle : Etiquette :
 Niveau de mesure :
 Rôle :
 Type :
 Largeur : Décimales :
 Valeurs non étiquetées :
 Grille d'étiquettes de valeur : Entrez ou modifiez des étiquettes dans la grille. Vous pouvez saisir des valeurs supplémentaires dans la partie inférieure.

	Modifié	Manquant	Effectif	Valeur	Etiquette
1	<input type="checkbox"/>	<input type="checkbox"/>	3588	0	Non
2	<input type="checkbox"/>	<input type="checkbox"/>	2811	1	Oui
3	<input type="checkbox"/>	<input type="checkbox"/>	1	99	
4	<input type="checkbox"/>	<input type="checkbox"/>			

Observations analysées : Copier les propriétés:
 Limite de la liste de valeurs : Valeurs non étiquetées:

- ▶ Dans la liste des variables analysées, sélectionnez *ownpc*.
Le niveau de mesure actuel de la variable sélectionnée est échelle. Vous pouvez sélectionner un autre niveau de mesure dans la liste déroulante ou consulter ceux proposés par la boîte de dialogue Définir les propriétés de variable.
- ▶ Cliquez sur Suggérer.
La boîte de dialogue Suggestion d'un niveau de mesure est affichée.

Figure 3-22
Boîte de dialogue Suggestion d'un niveau de mesure



Etant donné que la variable comporte peu de valeurs différentes et que toutes les observations analysées contiennent des valeurs entières, vous devez choisir un niveau de mesure ordinal ou nominal.

- Sélectionnez Ordinal, puis cliquez sur Continuer.

La variable sélectionnée comporte à présent un niveau de mesure ordinal.

La grille des étiquettes de valeurs affiche toutes les valeurs de données uniques de la variable sélectionnée, toutes les étiquettes de valeur définies pour ces valeurs, ainsi que le nombre d'occurrences (effectif) de chaque valeur dans les observations analysées.

La valeur saisie dans l'affichage de données, 99, figure dans la grille. L'effectif est de 1 car nous n'avons modifié la valeur que pour une observation. Quant à la colonne *Etiquette*, elle est vide car nous n'avons pas encore défini d'étiquette de valeur pour 99. Le X dans la première colonne de la liste des variables analysées indique par ailleurs que la variable sélectionnée comporte au moins une valeur observée sans étiquette de valeur définie.

- Dans la colonne *Etiquette* de la valeur 99, saisissez Pas de réponse.
- Cochez la case dans la colonne *Manquante* pour la valeur 99 afin d'identifier la valeur 99 comme **valeur utilisateur manquante**.

Les valeurs des données définies comme valeurs utilisateur manquantes sont repérées par un indicateur en vue d'un traitement spécial et sont exclues de la plupart des calculs.

Figure 3-23
Nouvelles propriétés définies pour la variable *ownpc*

Liste des variables analysées: No. Me. Rôle Variable
 ownrtv ownrvcr ownncd ownpda ownpc ownrfax

Variable actuelle : Etiquette :
 Niveau de mesure : Suggérer
 Rôle :
 Valeurs non étiquetées :
 Type : Largeur : Décimales :
 Attributs...

Grille d'étiquettes de valeur : Entrez ou modifiez des étiquettes dans la grille. Vous pouvez saisir des valeurs supplémentaires dans la partie inférieure.

	Modifié	Manquant	Effectif	Valeur	Etiquette
1	<input type="checkbox"/>	<input type="checkbox"/>	3588	0	Non
2	<input type="checkbox"/>	<input type="checkbox"/>	2811	1	Oui
3	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	1	99	Pas de réponse.
4	<input type="checkbox"/>	<input type="checkbox"/>			

Observations analysées : Copier les propriétés: Valeurs non étiquetées:
 Limite de la liste de valeurs :

Avant de terminer la modification des propriétés de la variable *ownpc*, appliquons un niveau de mesure, des étiquettes de valeur et des définitions de valeurs manquantes identiques aux autres variables de la liste.

- Dans la zone Copier les propriétés, cliquez sur Vers d'autres variables.

Figure 3-24
Boîte de dialogue Appliquer les étiquettes et le niveau

Sélectionner la variable dont les étiquettes de valeur, le niveau de mesure, le rôle, la définition de la valeur manquante et le format seront copiés dans la variable actuelle. (Les étiquettes de valeur et les valeurs manquantes sont fusionnées).

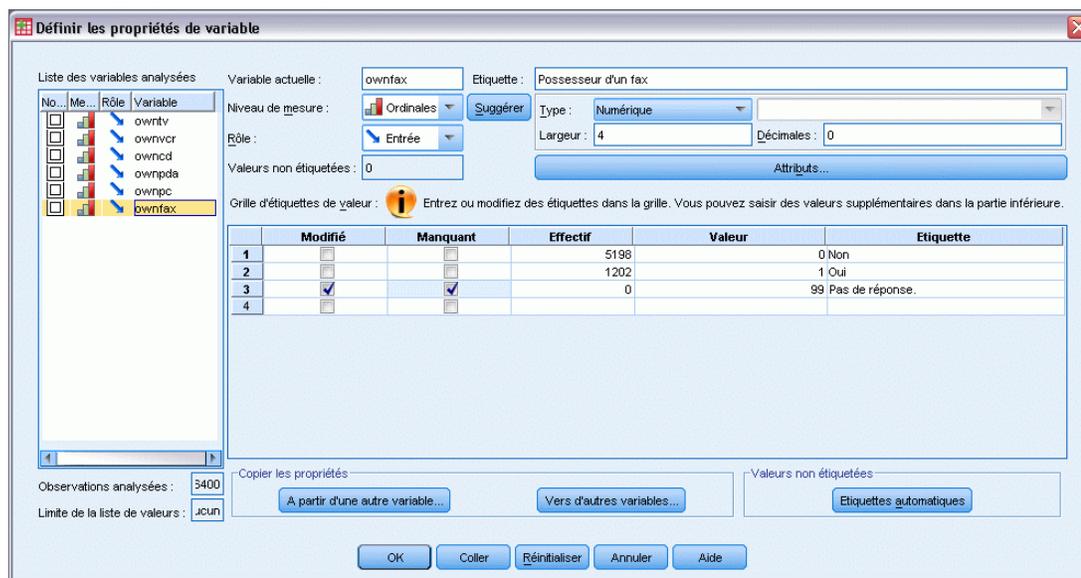
Variables :

- Possède un téléviseur [ownrtv]
- Possède un magnétoscope [ownrvcr]
- Possède matériel HiFi [ownncd]
- Possède un agenda électronique [ownpda]
- Possède un fax [ownrfax]

- Dans la boîte de dialogue Appliquer les étiquettes et le niveau, sélectionnez toutes les variables de la liste, puis cliquez sur Copier.

Si vous sélectionnez l'une des autres variables de la liste des variables analysées de la boîte de dialogue principale Définir les propriétés de variable, vous constaterez qu'il s'agit désormais d'une variable ordinale avec une valeur de 99 définie comme valeur manquante utilisateur et une étiquette de valeur *Pas de réponse*.

Figure 3-25
Nouvelles propriétés définies pour la variable ownfax



- Cliquez sur OK pour enregistrer toutes les propriétés de variable définies.

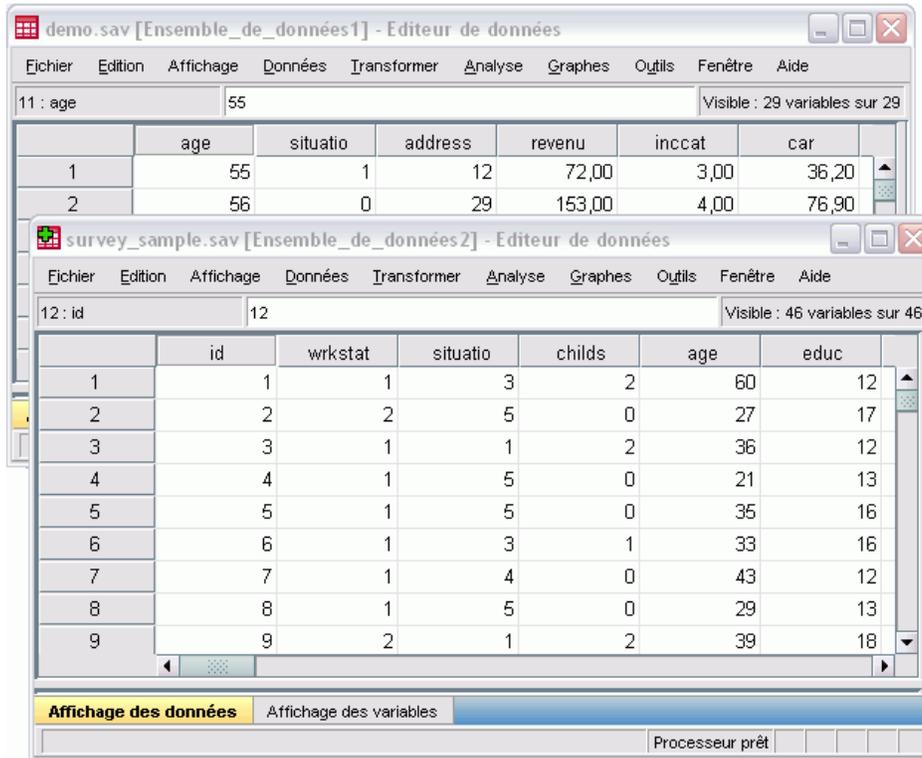
Utilisation des sources de données multiples

Avec la version 14.0, plusieurs sources de données peuvent être ouvertes en même temps, ce qui permet de facilement :

- Basculer entre les sources de données.
- Comparer les contenus des différentes sources de données.
- Copier et coller les données entre les sources de données.
- Créer de multiples sous-ensembles d'observations et/ou de variables pour analyse.
- Fusionner plusieurs sources de données à partir de différents formats de données (par exemple des feuilles de calcul, des bases de données ou des données texte) sans avoir à enregistrer préalablement chaque source de données.

Manipulation de base de plusieurs sources de données

Figure 4-1
Deux sources de données s'ouvrent en même temps



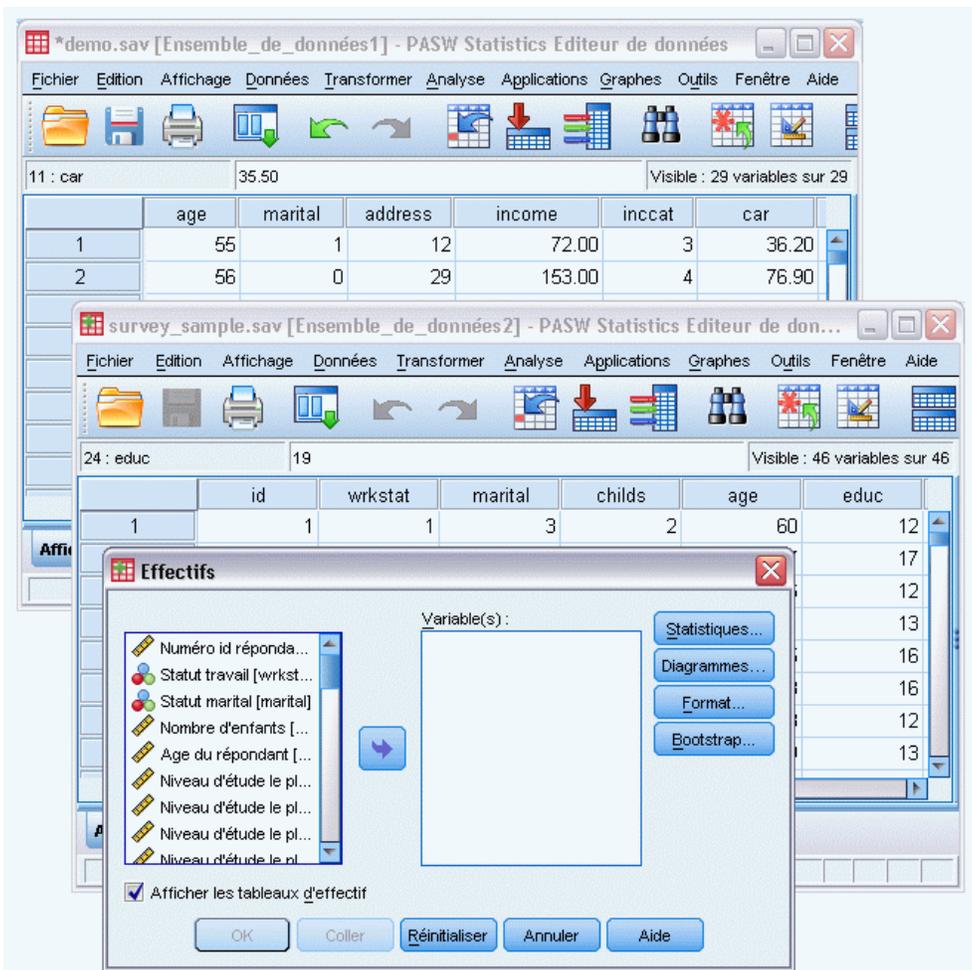
Par défaut, chaque source de données que vous ouvrez est affichée dans une nouvelle fenêtre de l'éditeur de données.

- Toute source de données ouverte au préalable le reste pour utilisation ultérieure.
- Lorsque vous ouvrez d'abord une source de données, elle devient automatiquement **l'ensemble de données actif**.
- Vous pouvez modifier l'ensemble de données actif en cliquant simplement n'importe où dans la fenêtre Data Editor de la source de données que vous souhaitez utiliser ou en sélectionnant la fenêtre Data Editor pour cette source de données à partir du menu Fenêtre.

- Seules les variables de l'ensemble de données actif sont disponibles pour analyse.

Figure 4-2

Liste de variables contenant les variables dans l'ensemble de données actif



- Vous ne pouvez pas modifier l'ensemble de données actif lorsque une boîte de dialogue ayant accès aux données est ouverte (y compris toutes les boîtes de dialogue qui affichent des listes de variables).
- Au moins une fenêtre Data Editor doit être ouverte lors d'une session. Lorsque vous fermez la dernière fenêtre de l'éditeur de données, IBM® SPSS® Statistics se ferme automatiquement, vous invitant au préalable à enregistrer vos modifications.

Utilisation de plusieurs ensembles de données dans une syntaxe de commande

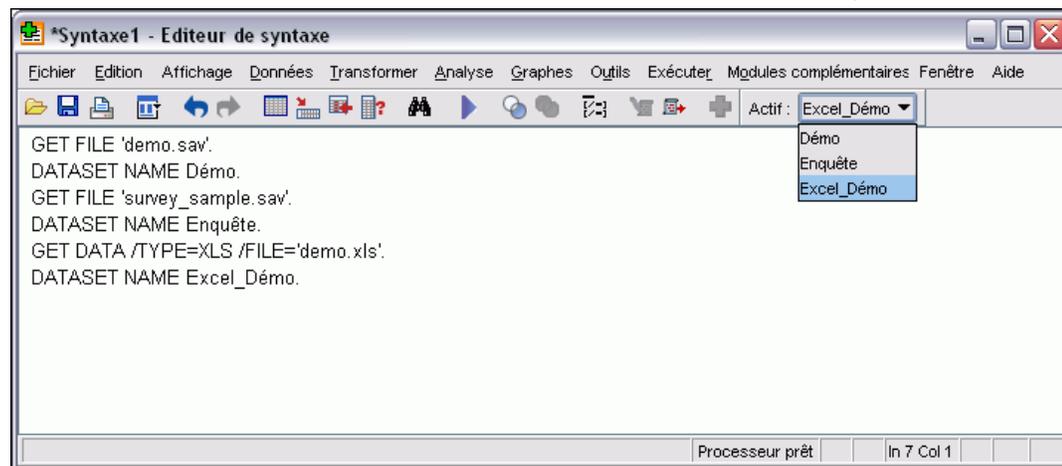
Si vous utilisez la syntaxe de commande pour ouvrir les sources de données (par exemple GET FILE ou GET DATA), vous devez utiliser la commande DATASET NAME pour nommer de façon explicite chaque ensemble de données afin d'avoir plusieurs sources de données ouvertes simultanément.

Lorsque vous utilisez une syntaxe de commande, le nom de l'ensemble de données actif apparaît dans la barre d'outils de la fenêtre de syntaxe. Toutes les actions suivantes permettent d'activer un autre ensemble de données :

- Utiliser la commande DATASET ACTIVATE.
- Cliquer n'importe où dans la fenêtre de l'éditeur de données d'un ensemble de données.
- Sélectionner le nom d'un ensemble de données dans la barre d'outils de la fenêtre de syntaxe.

Figure 4-3

Ensembles de données ouverts affichés dans la barre d'outils de la fenêtre de syntaxe



Copie et collage d'informations entre les ensembles de données

Vous pouvez copier à la fois des données et des attributs de définition de variable d'un ensemble de données à un autre de la même façon que vous copiez et collez des informations au sein d'un fichier de données unique.

- Le fait de copier et coller des cellules de données sélectionnées dans Affichage des données ne colle que les valeurs des données, sans les attributs de définition de variable.
- Le fait de copier et coller une variable entière dans Affichage des données en sélectionnant le nom de la variable en haut de la colonne colle toutes les données et tous les attributs de définition de variable pour cette variable.
- Le fait de copier et coller des attributs de définition de variable ou des variables entières dans Affichage des variables colle les attributs sélectionnés (ou la définition de la variable entière) mais ne colle pas les valeurs des données.

Attribution d'un nouveau nom aux ensembles de données

Lorsque vous ouvrez une source de données via les menus et les boîtes de dialogue, un nom d'ensemble de données *DataSet* est automatiquement attribué à chaque source de données ; *n* est une valeur entière séquentielle . Lorsque vous ouvrez une source de données via la syntaxe de commande, aucun nom d'ensemble de données n'est attribué à moins que vous en spécifiez un avec `DATASET NAME`. Pour fournir des noms d'ensembles de données plus descriptifs :

- ▶ Depuis les menus de la fenêtre Data Editor pour l'ensemble de données dont vous souhaitez modifier le nom, choisissez :
Fichier > Renommer l'ensemble de données...
- ▶ Entrez un nouveau nom conforme aux règles de dénomination des variables pour l'ensemble de données.

Suppression de plusieurs ensembles de données

Si vous préférez qu'un seul ensemble de données soit disponible à un moment donné et souhaitez supprimer la fonctionnalité d'ouverture de plusieurs ensembles de données :

- ▶ A partir des menus, sélectionnez :
Affichage > Options
- ▶ Cliquez sur l'onglet Général.

Sélectionnez (cochez) la case Ouvrir un seul ensemble de données à la fois.

Examen des statistiques récapitulatives pour chaque variable

Ce chapitre traite des mesures récapitulatives simples et de la façon dont le niveau de mesure d'une variable influence le type de statistiques devant être utilisé. Nous utiliserons le fichier de données *demo.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.](#)

Niveau de mesure

Différentes mesures récapitulatives sont adaptées à différents types de données, selon le niveau de mesure :

Qualitatives : Données ayant un nombre limité de valeurs ou de modalités distinctes (par exemple, sexe ou situation de famille). Elles sont parfois également qualifiées de **données qualitatives**. Les variables qualitatives peuvent être des données chaîne (alphanumérique) ou des variables numériques qui utilisent des codes chiffrés pour représenter les modalités (par exemple, 0 = *Célibataire* et 1 = *Marié*). Il existe deux types essentiels de données qualitatives :

- **Nominal**. Données qualitatives dont les modalités n'ont aucun ordre inhérent. Par exemple, une modalité d'emploi de type *ventes* n'est pas supérieure ou inférieure à une modalité d'emploi de type *marketing* ou *étude*.
- **Ordinal**. Données qualitatives dont les modalités possèdent un ordre significatif, mais pour lesquelles il n'existe aucune distance mesurable entre les modalités. Par exemple, les valeurs *élevée*, *moyenne* et *faible* doivent être classées dans un certain ordre, mais il est impossible de calculer la « distance » entre ces valeurs.

Echelle. Données mesurées sur une échelle d'intervalle ou de rapport, où les valeurs de données indiquent à la fois l'ordre des valeurs et la distance qui les sépare. Par exemple, un salaire de 58 160 € est supérieur à un salaire de 42 212 € et la distance entre les deux valeurs est de 15 948 €. Ces données sont aussi appelées données **quantitatives** ou données **continues**.

Mesures récapitulatives pour données qualitatives

Pour les données qualitatives, la mesure récapitulative la plus courante est le nombre ou le pourcentage d'observations dans chaque modalité. Le **mode** est la modalité ayant le plus grand nombre d'observations. Pour les données ordinales, la **médiane** (valeur au-dessus ou au-dessous

de laquelle se trouve la moitié des observations) peut également être une mesure récapitulative utile s'il existe un grand nombre de modalités.

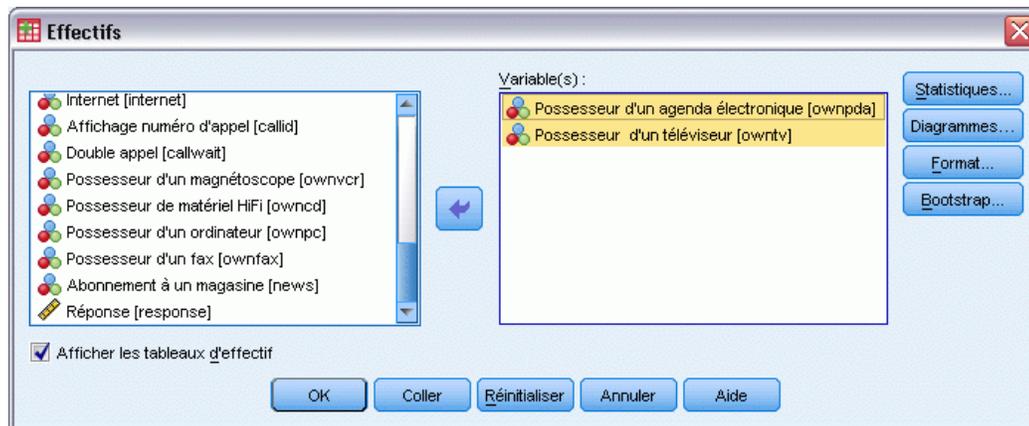
La procédure Fréquences produit des tableaux de fréquences qui affichent le nombre et le pourcentage d'observations pour chaque valeur observée d'une variable.

- A partir des menus, sélectionnez :
Analyse > Statistiques descriptives > Effectifs

Remarque : Cette fonction nécessite l'option Statistiques de base.

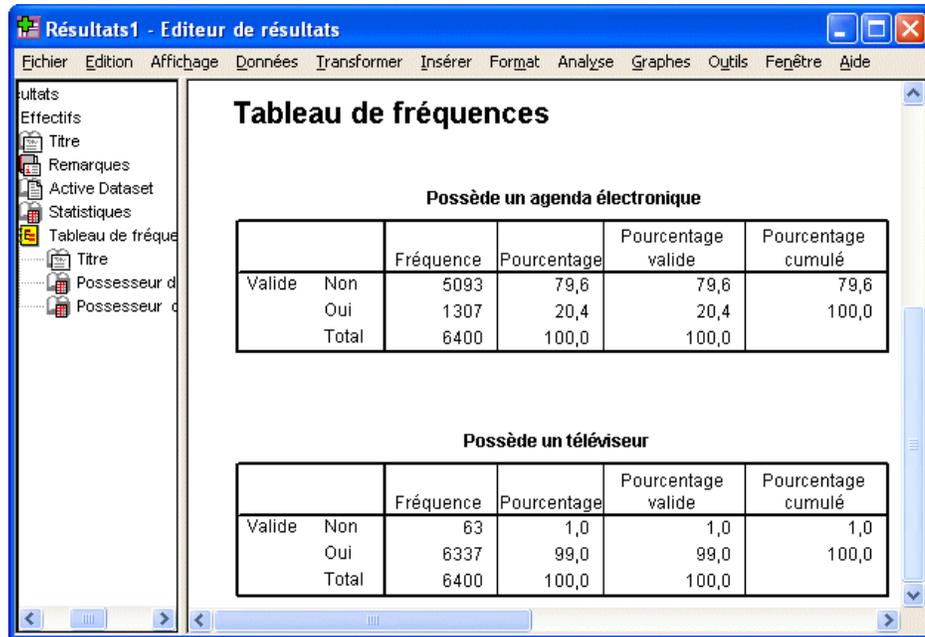
- Sélectionnez *Possède un agenda électronique [pda]* et *Possède un téléviseur [tv]*, et déplacez-les vers la liste Variable(s).

Figure 5-1
Variables qualitatives sélectionnées pour l'analyse



- Cliquez sur OK pour exécuter la procédure.

Figure 5-2
Tableaux de fréquences



Les tableaux de fréquences apparaissent dans la fenêtre du Viewer. Les tableaux d'effectifs révèlent que seuls 20,4 % des personnes possèdent un agenda électronique, mais que la quasi-totalité possèdent une télévision (99,0 %). Ces informations ne semblent pas vraiment pertinentes, mais il peut être intéressant d'en savoir plus sur le petit groupe de personnes qui ne possèdent pas de télévision.

Diagrammes pour données qualitatives

Vous pouvez afficher graphiquement les informations dans un tableau de fréquences avec un diagramme en bâtons ou un diagramme en secteurs.

- Ouvrez à nouveau la boîte de dialogue Fréquences. (Les deux variables doivent toujours être sélectionnées.)

Vous pouvez utiliser le bouton Rappeler boîte de dialogue de la barre d'outils pour revenir rapidement aux dernières procédures utilisées.

Figure 5-3
Bouton Rappeler boîte de dialogue



- Cliquez sur Diagrammes.

- Cliquez sur Diagrammes en bâtons, puis sur Poursuivre.

Figure 5-4

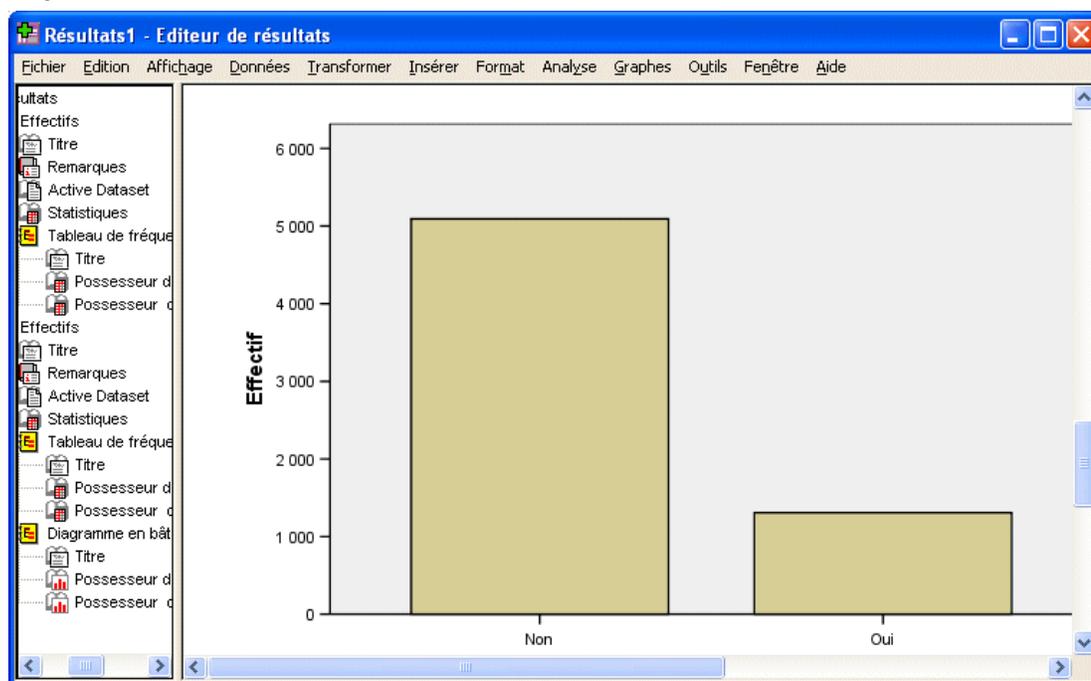
Boîte de dialogue Fréquences : Diagrammes



- Cliquez sur OK dans la boîte de dialogue principale pour exécuter la procédure.

Figure 5-5

Diagramme en bâtons



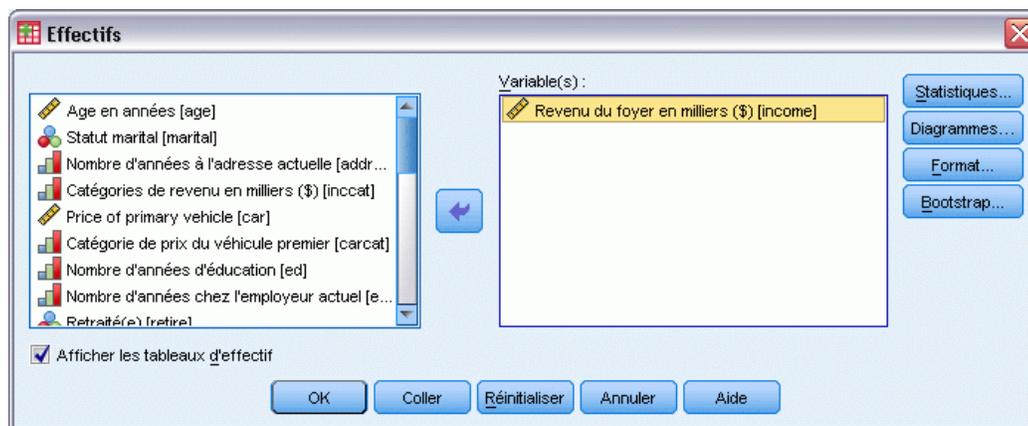
Outre les tableaux de fréquences, les mêmes informations sont à présent affichées sous forme de diagrammes en bâtons ; vous pouvez ainsi voir que la plupart des personnes n'ont pas d'agenda électronique alors que la quasi-totalité d'entre elles possèdent une télévision.

Mesures récapitulatives pour variables d'échelle

De nombreuses mesures récapitulatives sont disponibles pour les variables d'échelle, dont :

- **Mesures de la tendance centrale.** Les mesures les plus courantes de la tendance centrale sont la **moyenne** (moyenne arithmétique) et la **médiane** (valeur au-dessus ou au-dessous de laquelle se trouve la moitié des observations).
 - **Mesures de la dispersion.** Les statistiques qui mesurent la quantité de variation ou de dispersion dans les données comprennent l'écart-type, minimal et maximal.
- Ouvrez à nouveau la boîte de dialogue Fréquences.
- Cliquez sur Réinitialiser pour effacer les paramètres précédents.
- Sélectionnez la variable *Revenu du ménage en milliers [revenu]* et déplacez-la dans la liste Variable(s).

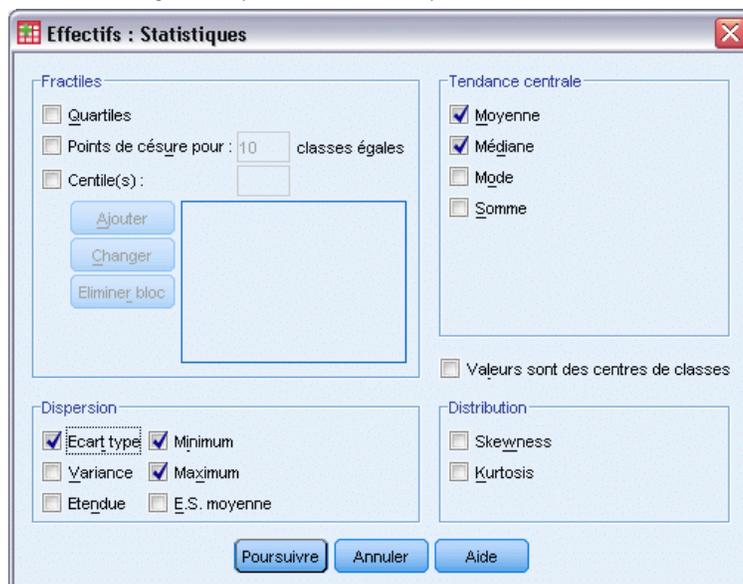
Figure 5-6
Variables d'échelle sélectionnées pour l'analyse



- Cliquez sur Statistiques.

- Sélectionnez Moyenne, Médiane, Ecart type, Minimum et Maximum.

Figure 5-7
Boîte de dialogue Fréquences : Statistiques



- Cliquez sur Poursuivre.
- Désélectionnez Afficher les tableaux d'effectif dans la boîte de dialogue principale Effectifs. (En général, les tableaux de fréquences ne sont pas très utiles pour les variables d'échelle car il peut exister presque autant de valeurs distinctes que d'observations dans le fichier de données.)
- Cliquez sur OK pour exécuter la procédure.

Le tableau statistique de fréquences est affiché dans la fenêtre du Viewer.

Figure 5-8
Tableau statistique de fréquences

Statistiques		
Revenu du ménage en milliers (revenu)		
N	Valide	6400
	Manquante	0
Moyenne		69,4748
Médiane		45,0000
Ecart-type		78,71856
Minimum		9,00
Maximum		1116,00

Dans cet exemple, la différence entre la moyenne et la médiane est importante. La moyenne est plus importante que la médiane de quasiment 25 000, ce qui indique que les valeurs ne sont pas distribuées normalement. Vous pouvez vérifier visuellement la distribution grâce à un histogramme.

Histogrammes pour variables d'échelle

- ▶ Ouvrez à nouveau la boîte de dialogue Fréquences.
- ▶ Cliquez sur Diagrammes.
- ▶ Cliquez sur Histogrammes et Avec courbe gaussienne.

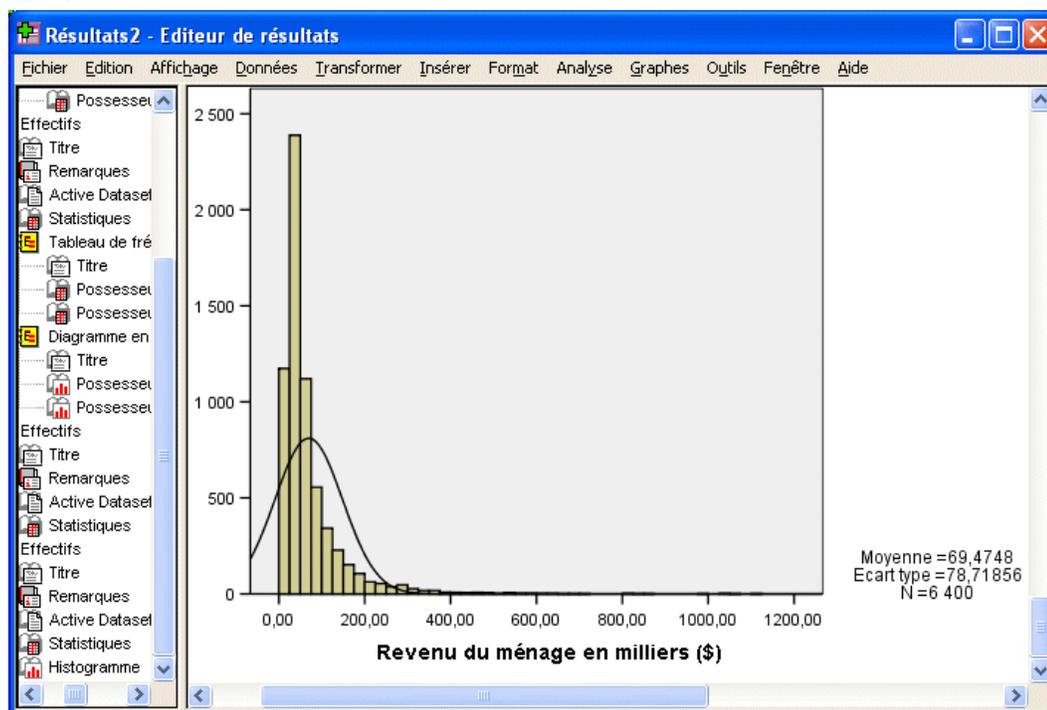
Figure 5-9

Boîte de dialogue Fréquences : Diagrammes



- ▶ Cliquez sur Poursuivre, puis sur OK dans la boîte de dialogue principale pour exécuter la procédure.

Figure 5-10
Histogramme



La grande majorité des observations est regroupée au bas de l'échelle, la plupart se trouvant au-dessous de 100 000. Quelques observations, cependant, se trouvent dans l'intervalle 500 000 et au-delà (elles sont si peu nombreuses que vous devez modifier l'historgramme pour les voir). Ces valeurs très élevées pour quelques observations seulement ont un effet significatif sur la moyenne mais peu d'effet, voire aucun, sur la médiane ; cela signifie que, dans cet exemple, la médiane est un meilleur indicateur de la tendance centrale.

Création et modification de diagrammes

Vous pouvez créer et modifier des types de diagramme divers et variés. Dans ce chapitre, nous allons créer et modifier des diagrammes en bâtons. Vous pouvez appliquer les principes à n'importe quel type de diagramme.

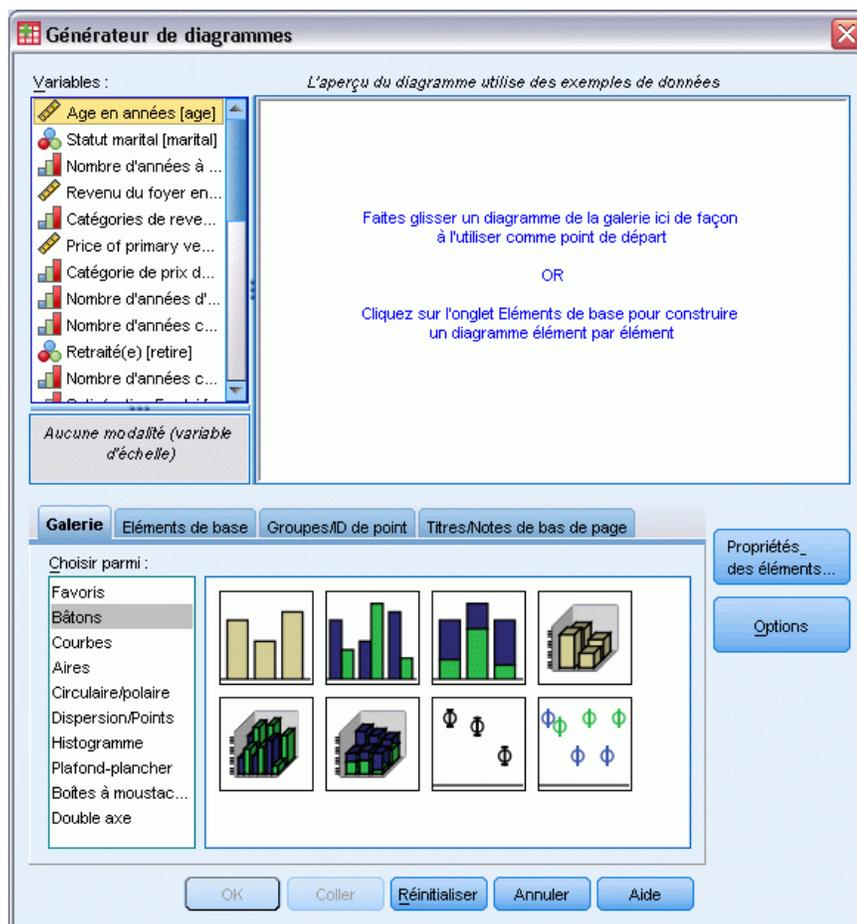
Création de diagrammes - Notions de base

Pour illustrer les notions de base de la création de diagrammes, nous allons créer un diagramme en bâtons du revenu moyen pour plusieurs niveaux de satisfaction professionnelle. Cet exemple utilise le fichier de données *demo.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.](#)

- ▶ A partir des menus, sélectionnez :
Graphes > Générateur de diagrammes...

La boîte de dialogue Générateur de diagrammes est une fenêtre interactive qui vous permet d'obtenir l'aperçu d'un diagramme avant que vous ne le génériez.

Figure 6-1
Boîte de dialogue Générateur de diagrammes



Utilisation de la galerie Générateur de diagrammes

- Cliquez sur l'onglet Galerie s'il n'est pas sélectionné.

La galerie inclut plusieurs diagrammes différents prédéfinis, qui sont organisés par type de diagramme. L'onglet Eléments de base fournit également des éléments de base (comme les axes et les éléments graphiques) pour créer des diagrammes en partant de zéro, mais il est plus facile d'utiliser la galerie.

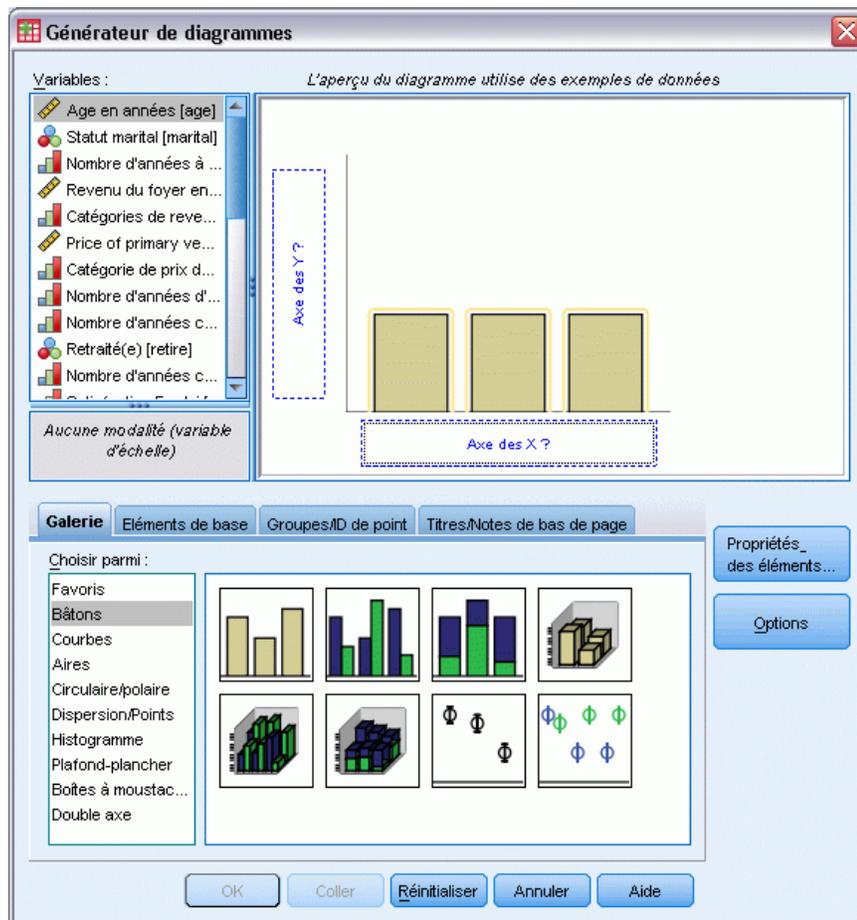
- Cliquez sur Bâton s'il n'est pas sélectionné.

Les icônes représentant les diagrammes en bâtons disponibles dans la galerie apparaissent dans la boîte de dialogue. Les images doivent fournir suffisamment d'informations pour identifier le type de diagramme spécifique. Si vous avez besoin de plus d'informations, vous pouvez également afficher une description de la note d'aide du diagramme en posant votre curseur sur une icône.

- Faites glisser l'icône du diagramme en bâtons simples sur le « canevas », qui est en fait la zone étendue au-dessus de la galerie. Le Générateur de diagrammes affiche un aperçu du diagramme sur le canevas. Notez que les données utilisées pour dessiner le diagramme ne sont pas vos données en cours. Il s'agit d'exemples.

Figure 6-2

Diagramme en bâtons sur le canevas du Générateur de diagrammes



Définition des variables et des statistiques

Bien qu'il y ait un diagramme sur le canevas, il n'est pas complet car il n'y a aucune variable ou statistique pour contrôler la hauteur des bâtons et pour spécifier la catégorie de variable correspondant à chaque bâton. Vous ne pouvez pas avoir de diagramme sans variable et statistique. Vous pouvez ajouter des variables en les glissant de la liste Variables qui se trouve à gauche du canevas.

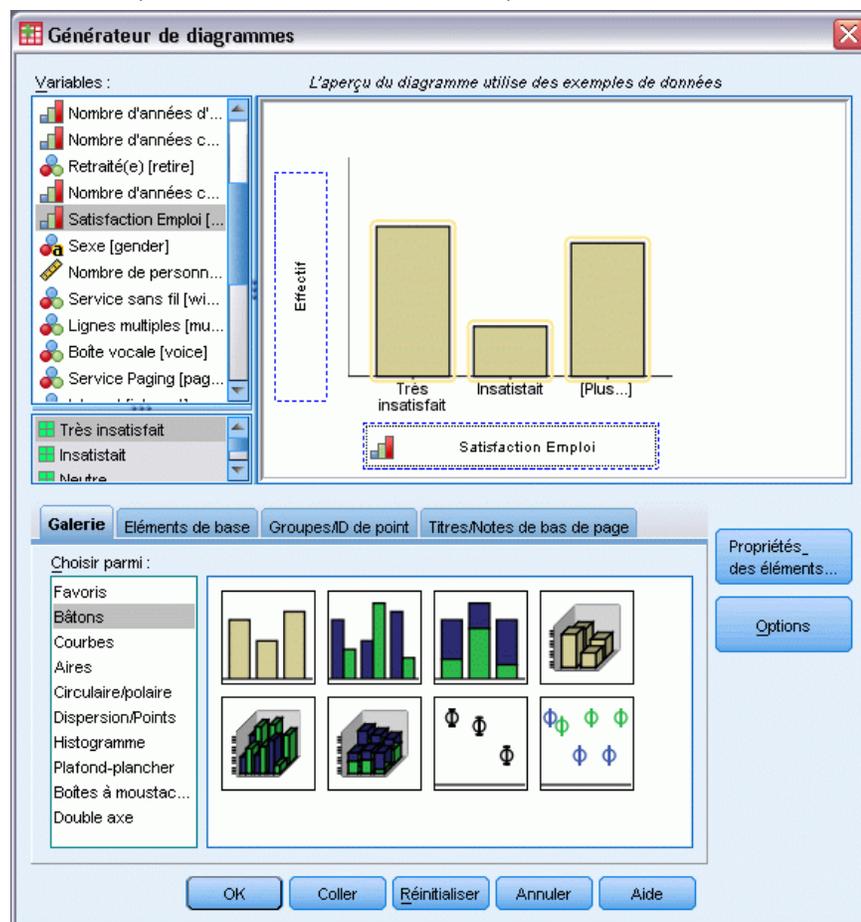
Le niveau de mesure d'une variable est important dans le Générateur de diagrammes. Vous allez utiliser la variable *Satisfaction professionnelle* de l'axe *x*. Cependant, l'icône (qui ressemble à une règle) à côté de la variable indique que son niveau de mesure est défini en tant que variable

d'échelle. Pour créer le diagramme correct, vous devez utiliser un niveau de mesure qualitatif. Plutôt que de revenir et de modifier le niveau de mesure dans l’Affichage des variables, vous pouvez le modifier temporairement dans le Générateur de diagrammes.

- ▶ Cliquez avec le bouton droit de la souris sur *Satisfaction professionnelle* dans la liste Variables et choisissez Ordinal. Ordinal correspond à un niveau de mesure approprié car les catégories dans *Satisfaction professionnelle* peuvent être ordonnées par niveau de satisfaction. Notez que l’icône change une fois le niveau de mesure modifié.
- ▶ Faites glisser *Satisfaction professionnelle* de la liste Variables dans la zone de déplacement de l’axe x.

Figure 6-3

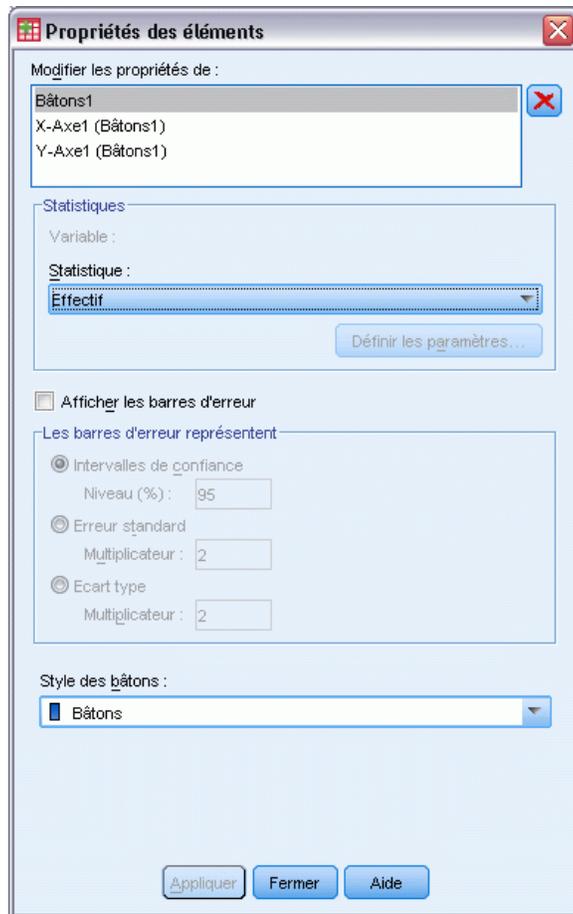
Satisfaction professionnelle dans la zone de déplacement de l’axe x.



La zone de déplacement de l’axe y prend par défaut la statistique *Effectif*. Si vous souhaitez utiliser une autre statistique (comme pourcentage ou moyenne), vous pouvez facilement en changer. Vous n’utiliserez aucune de ces statistiques dans cet exemple, mais nous allons revoir le processus au cas où vous devez changer cette statistique plus tard.

- ▶ Cliquez sur Propriété des éléments pour afficher la fenêtre Propriété des éléments.

Figure 6-4
Fenêtre Propriété des éléments



La fenêtre Propriété des éléments vous permet de modifier les propriétés des différents éléments du diagramme. Ces éléments comprennent les éléments graphiques (comme les bâtons du diagramme) et les axes du diagramme. Sélectionnez un des éléments dans Modifier les propriétés de liste pour modifier les propriétés associées à cet élément. Notez également le X rouge situé à droite de la liste. Ce bouton supprime un élément graphique du canevas. Puisque Bâton1 est sélectionné, les propriétés affichées s'appliquent aux éléments graphiques, en particulier les éléments de bâtons.

La liste déroulante Statistiques affiche les statistiques spécifiques qui sont disponibles. Les mêmes statistiques sont habituellement disponibles pour chaque type de diagramme. Notez que certaines statistiques requièrent que la zone de déplacement de l'axe y contienne une variable.

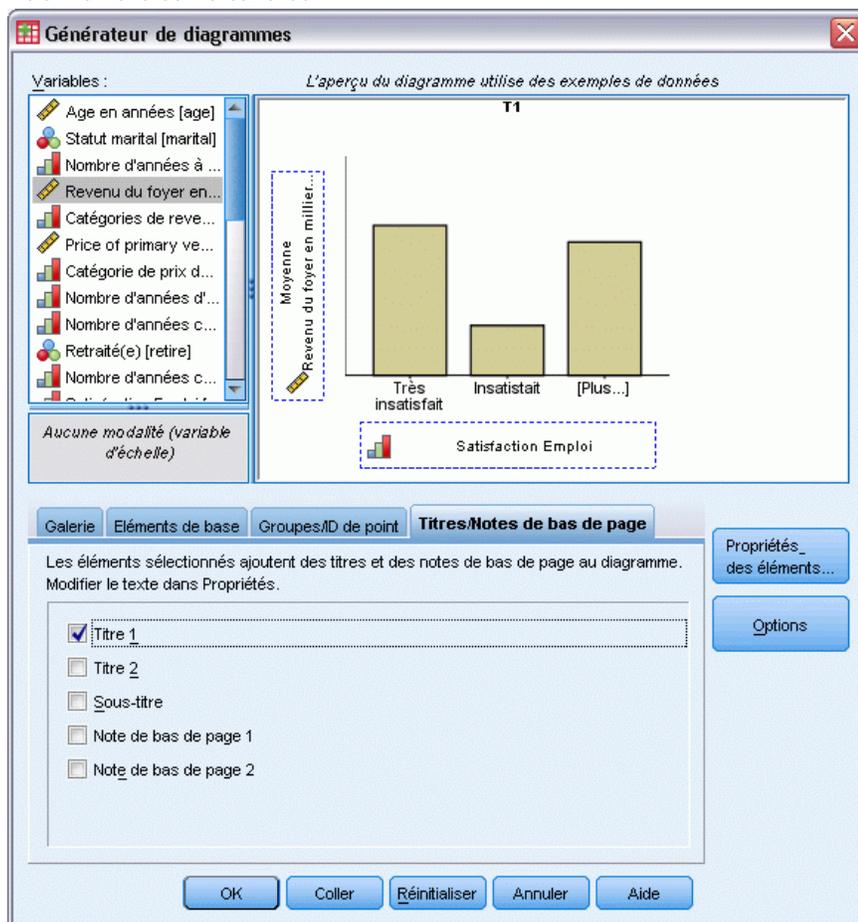
- Revenez à la boîte de dialogue Générateur de diagrammes et faites glisser *Revenu du ménage en milliers* de la liste Variables vers la zone de déplacement de l'axe y. Puisque la variable sur l'axe y est sous forme d'échelle et que la variable de l'axe x est qualitative (ordinal est un type de niveau de mesure qualitatif), la zone de déplacement de l'axe y prend par défaut la statistique *Moyenne*. Il s'agit des variables et des statistiques souhaitées, il n'y a donc aucun besoin de modifier les propriétés de l'élément.

Ajout de texte

Vous pouvez également ajouter des titres et des notes de bas de page au diagramme.

- ▶ Cliquez sur l'onglet Titres/Notes de bas de page.
- ▶ Sélectionnez Titre 1.

Figure 6-5
Titre 1 affiché sur le canevas



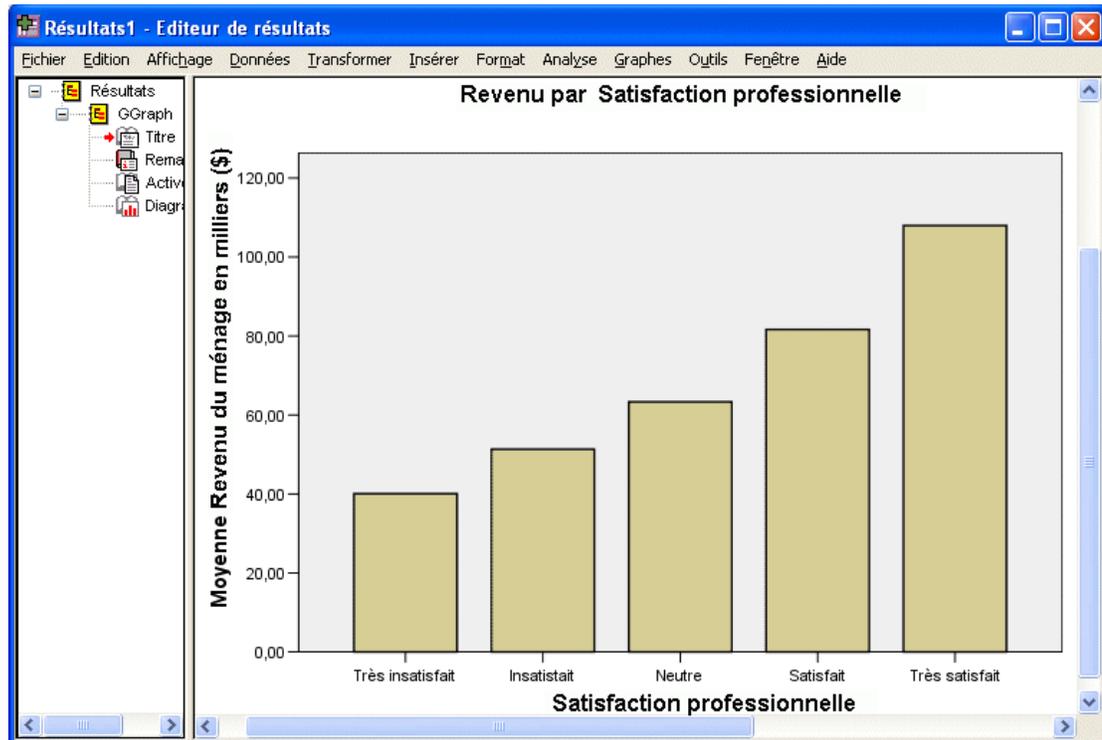
Le titre apparaît sur le canevas avec l'étiquette T1.

- ▶ Dans la fenêtre Propriété des éléments, sélectionnez Titre 1 dans Modifier les propriétés de liste.
- ▶ Dans la zone de texte Contenu, saisissez Revenu par satisfaction professionnelle. Il s'agit du texte que le titre affichera.
- ▶ Cliquez sur Appliquer pour enregistrer le texte. Bien que le texte ne soit pas affiché dans le Générateur de diagrammes, il apparaît lorsque vous générez le diagramme.

Création du diagramme

- Cliquez sur OK pour créer le diagramme en bâtons.

Figure 6-6
Diagramme en bâtons



Le diagramme en bâtons indique que les répondants les plus satisfaits de leur situation professionnelle ont tendance à avoir des revenus plus élevés.

Modification de diagrammes – Notions de base

Vous pouvez modifier les diagrammes de différentes façons. Pour l'exemple de diagramme en bâtons créé, vous allez effectuer les tâches suivantes :

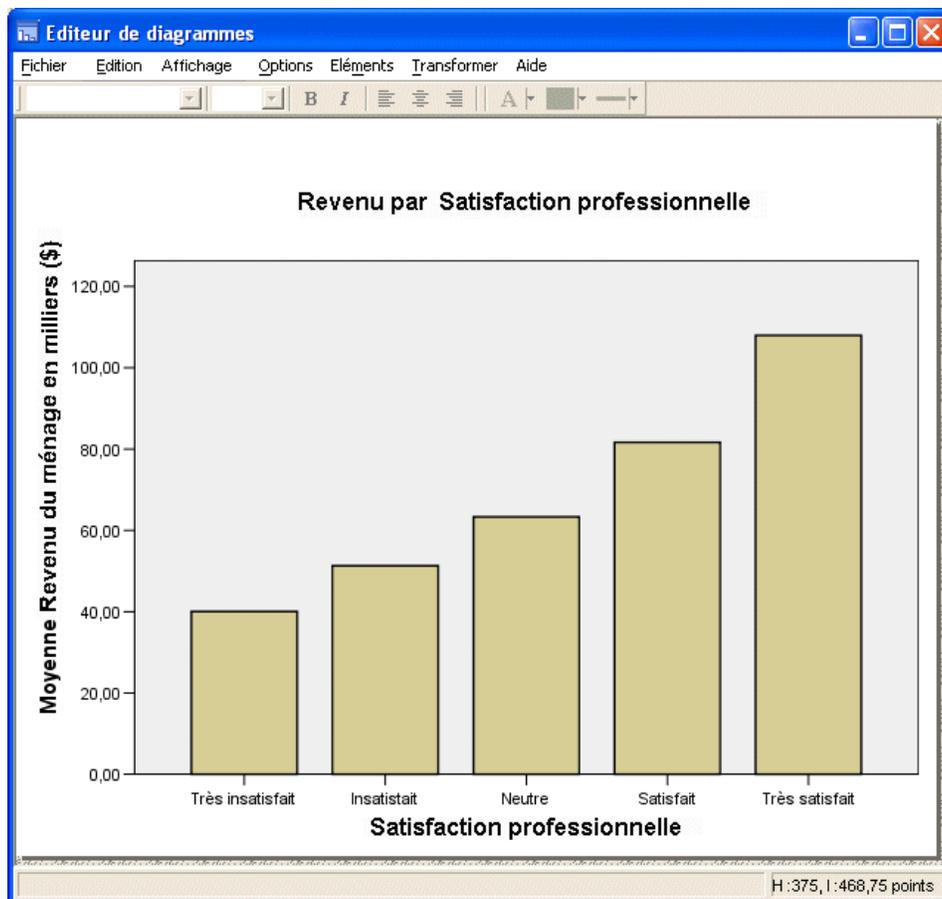
- Changement de couleur.
- Formatage des valeurs des étiquettes de graduation.
- Modification du texte.
- Affichage des étiquettes des valeurs de données.
- Utilisation de modèles de diagramme.

Pour modifier le diagramme, ouvrez-le dans l'éditeur de diagrammes.

- Double-cliquez sur le diagramme en bâtons pour l'ouvrir dans l'éditeur de diagrammes.

Figure 6-7

Diagramme en bâtons dans l'éditeur de diagrammes



Sélection d'éléments de diagramme

Pour modifier un élément de diagramme, sélectionnez-le.

- Cliquez sur l'un des bâtons. Les rectangles contenant les bâtons signifient qu'ils sont sélectionnés.

Il existe des règles générales pour sélectionner des éléments dans un diagramme simple :

- Si aucun des éléments graphiques n'est sélectionné, cliquez sur l'un d'eux afin de les sélectionner tous.
 - Si tous les éléments graphiques sont sélectionnés, cliquez sur l'un d'eux afin que la sélection porte sur lui seul. Pour sélectionner un autre élément graphique, vous pouvez cliquer dessus. Pour sélectionner plusieurs éléments graphiques, cliquez dessus tout en maintenant la touche Ctrl enfoncée.
- Pour désélectionner tous les éléments, appuyez sur la touche Echap.

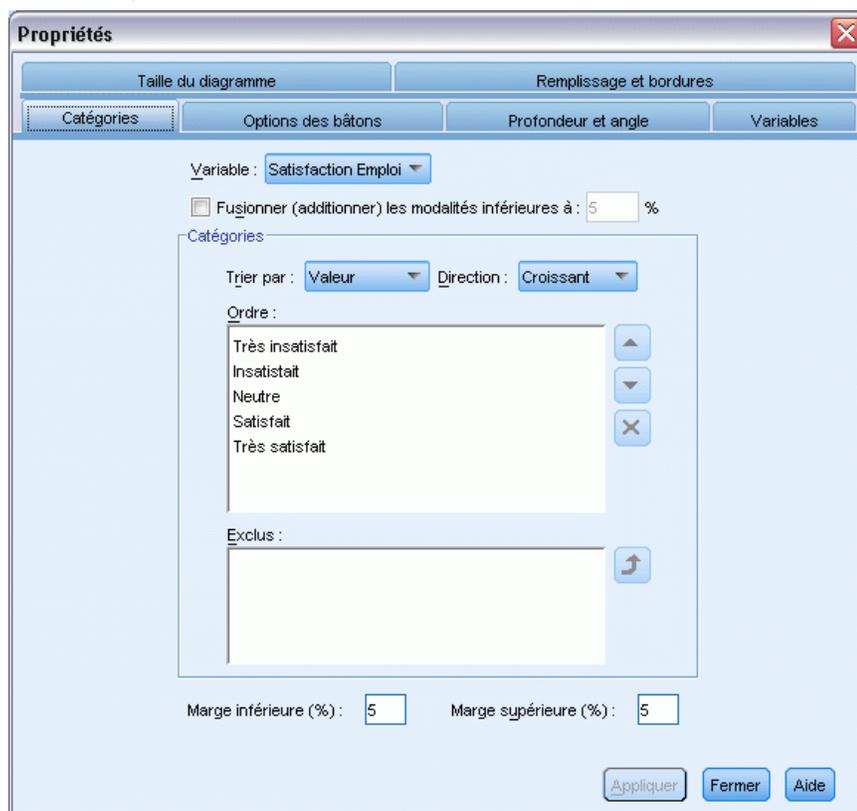
- Cliquez sur un bâton pour resélectionner tous les bâtons.

Utilisation de la fenêtre Propriétés

- A partir des menus de l'éditeur de diagrammes, sélectionnez :
Edition > Propriétés

Ce menu affiche la fenêtre Propriétés, qui contient les onglets s'appliquant aux bâtons sélectionnés. Ces onglets varient selon l'élément de diagramme que vous sélectionnez dans l'éditeur de diagrammes. Par exemple, si vous aviez sélectionné un cadre de texte au lieu de bâtons, plusieurs onglets apparaissent dans la fenêtre Propriétés. Ces onglets vous permettent d'effectuer la plupart des modifications de diagramme.

Figure 6-8
Fenêtre Propriétés

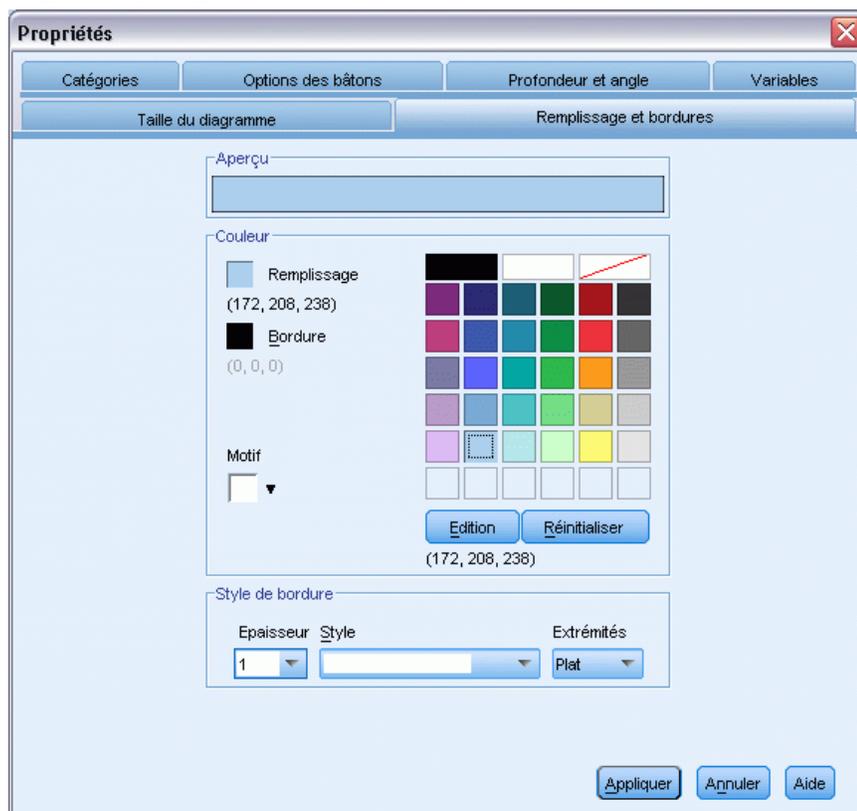


Changement de la couleur des bâtons

Vous allez d'abord changer la couleur des bâtons. En outre, vous spécifiez les attributs de couleur des éléments graphiques (à l'exception des courbes et des marques) dans l'onglet de remplissage et de bordures.

- ▶ Cliquez sur l'onglet Remplissage et bordures.
- ▶ Cliquez sur l'échantillon en regard de l'option de remplissage pour indiquer que vous souhaitez modifier la couleur de remplissage des bâtons. Les numéros figurant sous l'échantillon font référence aux composantes rouge, verte et bleue de la couleur actuelle.
- ▶ Cliquez sur la couleur bleu clair, au deuxième rang à gauche sur la deuxième ligne en partant du bas.

Figure 6-9
Onglet Remplissage et bordures



- ▶ Cliquez sur Appliquer.

Les bâtons du diagramme apparaissent désormais en bleu clair.

Figure 6-10
Modification du diagramme en bâtons bleus



Formatage des valeurs des étiquettes de graduation

Notez que les valeurs sur l'axe des y sont exprimées en milliers. Pour améliorer la présentation du diagramme et faciliter son interprétation, nous allons modifier le format numérique des étiquettes de graduation, puis le titre de l'axe en conséquence.

- ▶ Sélectionnez les étiquettes de graduation sur l'axe des y en cliquant dessus.
- ▶ Pour rouvrir la fenêtre Propriétés (si vous l'avez fermée précédemment), sélectionnez les options suivantes :
Edition > Propriétés

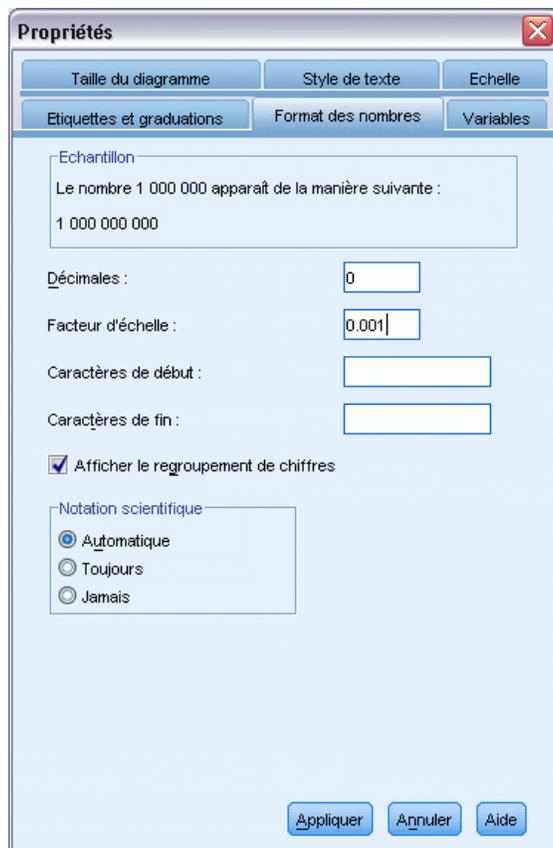
Remarque : A ce stade, nous supposons que la fenêtre Propriétés est ouverte. Si vous avez fermé la fenêtre Propriétés, suivez l'étape précédente pour la rouvrir. Il est également possible d'utiliser le raccourci clavier Ctrl+T afin de rouvrir cette fenêtre.

- ▶ Cliquez sur l'onglet Format numérique.
- ▶ Si vous ne souhaitez pas afficher les décimales sur les étiquettes de graduation, entrez 0 dans la zone de texte Décimales.
- ▶ Saisissez 0,001 dans la zone de texte Facteur d'échelle. Le facteur d'échelle correspond au diviseur de la valeur affichée, utilisé par l'éditeur de diagrammes. Etant donné que 0,001 représente une

fraction, employer ce diviseur *incrémenté* de 1 000 les valeurs des étiquettes de graduation. Par conséquent, elles ne sont plus exprimées en milliers et leur mise à l'échelle est annulée.

- Sélectionnez Afficher le regroupement de chiffres. La fonction de regroupement des chiffres utilise un caractère spécial (fourni par les paramètres régionaux de votre ordinateur) pour marquer la position des milliers dans la valeur concernée.

Figure 6-11
Onglet Format numérique

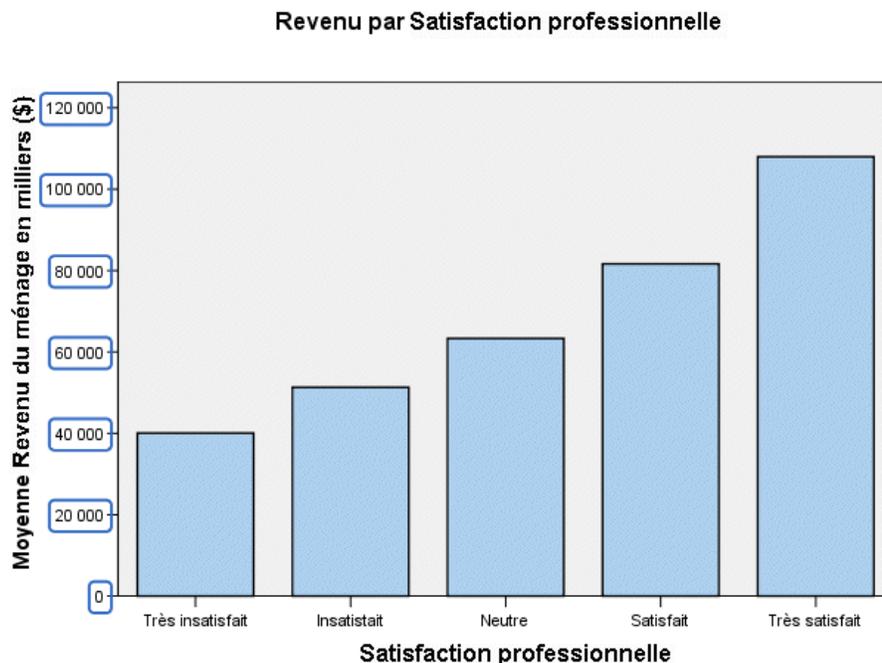


- Cliquez sur Appliquer.

Les étiquettes de graduation reflètent ce nouveau formatage numérique : Aucune décimale n'apparaît, les valeurs ne sont plus mises à l'échelle et les milliers sont signalés par un caractère spécial.

Figure 6-12

Diagramme en bâtons modifié contenant le nouveau format numérique



Modification de texte

Maintenant que vous avez modifié le format numérique des étiquettes de graduation, le titre de l'axe ne correspond plus. Vous devrez le modifier par la suite en tenant compte du nouveau format numérique.

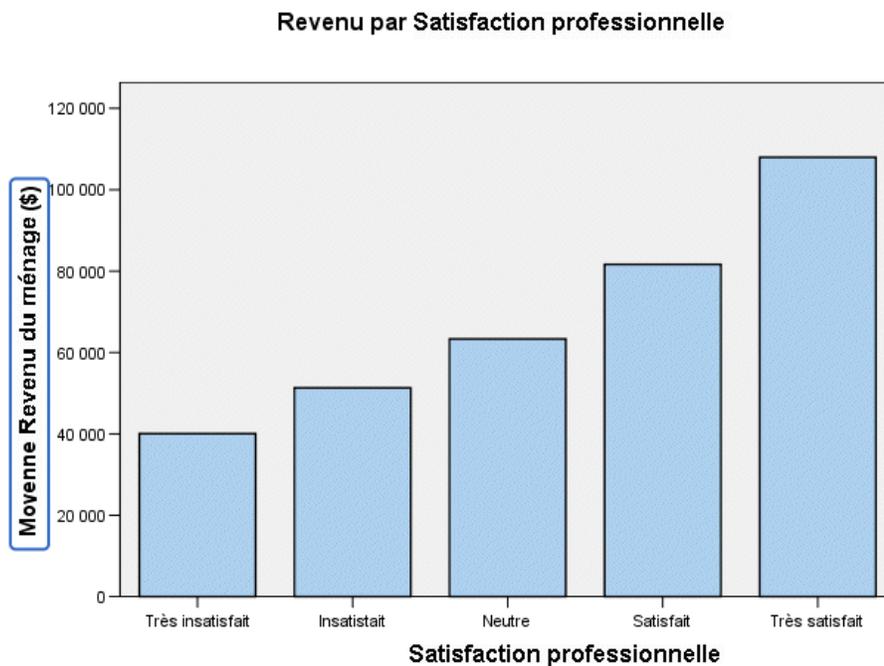
Remarque : Vous n'avez pas besoin d'ouvrir la fenêtre Propriétés pour modifier le texte. Vous pouvez modifier du texte directement à l'intérieur des diagrammes.

- ▶ Cliquez sur le titre de l'axe des y pour le sélectionner.
- ▶ Cliquez à nouveau dessus afin d'activer le mode d'édition. Une fois en mode d'édition, l'éditeur de diagrammes positionne horizontalement le texte après rotation. Il affiche également un curseur en forme de barre rouge qui clignote (non illustré dans cet exemple).

- ▶ Supprimez le texte suivant :
en milliers
- ▶ Appuyez sur Entrée pour désactiver le mode d'édition et mettre à jour le titre de l'axe. Ce titre décrit désormais avec précision le contenu des étiquettes de graduation.

Figure 6-13

Diagramme en bâtons contenant le titre modifié de l'axe des y

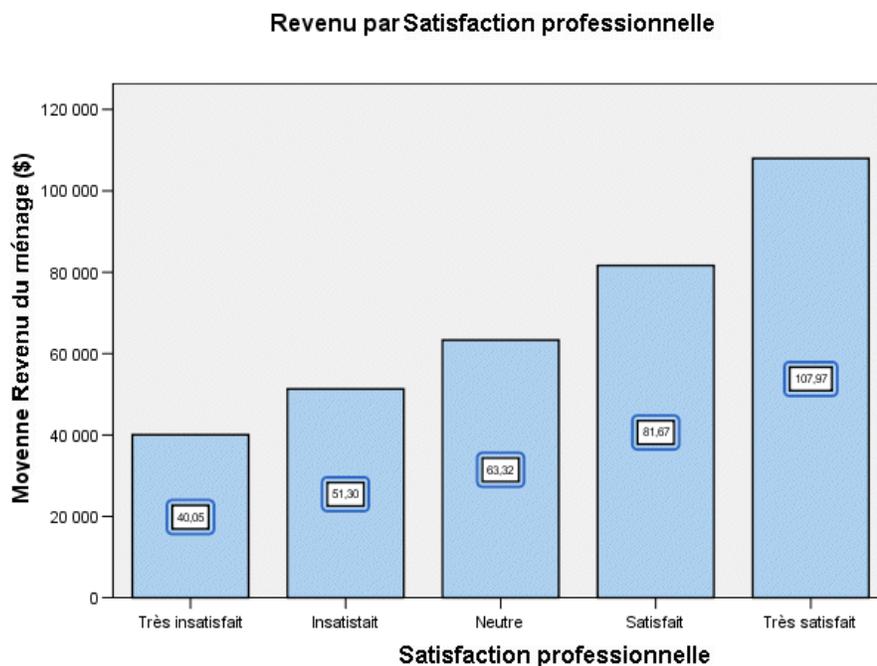


Affichage des étiquettes des valeurs de données

Une autre tâche courante consiste à afficher les valeurs précises associées aux éléments graphiques (sous forme de bâtons dans cet exemple). Ces valeurs apparaissent dans les étiquettes de données.

- ▶ A partir des menus de l'éditeur de diagrammes, sélectionnez :
Éléments > Afficher les étiquettes de données

Figure 6-14
Diagramme en bâtons contenant les étiquettes des valeurs de données



Chaque bâton du diagramme indique désormais le revenu moyen exact du ménage. Notez que les unités sont en milliers. Par conséquent, vous pouvez réutiliser l'onglet Format numérique pour modifier le facteur d'échelle.

Utilisation de modèles

Si vous modifiez régulièrement vos diagrammes, vous pouvez utiliser un modèle de diagramme pour réduire le temps nécessaire à leur création et à leur modification. Un modèle de diagramme enregistre les attributs d'un diagramme spécifique. Vous pouvez par la suite appliquer ce modèle lorsque vous créez ou modifiez un diagramme.

Nous allons enregistrer le diagramme actuel en tant que modèle, puis appliquer ce modèle lors de la création d'un nouveau diagramme.

- A partir des menus, sélectionnez :
Fichier > Enregistrer modèle de diagramme...

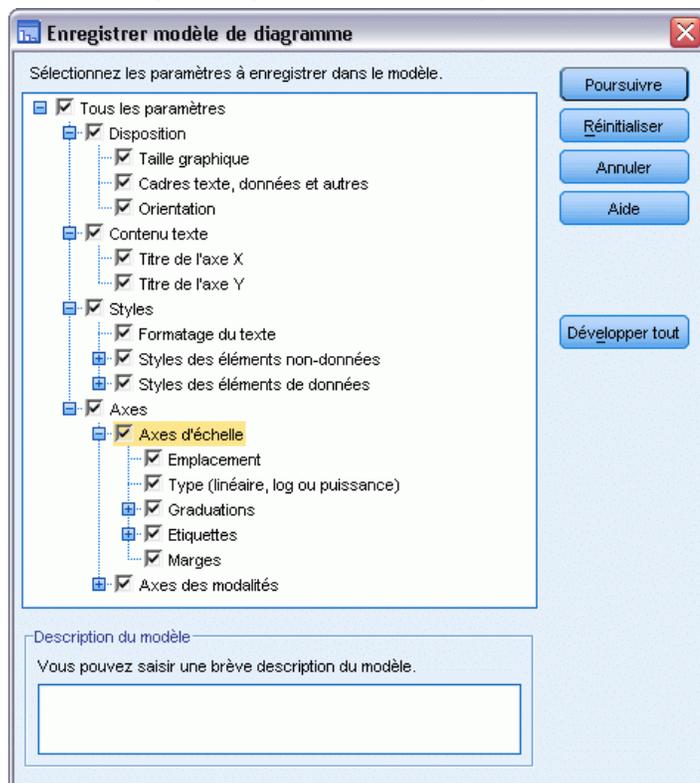
La boîte de dialogue Enregistrer modèle de diagramme vous permet de spécifier les attributs de diagramme à ajouter au modèle.

Si vous développez l'un des éléments de l'arbre, vous voyez apparaître les attributs susceptibles d'être enregistrés avec le diagramme. Par exemple, si vous développez la partie Axes d'échelle de l'arbre, vous voyez tous les attributs des étiquettes des valeurs de données que contiendra le modèle. Vous pouvez sélectionner n'importe quel attribut pour l'inclure dans ce modèle.

- Sélectionnez Tous les paramètres afin d'ajouter à ce modèle tous les attributs de diagramme disponibles.

Vous pouvez également saisir la description du modèle. Cette description apparaîtra lorsque vous appliquerez le modèle.

Figure 6-15
Boîte de dialogue Enregistrer modèle de diagramme



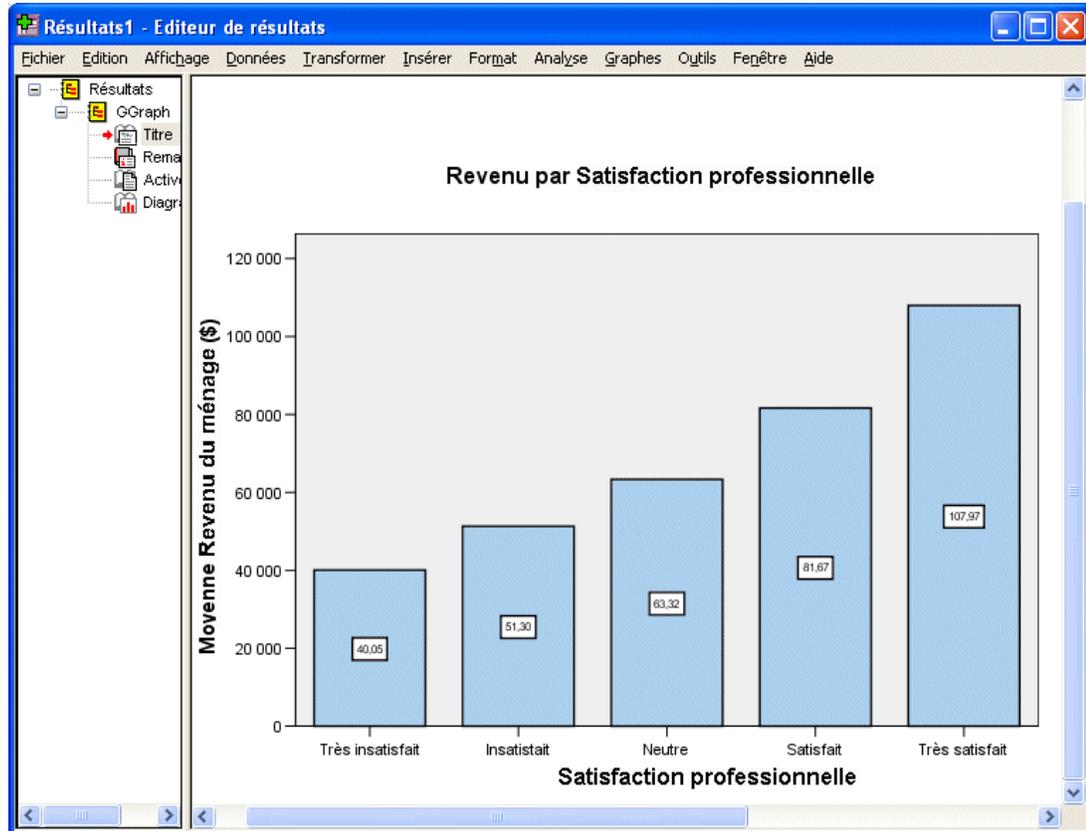
- Cliquez sur Poursuivre.
- Dans la boîte de dialogue Enregistrer le modèle, indiquez l'emplacement et le nom de fichier du modèle.
- Lorsque vous avez terminé, cliquez sur Enregistrer.

Vous pouvez appliquer le modèle quand vous créez un diagramme ou dans l'éditeur de diagrammes. Dans l'exemple suivant, nous allons l'appliquer durant la création d'un diagramme.

- Fermez l'éditeur de diagrammes. Le diagramme en bâtons mis à jour apparaît dans le Viewer.

Figure 6-16

Diagramme en bâtons mis à jour dans le Viewer



- A partir du menu du Viewer, sélectionnez :
Graphes > Générateur de diagrammes...

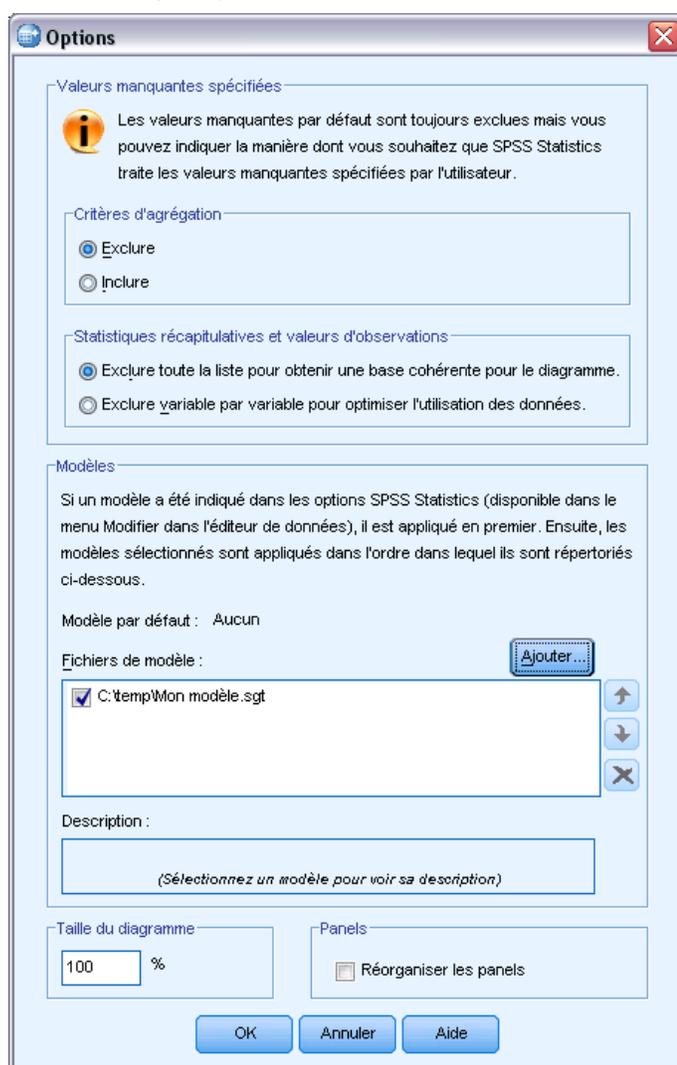
La boîte de dialogue Générateur de diagrammes « mémorise » les variables entrées lors de la création du premier diagramme. Toutefois, dans le cas présent, vous allez créer un diagramme légèrement différent pour voir comment l'application d'un modèle formate un diagramme.

- Supprimez *Satisfaction professionnelle* de l'axe *x* en le glissant de la zone de déplacement vers la liste Variables. Vous pouvez également cliquer sur la zone de déplacement et appuyer sur Supprimer.
- Cliquez avec le bouton droit de la souris sur *Niveau d'éducation* dans la liste Variables et choisissez Ordinal.
- Faites glisser *Niveau d'éducation* de la liste Variables vers la zone de déplacement de l'axe *x*.
Puisque le titre n'est plus pertinent, nous allons le supprimer.
- Dans l'onglet Titres/Notes de bas de page, désélectionnez Titre 1.

Nous allons maintenant indiquer le modèle à appliquer au nouveau diagramme.

- ▶ Cliquez sur Options.
- ▶ Dans le groupe Modèles de la boîte de dialogue Options, cliquez sur Ajouter.
- ▶ Dans la boîte de dialogue Trouver les fichiers du modèle, repérez le fichier du modèle préalablement enregistré à l'aide de la boîte de dialogue Enregistrer modèle de diagramme.
- ▶ Sélectionnez ce fichier et cliquez sur Ouvrir.

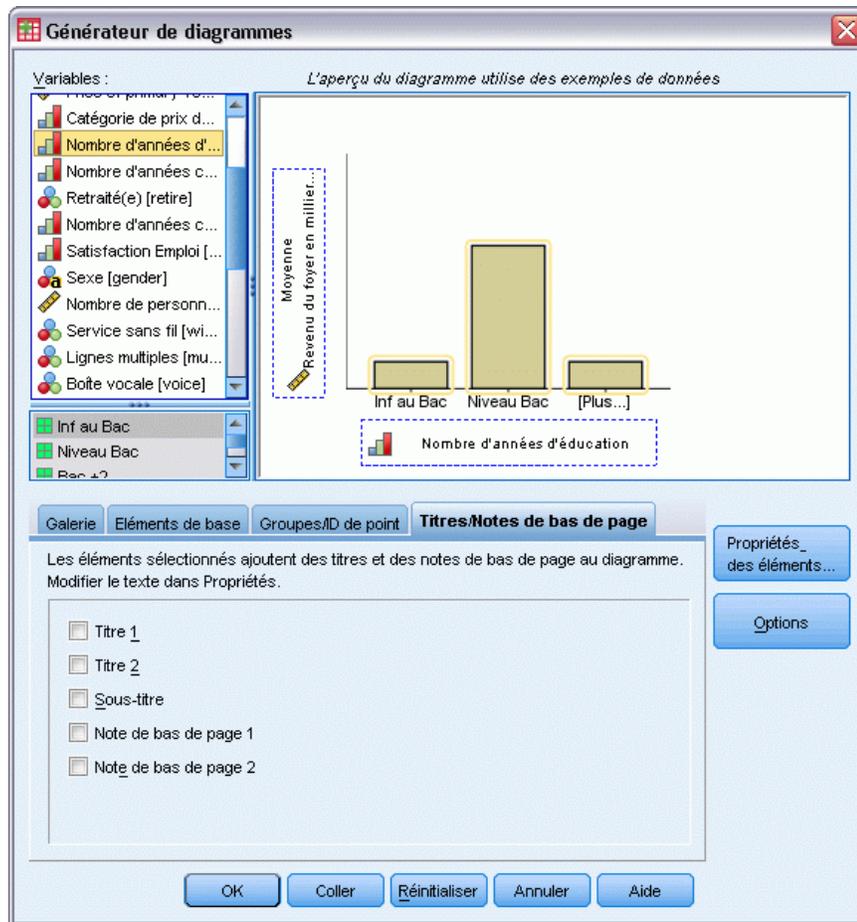
Figure 6-17
Boîte de dialogue Options avec le modèle



La boîte de dialogue Options indique le chemin d'accès du modèle sélectionné.

- ▶ Cliquez sur OK pour fermer la boîte de dialogue Options.

Figure 6-18
Générateur de diagrammes avec des zones de déplacement complétées

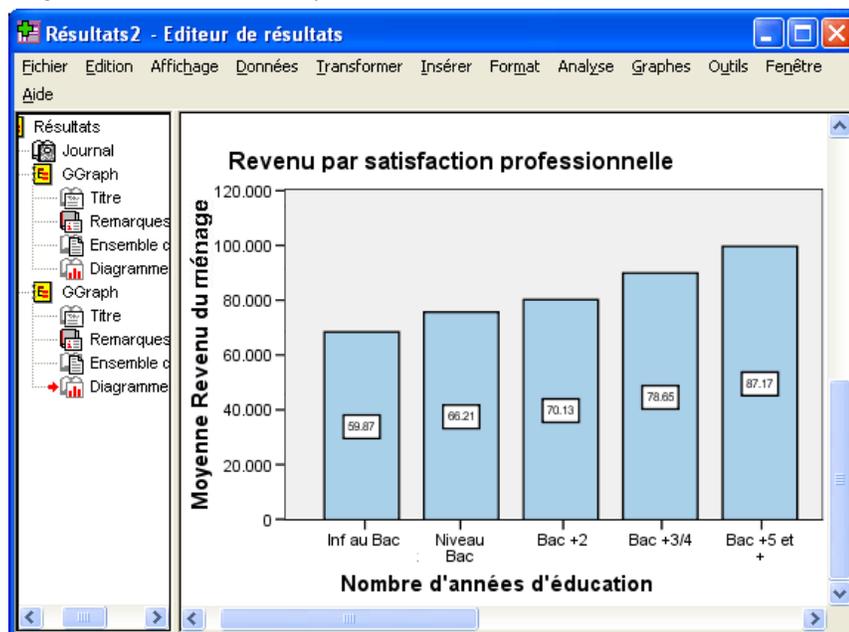


- Cliquez sur OK dans la boîte de dialogue Générateur de diagrammes pour créer le diagramme et appliquer le modèle.

Le formatage du nouveau diagramme correspond à celui du diagramme que vous avez créé et modifié précédemment. Même si les variables sur l'axe des x sont différentes, les diagrammes, eux, sont semblables. Notez que le titre du diagramme précédent a été conservé dans le modèle, même si vous avez supprimé le titre dans le Générateur de diagrammes.

Si vous souhaitez appliquer des modèles après avoir créé un diagramme, vous pouvez le faire dans Chart Editor (choisissez l'option Appliquer modèle de diagramme du menu Fichier).

Figure 6-19
Diagramme en bâtons mis à jour dans le Viewer



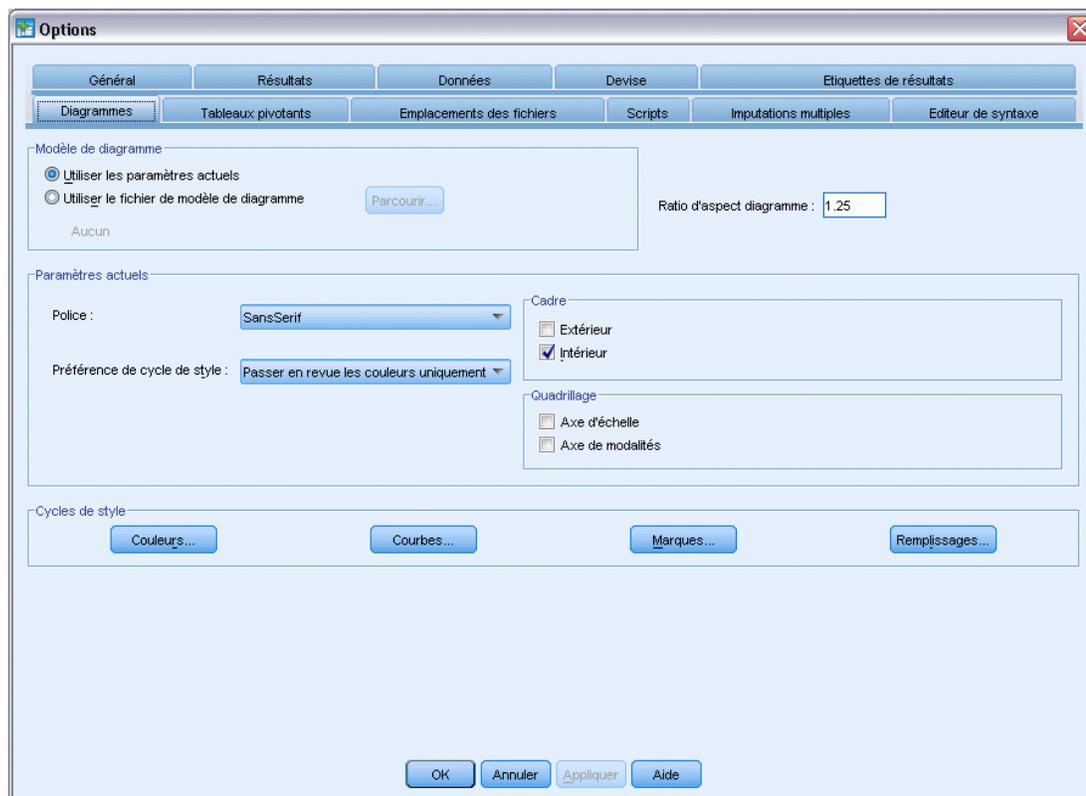
Définition des options de diagramme

En plus des modèles permettant de formater des diagrammes, vous pouvez utiliser les différentes options pour contrôler plusieurs aspects du mode de création des diagrammes.

- Dans les menus de l'éditeur de données ou du Viewer, sélectionnez :
Edition > Options...

La boîte de dialogue Options contient de nombreux paramètres de configuration. Cliquez sur l'onglet Diagrammes pour visualiser les options disponibles.

Figure 6-20
Onglet Diagrammes de la boîte de dialogue Options



Les options déterminent le mode de création du diagramme. Pour chaque nouveau diagramme, vous pouvez spécifier :

- Si les paramètres actuels ou un modèle doivent être utilisés.
- Le rapport hauteur/largeur (proportions).
- Si vous ne vous servez d'aucun modèle, les paramètres à utiliser en vue du formatage.
- Les cycles de style des éléments graphiques.

Les cycles de style vous permettent d'indiquer le style des éléments graphiques des nouveaux diagrammes. Dans cet exemple, nous allons consulter les détails du cycle de style des couleurs.

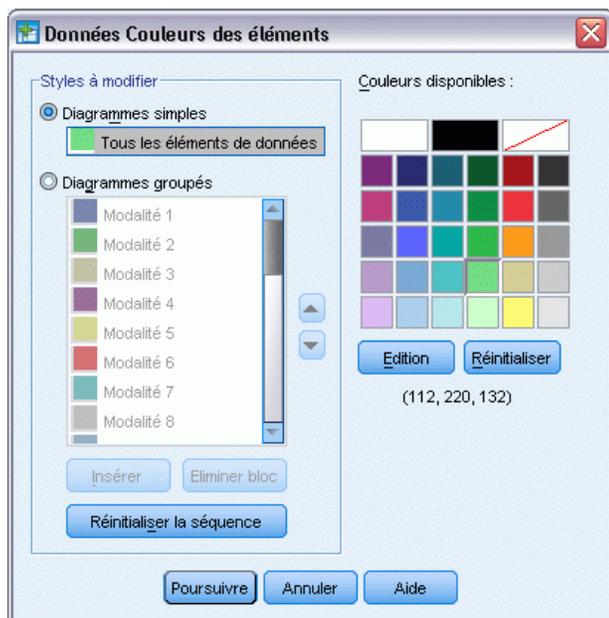
- ▶ Cliquez sur Couleurs pour ouvrir la boîte de dialogue Données Couleurs des éléments.

Pour un diagramme simple, l'éditeur de diagrammes utilise le style que vous spécifiez. Pour un diagramme regroupé, l'éditeur de diagrammes utilise un ensemble de styles qu'il applique à chaque groupe (modalité) du diagramme.

- ▶ Sélectionnez Diagrammes simples.

- Sélectionnez la couleur vert clair, qui est la troisième en partant de la droite dans la deuxième ligne à partir du bas.

Figure 6-21
Boîte de dialogue Données Couleurs des éléments



- Cliquez sur Poursuivre.
- Dans la boîte de dialogue Options, cliquez sur OK pour enregistrer les modifications apportées au cycle de style des couleurs.

Les éléments graphiques dans tout nouveau diagramme simple apparaissent maintenant en vert clair.

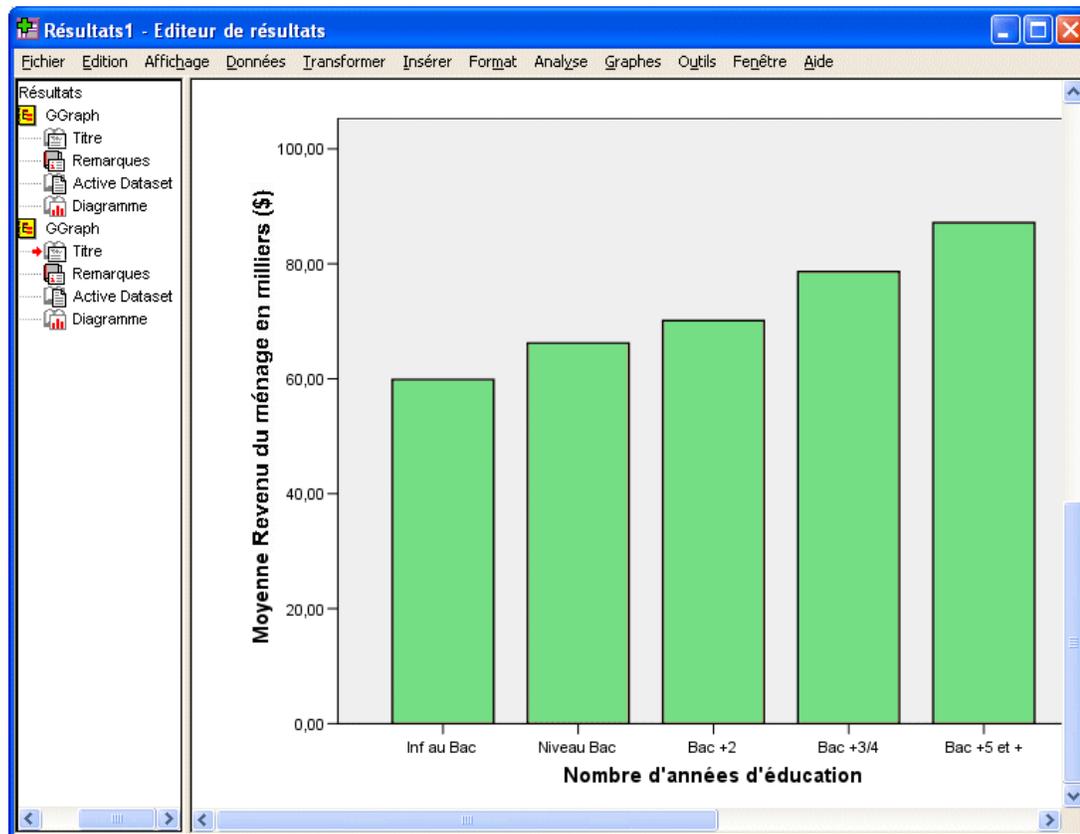
- Dans les menus de l'éditeur de données ou du Viewer, sélectionnez :
Graphes > Générateur de diagrammes...

Le Générateur de diagrammes affiche le dernier diagramme créé. Souvenez-vous que ce diagramme était associé à un modèle. Nous ne souhaitons plus utiliser ce modèle.

- Cliquez sur Options.
- Désélectionnez (décochez) le modèle que vous avez ajouté précédemment. Notez que vous pouvez également cliquer sur le X rouge pour supprimer le modèle. Désélectionner plutôt que supprimer permet de garder le modèle disponible pour une utilisation ultérieure.
- Cliquez sur OK pour créer le diagramme.

Les bâtons du nouveau diagramme sont en vert clair. Ce diagramme est différent du dernier pour plusieurs raisons. Il n'y a pas de titre ; les étiquettes des axes sont exprimées en milliers ; il n'y a pas d'étiquette de données. Des différences sont survenues parce que le modèle ne s'appliquait pas au diagramme.

Figure 6-22
Diagramme en bâtons mis à jour dans le Viewer

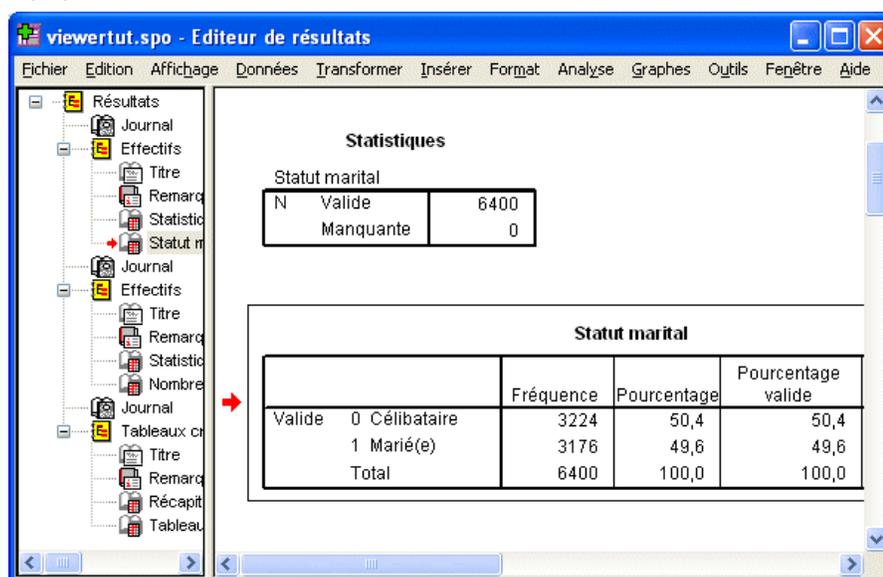


Utilisation du résultat

Les résultats des procédures statistiques apparaissent dans le Viewer. En fonction des choix effectués lors de l'exécution de la procédure, ces résultats peuvent prendre la forme d'un texte, ou de tableaux, de diagrammes ou de graphiques statistiques. Cette section utilise les fichiers *viewertut.spv* et *demo.sav*. Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.

Utilisation du Viewer

Figure 7-1
Viewer



La fenêtre du Viewer est divisée en deux panneaux. Le **panneau de légende** contient la légende de toutes les informations stockées dans le Viewer. Le **panneau de contenu** comporte les tableaux statistiques, les graphiques et les textes.

A l'aide des barres de défilement horizontale et verticale, parcourez le contenu de la fenêtre. Pour faciliter la navigation, cliquez sur un item dans le panneau de légende pour l'afficher dans le panneau de contenu.

Si une partie seulement du tableau apparaît dans le Viewer ou si l'affichage de la légende est trop serré, vous pouvez redimensionner la fenêtre.

- Cliquez sur la bordure droite du panneau de légende et faites-la glisser pour modifier sa largeur.

Dans le panneau de légende, lorsqu'une icône en regard d'un élément représente un livre ouvert, cela indique qu'il est actuellement visible dans le Viewer, même s'il n'est pas visible dans le panneau de contenu.

- Pour masquer un tableau ou un diagramme, double-cliquez sur l'icône du livre dans le panneau de légende.

L'icône représente maintenant un livre fermé, ce qui signifie que les informations associées sont masquées.

- Pour que le résultat apparaisse à nouveau, double-cliquez sur l'icône du livre fermé.

Vous pouvez également masquer tous les résultats d'une procédure statistique donnée ou l'intégralité des résultats dans le Viewer.

- Cliquez sur la case comportant le signe moins (-), à gauche de la procédure dont vous souhaitez masquer les résultats, ou cochez la case située en regard de l'élément figurant tout en haut du panneau de légende pour masquer la totalité des résultats.

L'arborescence de la légende est réduite, indiquant visuellement que les résultats sont masqués.

Vous pouvez également modifier l'ordre de l'affichage des résultats.

- Dans le panneau de légende, cliquez sur les items à déplacer.
- Faites glisser les items sélectionnés vers un nouvel emplacement de la légende et relâchez le bouton de la souris.

Figure 7-2
Résultats réorganisés dans le Viewer

Récapitulatif du traitement des observations

	Observations			
	Valide		Manquante	
	N	Pourcentage	N	Pourcentage
Possesseur d'un agenda électronique * Sexe * Internet	6145	96,0%	255	4,0%

Tableau croisé Possesseur d'un agenda électronique

Internet	

Vous pouvez également déplacer les éléments de résultat en cliquant dessus et en les faisant glisser dans le panneau de contenu.

Utilisation de l'éditeur de tableau pivotant

Les résultats de la plupart des procédures statistiques sont affichés dans **des tableaux pivotants**.

Accès aux définitions de résultats

De nombreux termes statistiques sont affichés dans le résultat. Vous pouvez directement accéder aux définitions de ces termes dans le Viewer.

- ▶ Activez le tableau croisé (double-cliquez dessus) *Possède un agenda électronique * Sexe * Internet*.
- ▶ Cliquez avec le bouton droit de la souris sur *Effectif théorique* et sélectionnez *Qu'est-ce que c'est ?* dans le menu contextuel.

La définition apparaît dans une fenêtre contextuelle.

Figure 7-3
Définition contextuelle

The screenshot shows the SPSS Viewer window titled 'viewertut.spv [Document2] - Viewer'. The main content is a pivot table titled 'Possède un agenda électronique * Sexe * Internet'. The table is structured as follows:

				Sexe	
				f Femme	m Homme
Internet					
0 Aucun	Possesseur d'un agenda électronique	0 Non	Effectif	1897	1956
			Effectif théorique	1903,4	1956,0
			% dans Possesseur d'un agenda électronique	49,3%	50,0%
			Effectif théorique	320,6	329,0
			% dans Possesseur d'un agenda électronique	50,3%	49,7%
	Total		Effectif	2224	2285
			Effectif théorique	2224,0	2285,0
			% dans Possesseur d'un agenda électronique	49,3%	50,7%
1 Dial-up	Possesseur d'un agenda électronique	0 Non	Effectif	513	512
			Effectif théorique	512,5	512,0
			% dans Possesseur d'un agenda électronique	50,0%	50,0%
		1 Oui	Effectif	305	305
			Effectif théorique	305,5	305,0

A context menu is open over the 'Effectif théorique' cell in the '0 Non' row. The menu text reads: 'Nombre d'observations d'un ensemble de valeurs de facteurs prévu par le modèle sélectionné. Ce nombre est toujours positif.'

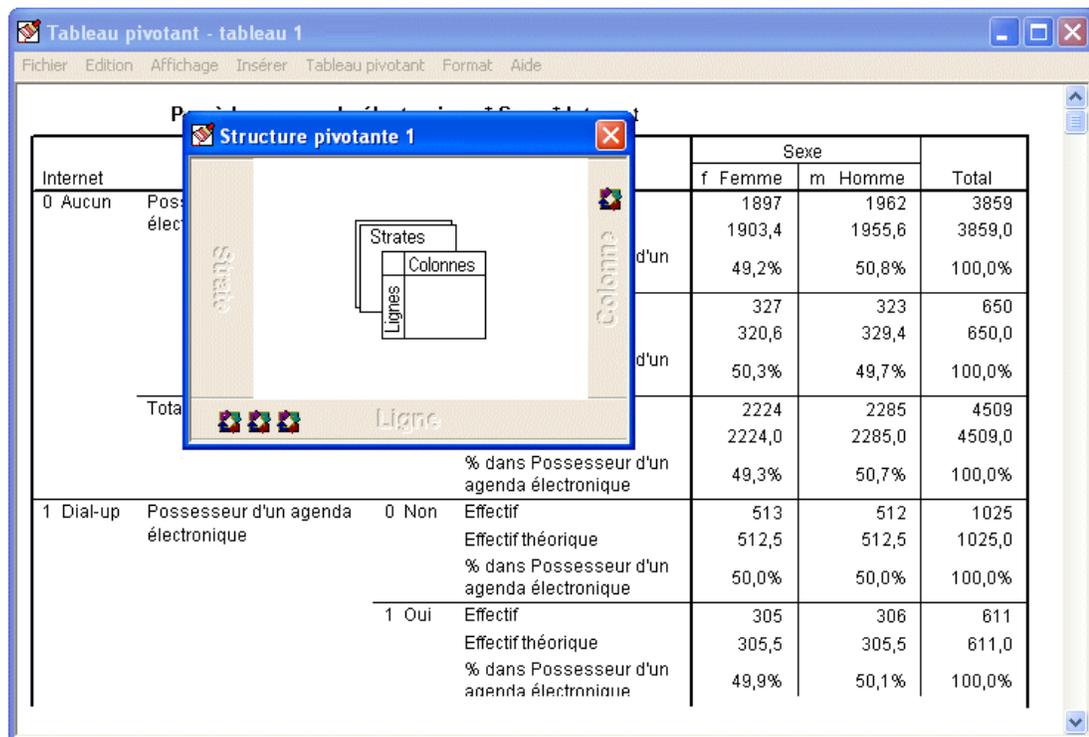
Tableaux pivotants

Les tableaux produits par défaut risquent de ne pas présenter les informations aussi clairement que vous le voudriez. Avec les tableaux pivotants, vous pouvez transposer des lignes et des colonnes (« retourner » le tableau), ajuster l'ordre des données d'un tableau et modifier le tableau de nombreuses autres façons. Vous pouvez, par exemple, changer un tableau court et large en un tableau long et étroit, en transposant les lignes et les colonnes. La modification de la mise en page d'un tableau n'affecte pas les résultats. C'est, au contraire, une façon d'afficher vos informations d'une manière différente, plus attrayante.

- ▶ Si il n'est pas déjà activé, double-cliquez sur le tableau croisé *Possède un agenda électronique * Sexe * Internet* pour l'activer.
- ▶ Si la fenêtre Structures pivotantes n'est pas visible, à partir des menus, sélectionnez :
Tableau pivotant > Structure pivotante

Les structures pivotantes vous permet de déplacer les données entre les colonnes, les lignes et les strates.

Figure 7-4
Structure pivotante



- Faites glisser l'élément *Statistiques* de la dimension de ligne à la dimension de colonne, sous *Sexe*. Le tableau est immédiatement reconfiguré pour refléter vos modifications.

Figure 7-5

Déplacement des lignes vers les colonnes

The screenshot shows a Tableau pivot table window titled "Tableau pivotant - tableau 1". The table is titled "Possède un agenda électronique * Sexe * Internet". The table has three columns: "Effectif", "Effectif théorique", and "% dans Possède agenda élec". The rows are grouped by "Internet" (0 Aucun, 1 Dial-up) and "Sexe" (0 Non, 1 Oui). The "Statistiques" dimension is placed on the columns shelf.

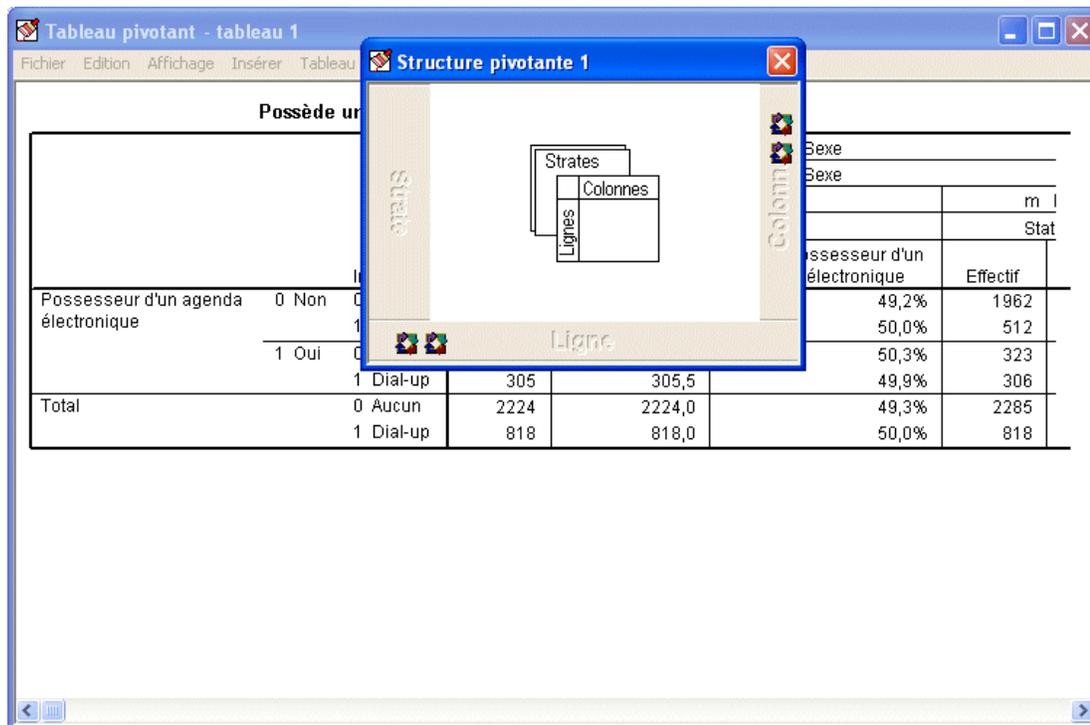
			Sexe		
			f Femme		
			Statistiques		
Internet			Effectif	Effectif théorique	% dans Possède agenda élec
0 Aucun	Possesseur d'un agenda électronique	0 Non	1897	1903,4	
		1 Oui	327	320,6	
	Total		2224	2224,0	
1 Dial-up	Possesseur d'un agenda électronique	0 Non	513	512,5	
		1 Oui	305	305,5	
	Total		818	818,0	

The structure window "Structure pivotante 1" shows the dimensions: Strates, Lignes, and Colonnes. The "Lignes" dimension is currently on the left shelf, and "Statistiques" is on the right shelf.

L'ordre des éléments dans la structure pivotante reflète l'ordre d'affichage des éléments dans le tableau.

- Faites glisser l'élément *Possède un agenda électronique* avant l'élément *Internet* dans la dimension de ligne pour inverser l'ordre de ces deux lignes.

Figure 7-6
Permuter les lignes



Création et affichage de strates

Les strates s'avèrent utiles pour les grands tableaux ayant des modalités d'informations emboîtées. En créant des strates, vous simplifiez le tableau qui devient plus facile à lire.

- Faites glisser l'élément *Sexe* de la dimension de colonne à la dimension de strate.

Figure 7-7

Icône de rotation Sexe dans la dimension Strate

Sexe **Femme** ▼

	Internet		Effectif	Effectif théorique	% compris dans Possesseur d'un agenda électronique
Possesseur d'un agenda électronique	Non	Non	1164	1128,4	79,3%
Total					

Structure pivotante

Statistiques

COLONNE

STRATE

LIGNE

Pour afficher une autre strate, sélectionnez une modalité dans la liste déroulante du tableau.

Figure 7-8

Choix d'une strate

Tableau pivotant - tableau 1

Fichier Edition Affichage Insérer Tableau pivotant Format Aide

Possède un agenda électronique * Sexe * Internet

Sexe				Statistiques		
Sexe Femme				Effectif	Effectif théorique	% dans Possesseur agenda électronique
Internet	Total					
0 Aucun	Possesseur d'un agenda électronique	0 Non	1 Oui	1897	1903,4	4
	Total			327	320,6	5
				2224	2224,0	4
1 Dial-up	Possesseur d'un agenda électronique	0 Non	1 Oui	513	512,5	5
	Total			305	305,5	4
				818	818,0	5

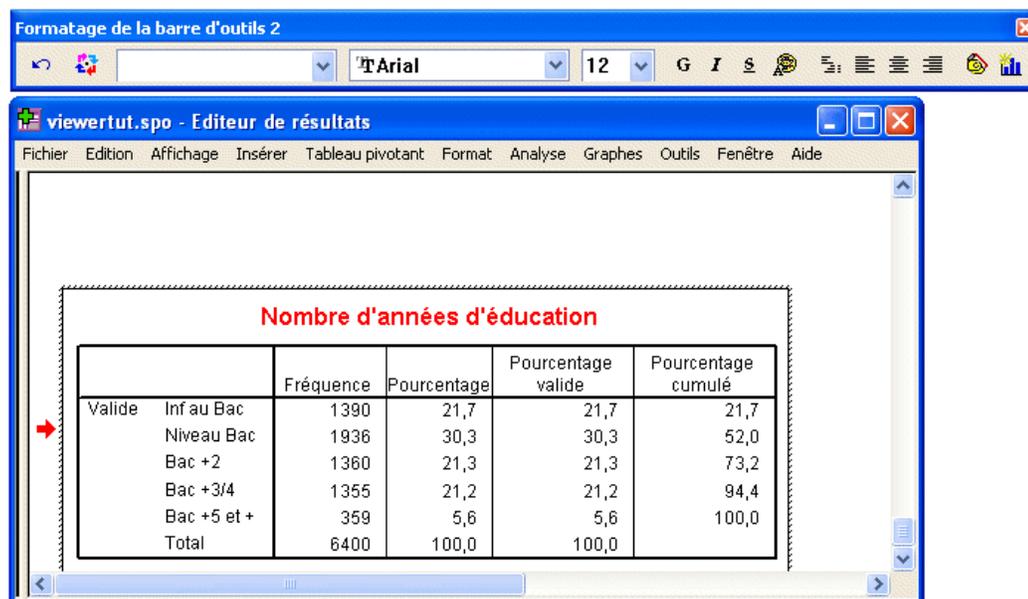
Modification de tableaux

A moins que vous ayez pris le temps de créer un modèle de tableaux personnalisé, les tableaux pivotants sont créés avec un format standard. Vous pouvez modifier le format d'un texte à l'intérieur d'un tableau. Les formats modifiables sont entre autres le nom, la taille, le style (italique ou gras) et la couleur de la police.

- ▶ Double-cliquez sur le tableau *Niveau d'éducation*.
- ▶ Si la barre d'outils de formatage n'est pas visible, à partir des menus, sélectionnez :
Affichage > Barre d'outils
- ▶ Cliquez sur le texte du titre, *Niveau d'éducation*.
- ▶ Dans la liste déroulante des tailles de police de la barre d'outils, sélectionnez 12.
- ▶ Pour modifier la couleur du texte du titre, cliquez sur l'outil Couleur du texte et sélectionnez une couleur.

Figure 7-9

Texte du titre reformaté dans le tableau pivotant



Vous pouvez également modifier le contenu des tableaux et des étiquettes. Vous pouvez, par exemple, modifier le titre de ce tableau.

- ▶ Double-cliquez sur le titre.
- ▶ Saisissez Niveau d'enseignement dans la nouvelle étiquette.

Remarque : Si vous modifiez les valeurs d'un tableau, les totaux et les autres statistiques ne sont pas recalculés.

Masquage de lignes et de colonnes

Certaines données affichées dans un tableau ne sont pas forcément utiles ou rendent le tableau trop complexe. Vous pouvez masquer des lignes et des colonnes entières sans perdre de données.

- ▶ S'il n'est pas déjà activé, double-cliquez sur le tableau *Niveau d'éducation* pour l'activer.
- ▶ Cliquez sur la colonne *Pourcentage valide* pour la sélectionner.
- ▶ Dans le menu Edition ou le menu contextuel, choisissez :
Sélectionnez > Cellules de données et d'étiquettes
- ▶ Dans le menu Affichage, choisissez Masquer ou dans le menu contextuel, choisissez Masquer la modalité.

La colonne est masquée, mais non supprimée.

Figure 7-10

Colonne *Pourcentage valide* masquée dans le tableau

Niveau d'éducation		Fréquence	Pourcentage	Pourcentage cumulé
Valide	Inf au Bac	1390	21,7	21,7
	Niveau Bac	1936	30,3	52,0
	Bac +2	1360	21,3	73,2
	Bac +3/4	1355	21,2	94,4
	Bac +5 et +	359	5,6	100,0
	Total	6400	100,0	

Pour afficher de nouveau la colonne :

- ▶ A partir des menus, sélectionnez :
Affichage > Montrer Tout

A l'instar des colonnes, les lignes peuvent être masquées et affichées.

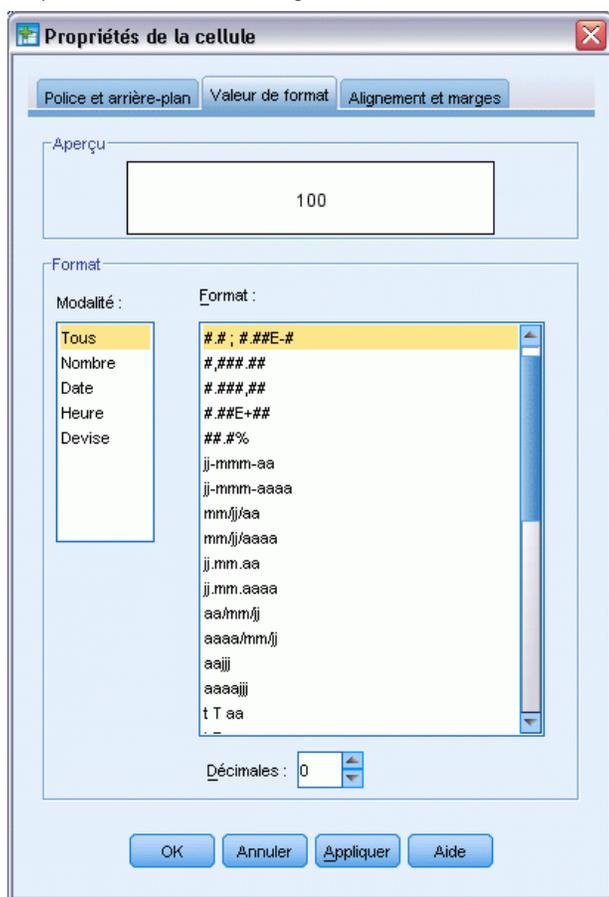
Modification des formats d'affichage des données

Vous pouvez modifier le format d'affichage des données dans les tableaux pivotants en toute facilité.

- ▶ S'il n'est pas déjà activé, double-cliquez sur le tableau *Niveau d'éducation* pour l'activer.

- ▶ Cliquez sur la colonne *Pourcentage* pour la sélectionner.
- ▶ Dans le menu Edition ou le menu contextuel, choisissez :
Sélectionnez > Cellules de Données
- ▶ Dans le menu Format ou le menu contextuel, choisissez Propriétés de la cellule.
- ▶ Cliquez sur l'onglet Valeur de format.
- ▶ Saisissez 0 dans le champ Décimales pour masquer tous les séparateurs décimaux dans cette colonne.

Figure 7-11
Propriétés des Cellules, onglet Valeur de format



Vous pouvez également modifier le type de données et le formatage dans cette boîte de dialogue.

- ▶ Sélectionnez le type souhaité dans la liste Modalité et le format de ce type dans la liste Format.

- Cliquez sur OK ou Appliquer pour appliquer les modifications que vous avez apportées.

Figure 7-12

Décimales masquées dans la colonne *Pour cent*

Niveau d'éducation		Fréquence	Pourcentage	Pourcentage cumulé
Valide	Inf au Bac	1390	22	21,7
	Niveau Bac	1936	30	52,0
	Bac +2	1360	21	73,2
	Bac +3/4	1355	21	94,4
	Bac +5 et +	359	6	100,0
	Total	6400	100	

Les décimales sont à présent masquées dans la colonne *Pourcentage*.

Modèles de tableaux

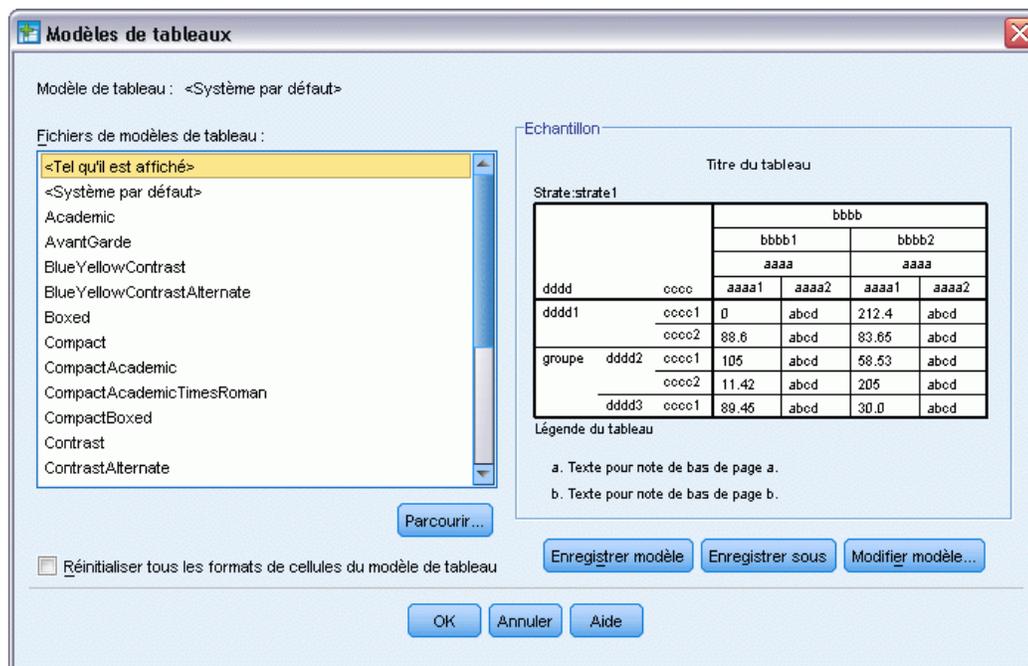
Le format de vos tableaux est primordial pour que les résultats soient clairs, précis et significatifs. Si votre tableau est complexe, les informations qu'il contient risquent d'être difficiles à comprendre.

Utilisation de formats prédéfinis

- Double-cliquez sur le tableau *Situation familiale*.
- A partir des menus, sélectionnez :
Format > Modèles de tableaux

La boîte de dialogue Modèles de tableaux répertorie différents styles prédéfinis. Sélectionnez un style dans la liste pour afficher son aperçu dans la fenêtre Exemple à droite.

Figure 7-13
Boîte de dialogue Modèles de tableaux



Vous pouvez utiliser un style tel qu'il est ou le modifier pour l'adapter à vos besoins.

- Pour utiliser un style existant, sélectionnez-le et cliquez sur OK.

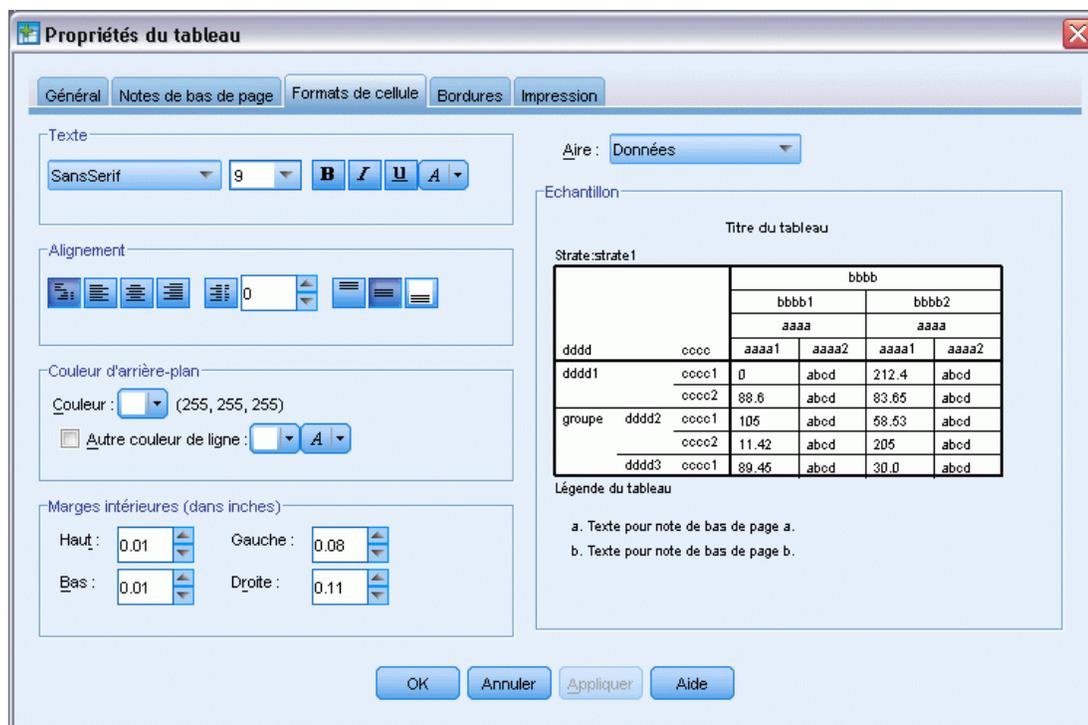
Personnalisation des styles Modèles de tableau

Vous pouvez personnaliser un format pour l'adapter à vos besoins. La quasi-totalité des aspects d'un tableau peut être personnalisée, de la couleur d'arrière-plan aux styles de bordure.

- Double-cliquez sur le tableau *Situation familiale*.
- A partir des menus, sélectionnez :
Format > Modèles de tableaux
- Sélectionnez le style correspondant le mieux au format voulu et cliquez sur Modifier modèle.

- Cliquez sur l'onglet Formats de cellule pour afficher les options de formatage.

Figure 7-14
Boîte de dialogue Propriétés du tableau



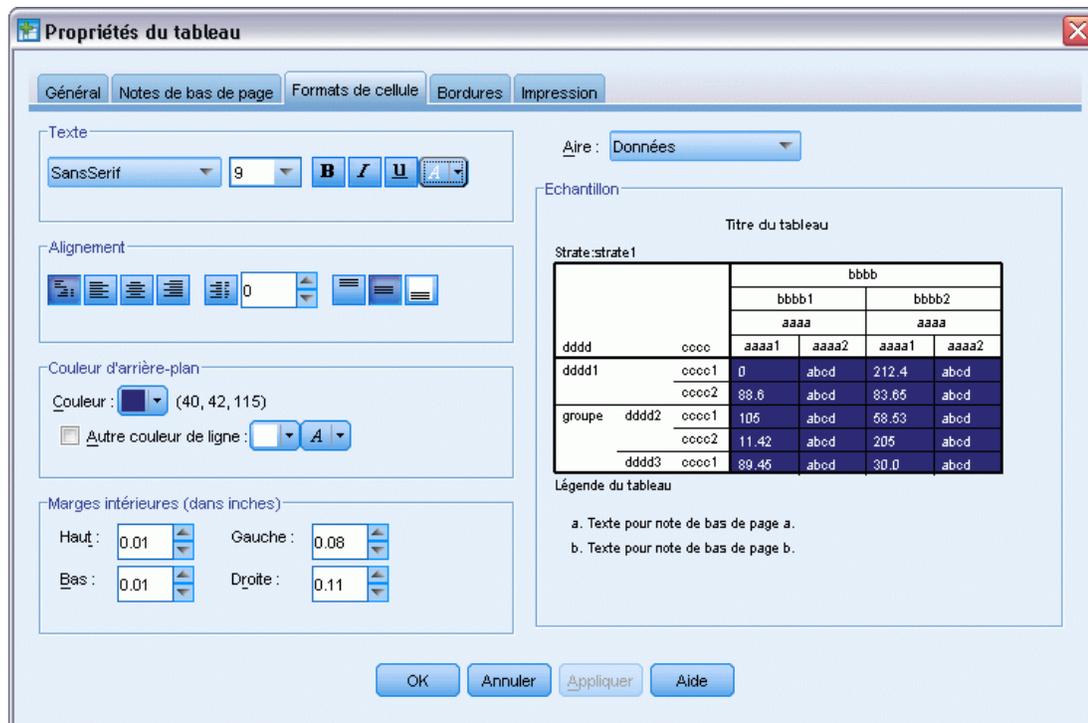
Les options de formatage comprennent le nom, la taille, le style et la couleur de police. Elles incluent également l'alignement, la couleur du texte et de l'arrière-plan, et la taille des marges.

La fenêtre Echantillon à droite fournit un aperçu de la modification de l'aspect du tableau. Chaque zone du tableau peut avoir des styles de format différents. Vous pouvez, par exemple, ne pas appliquer le même style au titre et aux données. Pour sélectionner une zone de tableau à modifier, sélectionnez le nom de la zone dans la liste déroulante Aires ou cliquez sur la zone dans la fenêtre Exemple.

- Sélectionnez Données dans la liste déroulante Aires.
- Sélectionnez une nouvelle couleur dans la palette déroulante Arrière-plan.
- Ensuite, sélectionnez une nouvelle couleur de texte.

La fenêtre Exemple illustre le nouveau style.

Figure 7-15
Modification des formats de cellule de tableau



- ▶ Cliquez sur OK pour revenir à la boîte de dialogue Modèles de tableaux.
- Vous pouvez enregistrer le nouveau style pour l'utiliser ultérieurement dans d'autres tableaux.
- ▶ Cliquez sur Enregistrer sous.
- ▶ Accédez au répertoire cible souhaité et saisissez un nom pour le nouveau style dans la zone Nom du fichier.
- ▶ Cliquez sur Enregistrer.
- ▶ Cliquez sur OK pour appliquer vos modifications et revenir dans le Viewer.

Le tableau contient à présent le format personnalisé que vous avez spécifié.

Figure 7-16
Modèle de tableau personnalisé

The screenshot shows the SPSS 'viewertut.spo - Editeur de résultats' window. At the top, there is a menu bar with options: 'Fichier', 'Edition', 'Affichage', 'Insérer', 'Tableau pivotant', 'Format', 'Analyse', 'Graphes', 'Outils', 'Fenêtre', and 'Aide'. Below the menu bar, the title 'Statistiques' is displayed. Underneath, a small table shows the distribution of 'Statut marital':

Statut marital		
N	Valide	6400
	Manquante	0

Below this, a larger pivot table is shown with the title 'Statut marital'. The table has five columns: 'Fréquence', 'Pourcentage', 'Pourcentage valide', and 'Pourcentage cumulé'. The rows represent marital status categories. A red arrow points to the first column of this table.

Statut marital					
		Fréquence	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	0 Célibataire	3224	50,4	50,4	50,4
	1 Marié(e)	3176	49,6	49,6	100,0
	Total	6400	100,0	100,0	

Modification des formats par défaut d'un tableau

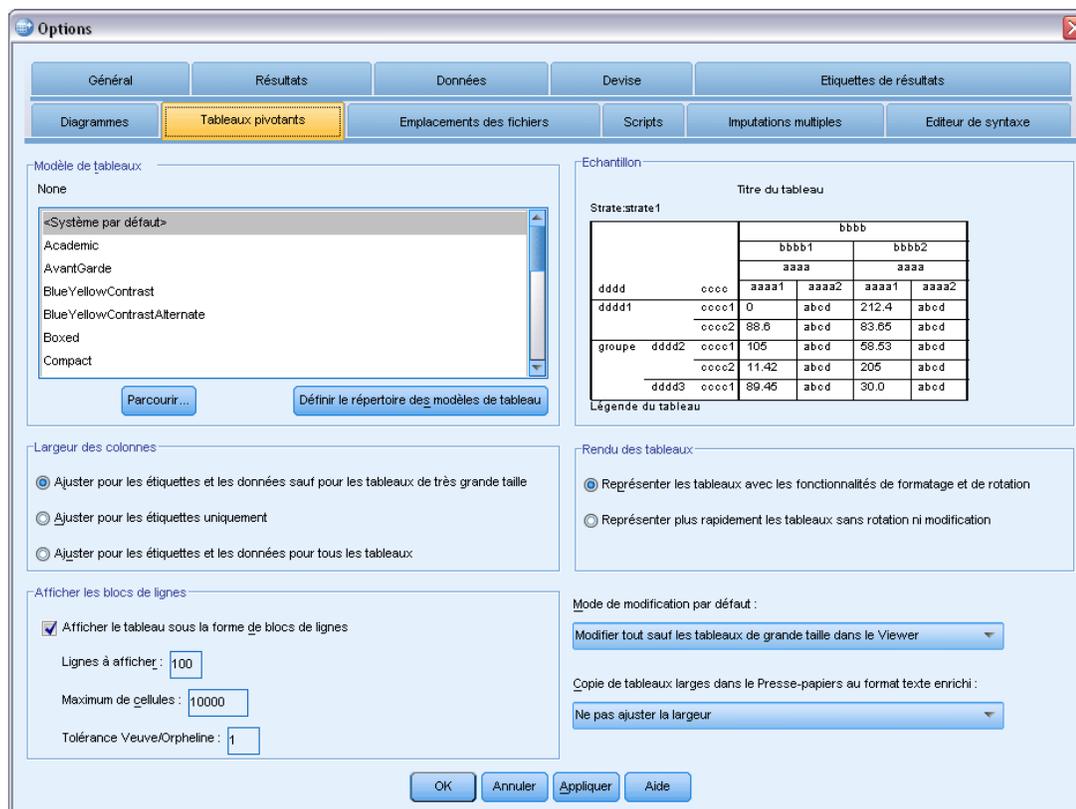
Bien que vous puissiez modifier le format des tableaux créés, il peut s'avérer plus efficace de modifier une fois pour toutes le modèle de tableau par défaut.

Pour modifier le style du modèle de tableau par défaut pour vos tableaux pivotants, à partir des menus, sélectionnez :

Edition > Options...

- Cliquez sur l'onglet Tableaux pivotants de la boîte de dialogue Options.

Figure 7-17
Boîte de dialogue Options



- Sélectionnez le style du modèle de tableau à utiliser pour tous les nouveaux tableaux.

La fenêtre Echantillon à droite affiche un aperçu de chaque modèle de tableau.

- Cliquez sur OK pour enregistrer les paramètres et fermer la boîte de dialogue.

Tous les tableaux créés après la modification du modèle de tableau par défaut suivent automatiquement les nouvelles règles de formatage.

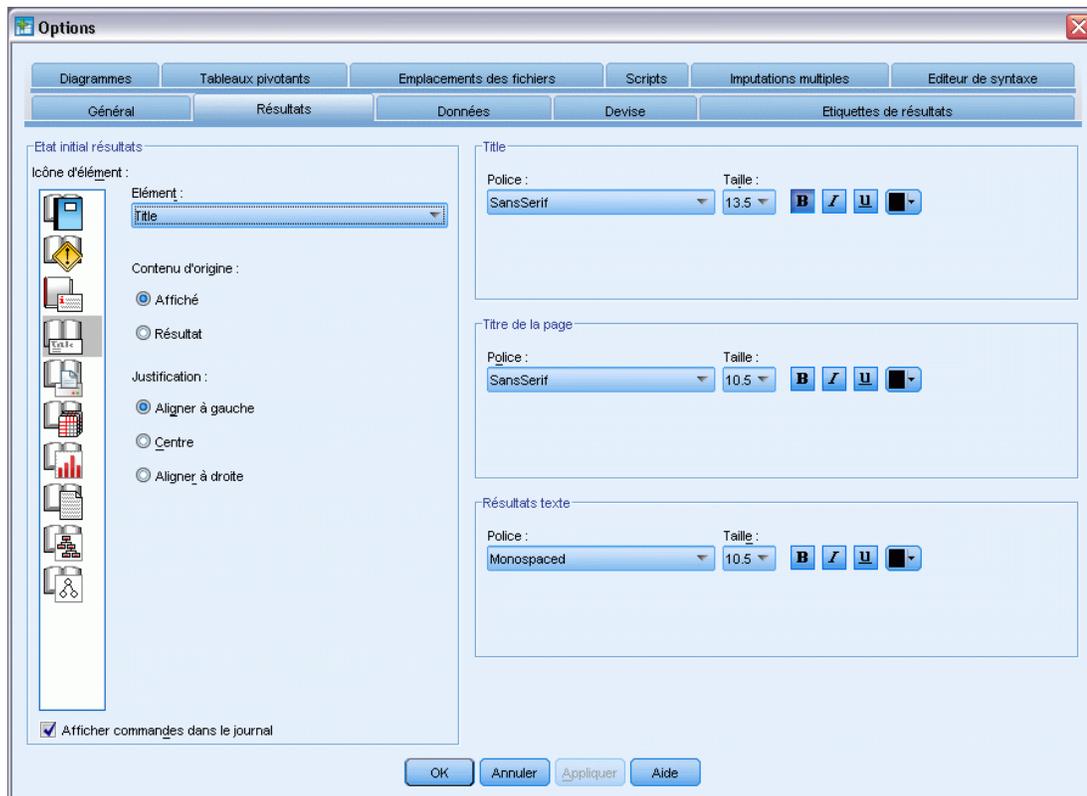
Personnalisation des paramètres de l'affichage initial

Les paramètres de l'affichage initial comprennent l'alignement des objets dans le Viewer, qu'ils soient affichés ou masqués par défaut, et la largeur de la fenêtre du Viewer. Pour modifier ces paramètres :

- A partir des menus, sélectionnez :
Edition > Options...

- Cliquez sur l'onglet Viewer.

Figure 7-18
Options du Viewer



Vos paramètres sont appliqués objet par objet. Vous pouvez, par exemple, personnaliser le mode d'affichage des diagrammes sans modifier celui des tableaux. Il suffit de sélectionner l'objet à personnaliser et d'apporter les modifications souhaitées.

- Cliquez sur l'icône Titre pour afficher ses paramètres.
- Cliquez sur Centre pour afficher tous les titres au centre (horizontal) du Viewer.

Vous pouvez également masquer des items, tels que le journal et les messages d'avertissement, qui ont tendance à rendre les résultats confus. Double-cliquez sur une icône pour modifier automatiquement la propriété d'affichage de cet objet.

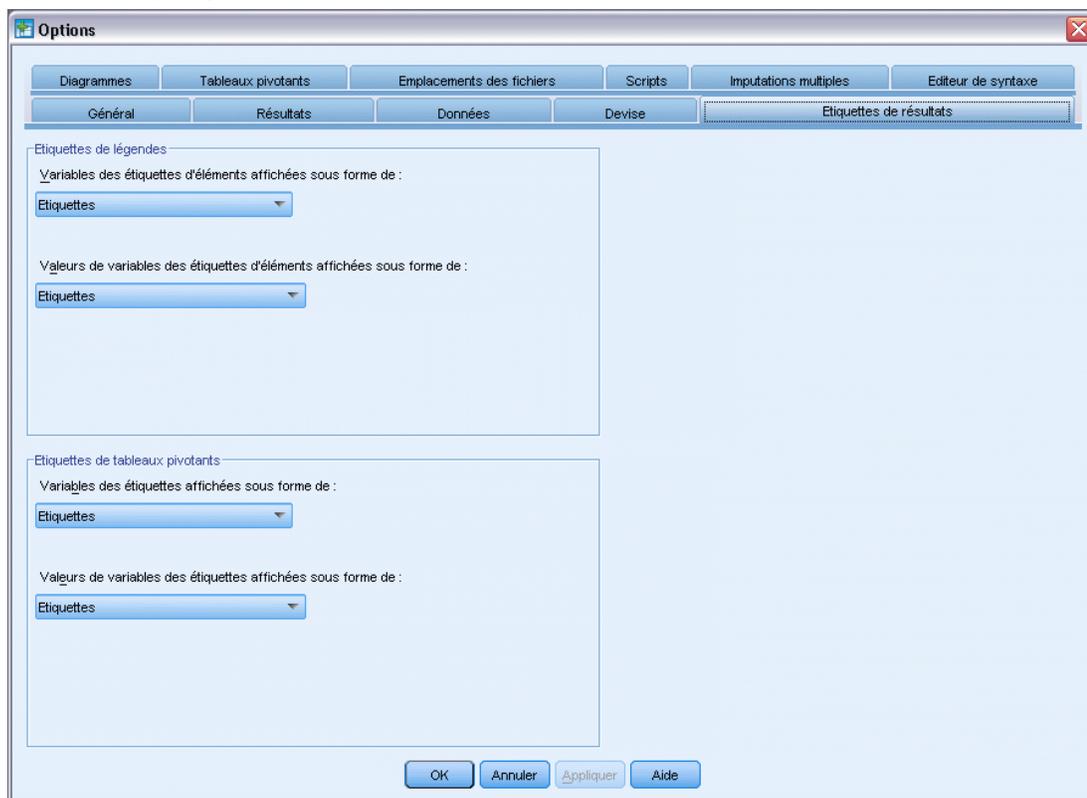
- Double-cliquez sur l'icône Avertissement pour masquer les messages d'avertissement dans les résultats.
- Cliquez sur OK pour enregistrer vos modifications et fermer la boîte de dialogue.

Affichage d'étiquettes de variable et de valeur

Dans la plupart des cas, il est plus judicieux d'afficher les étiquettes des variables et des valeurs plutôt que le nom de la variable et la valeur réelle des données. Toutefois, il peut arriver que vous vouliez afficher les noms et les étiquettes.

- ▶ A partir des menus, sélectionnez :
Edition > Options...
- ▶ Cliquez sur l'onglet Etiquettes de résultats.

Figure 7-19
Options d'étiquetage des résultats

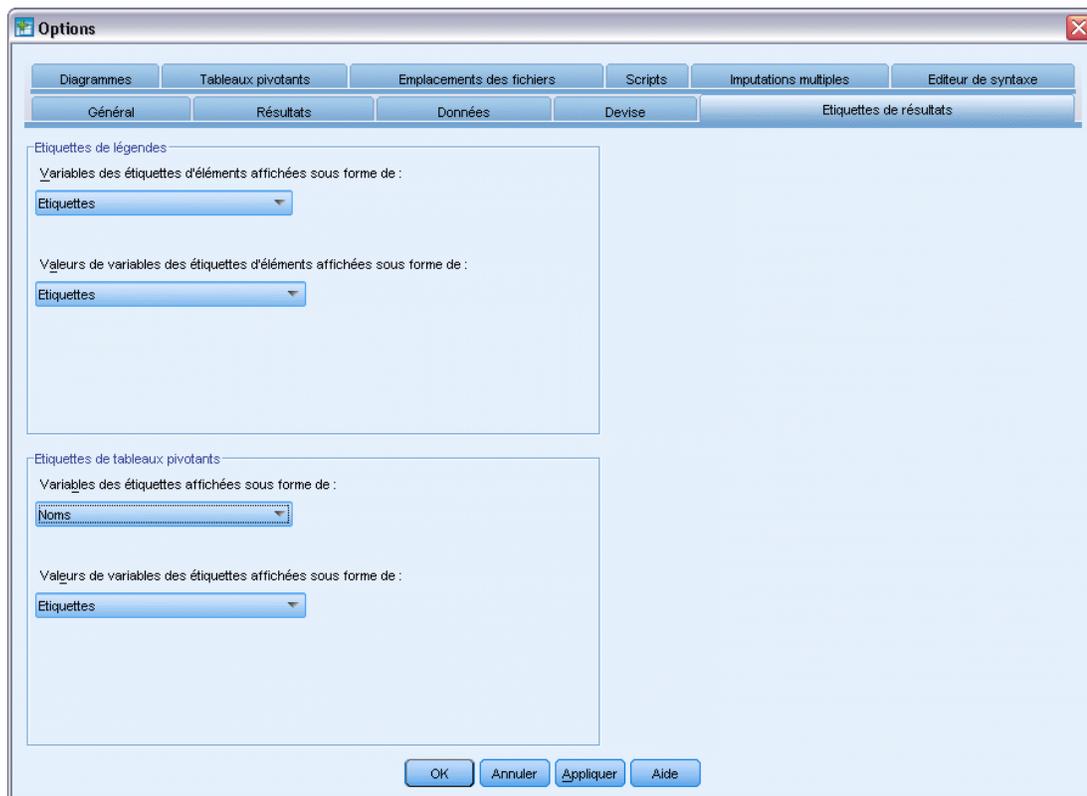


Vous pouvez indiquer des paramètres différents pour les panneaux de légende et de contenu. Par exemple, pour afficher des étiquettes dans la légende, et des noms de variable et des valeurs de données dans le contenu :

- ▶ Dans le groupe Etiquette tableau pivotant, sélectionnez Noms dans la liste déroulante Noms des variables en étiquettes pour afficher des noms de variable à la place des étiquettes.

- Sélectionnez ensuite Valeurs dans la liste déroulante Valeurs des variables en étiquettes pour afficher des valeurs de données à la place des étiquettes.

Figure 7-20
Paramètres Etiquette tableau pivotant



Les tableaux suivants générés au cours de la session refléteront ces modifications.

Figure 7-21
Noms et valeurs de variable affichés

Statut marital					
		Fréquence	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	0 Célibataire	3224	50,4	50,4	50,4
	1 Marié(e)	3176	49,6	49,6	100,0
	Total	6400	100,0	100,0	

Utilisation de résultats dans d'autres applications

Vos résultats peuvent être utilisés dans de nombreuses applications. Vous pouvez, par exemple, inclure un diagramme ou un tableau dans une présentation ou un rapport.

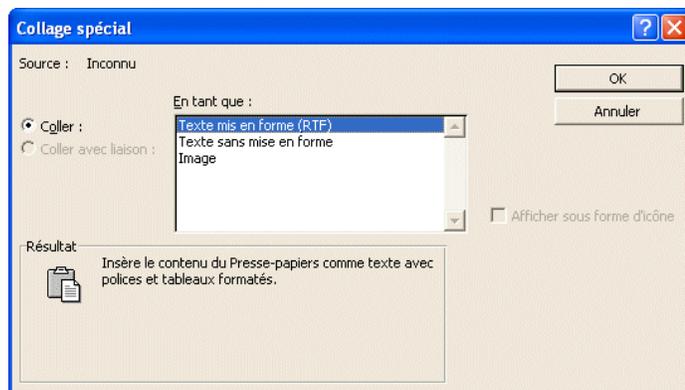
Les exemples suivants sont propres à Microsoft Word, mais ils peuvent fonctionner de la même façon dans d'autres applications de traitement de texte.

Collage des résultats sous forme de tableaux Word

Vous pouvez coller des tableaux pivotants dans Word en tant que tableaux natifs Word. Tous les attributs de tableau, comme la taille et la couleur de la police, sont conservés. Du fait que le tableau est collé dans le format de tableau Word, vous pouvez le modifier dans Word comme tout autre tableau.

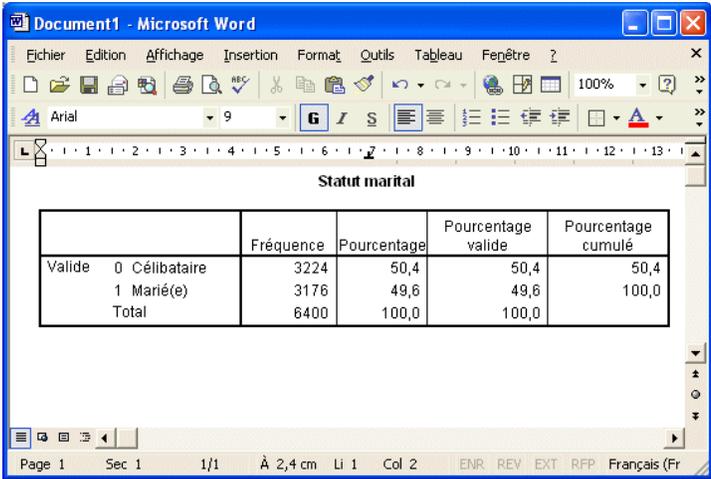
- ▶ Cliquez sur le tableau *Situation familiale* dans le Viewer.
- ▶ A partir des menus, sélectionnez :
Edition > Copier
- ▶ Ouvrez l'application de traitement de texte.
- ▶ A partir des menus du traitement de texte, sélectionnez :
Edition > Collage spécial
- ▶ Sélectionnez Texte mis en forme (RTF) dans la boîte de dialogue Collage spécial.

Figure 7-22
Boîte de dialogue Collage spécial



- ▶ Cliquez sur OK pour coller vos résultats dans le document actuel.

Figure 7-23
Tableau pivotant affiché dans Word



The screenshot shows a Microsoft Word window titled 'Document1 - Microsoft Word'. The menu bar includes 'Fichier', 'Edition', 'Affichage', 'Insertion', 'Format', 'Outils', 'Tableau', and 'Fenêtre'. The toolbar shows various icons for file operations and editing. The font is set to Arial, size 9. The pivot table is titled 'Statut marital' and is displayed in a grid format. The table has four columns: 'Fréquence', 'Pourcentage', 'Pourcentage valide', and 'Pourcentage cumulé'. The rows represent different marital statuses: 'Valide 0 Célibataire', '1 Marié(e)', and 'Total'. The data is as follows:

	Fréquence	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide 0 Célibataire	3224	50,4	50,4	50,4
1 Marié(e)	3176	49,6	49,6	100,0
Total	6400	100,0	100,0	

The status bar at the bottom indicates 'Page 1', 'Sec 1', '1/1', 'À 2,4 cm', 'Li 1', 'Col 2', 'ENR', 'REV', 'EXT', 'RFP', and 'Français (Fr)'.

Le tableau est à présent affiché dans votre document. Vous pouvez appliquer des formats personnalisés, modifier les données et redimensionner le tableau en fonction de vos besoins.

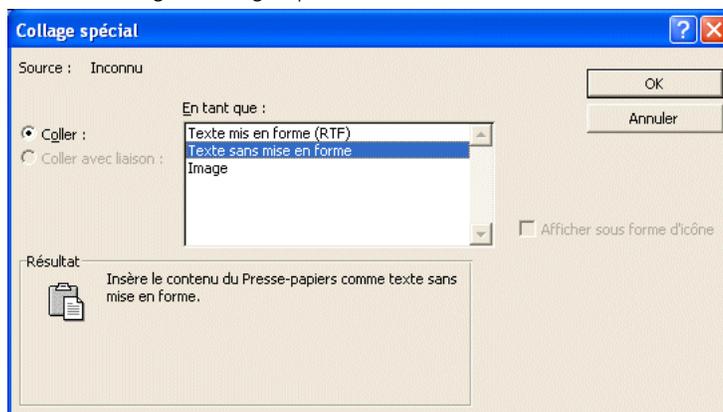
Collage des résultats sous forme de texte

Les tableaux pivotants peuvent être copiés dans d'autres applications sous forme de texte brut. Avec cette méthode, les styles de format ne sont pas conservés, mais vous pouvez modifier les données du tableau après l'avoir collé dans l'application cible.

- ▶ Cliquez sur le tableau *Situation familiale* dans le Viewer.
- ▶ A partir des menus, sélectionnez :
Edition > Copier
- ▶ Ouvrez l'application de traitement de texte.
- ▶ A partir des menus du traitement de texte, sélectionnez :
Edition > Collage spécial

- Sélectionnez Texte sans mise en forme dans la boîte de dialogue Collage spécial.

Figure 7-24
Boîte de dialogue Collage spécial



- Cliquez sur OK pour coller vos résultats dans le document actuel.

Figure 7-25
Tableau pivotant affiché dans Word

Statut marital	Fréquence	Pourcentage	Pourcentage valide	Pourcentage
Valide				
0 Célibataire	3224	50,4	50,4	50,4
1 Marié(e)	3176	49,6	49,6	100,0
Total	6400	100,0	100,0	

Chaque colonne du tableau est séparée par des tabulations. Vous pouvez modifier la largeur des colonnes en ajustant les marques de tabulation dans votre application de traitement de texte.

Exportation des résultats vers un fichier Microsoft Word, PowerPoint ou Excel

Vous pouvez exporter les résultats vers un fichier Microsoft Word, PowerPoint ou Excel. Vous pouvez exporter les éléments sélectionnés ou tous les éléments du Viewer. Cette section utilise les fichiers *msouttut.spv* et *demo.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.](#)

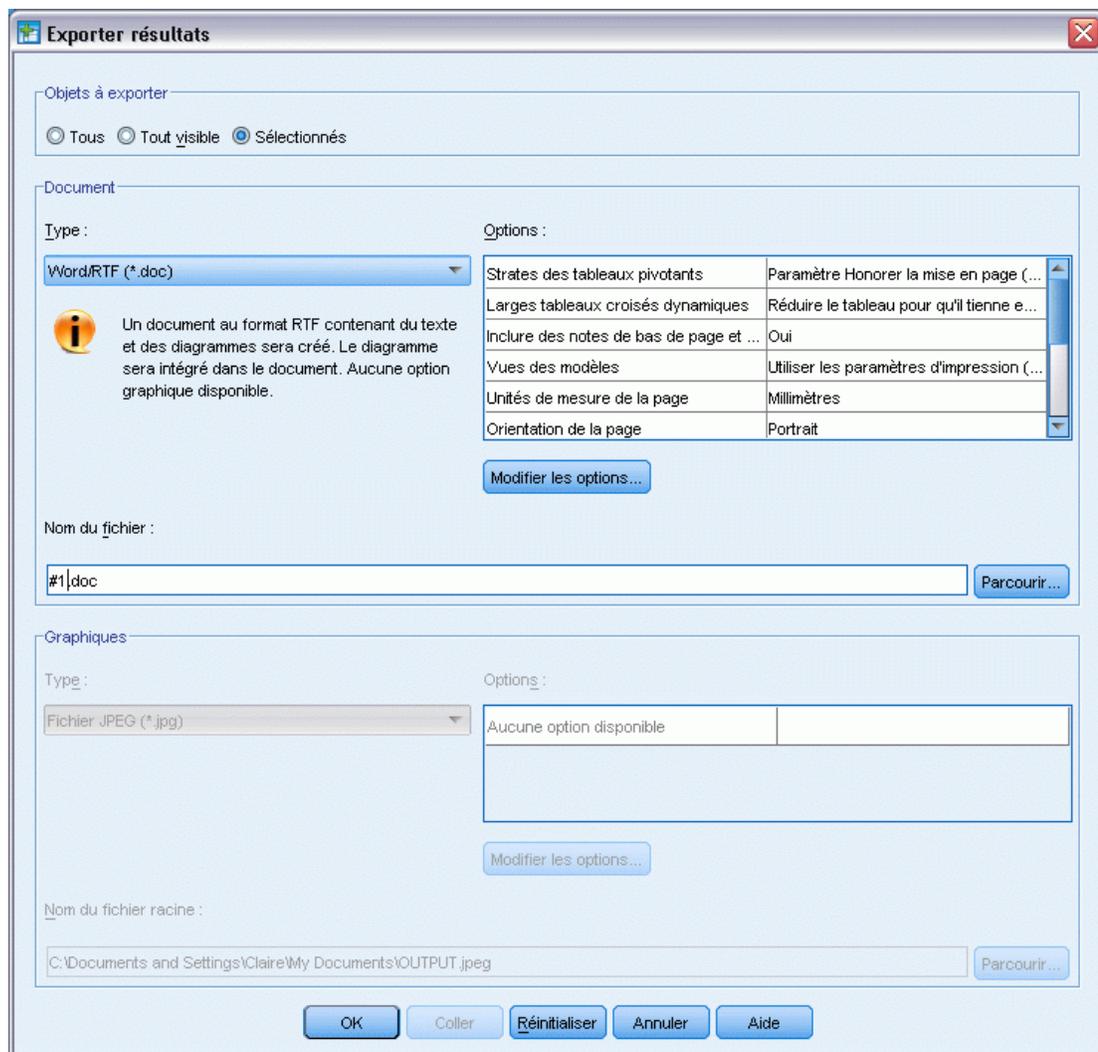
Remarque : L'exportation vers PowerPoint n'est disponible que sous les systèmes d'exploitation Windows et n'est pas disponible dans la version Student.

Dans le panneau de légende du Viewer, vous pouvez sélectionner les éléments à exporter. La sélection d'éléments spécifiques n'est pas obligatoire.

- A partir du menu du Viewer, sélectionnez :
Fichier > Exporter...

Au lieu d'exporter tous les objets du Viewer, vous pouvez n'exporter que les objets visibles (livres ouverts dans le panneau de légende) ou ceux sélectionnés dans le panneau de légende. Si vous n'avez sélectionné aucun élément du panneau de légende, vous n'avez pas la possibilité d'exporter les objets sélectionnés.

Figure 7-26
Boîte de dialogue Exporter résultats

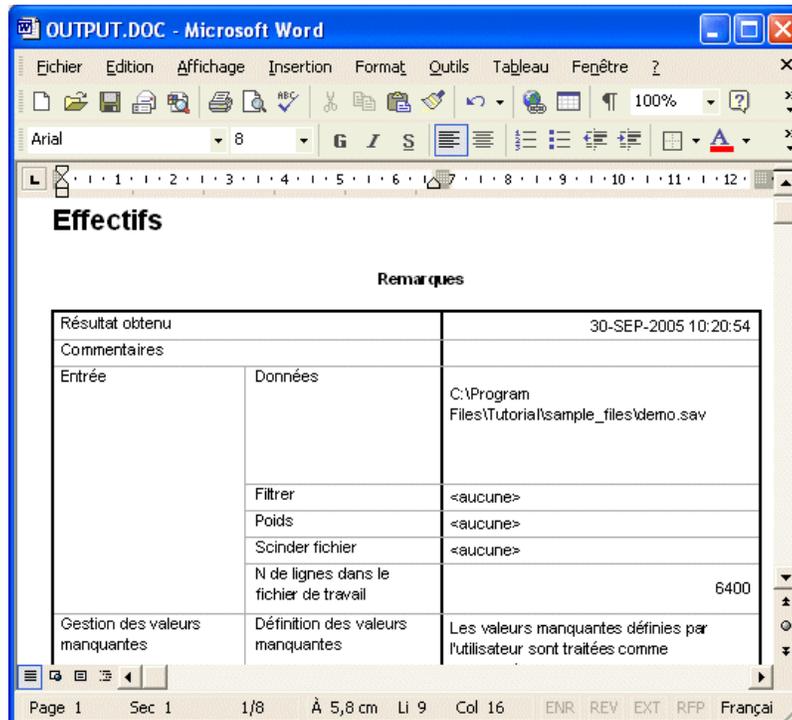


- Dans le groupe Objets à exporter, sélectionnez Tous.

- ▶ Dans la liste déroulante Type, sélectionnez Word/RTF (*.doc).
- ▶ Cliquez sur OK pour générer le fichier Word.

Lorsque vous ouvrez le fichier obtenu dans Word, vous pouvez observer le mode d'exportation des résultats. Les notes, qui ne sont pas des objets visibles, apparaissent dans Word car vous avez choisi d'exporter tous les objets.

Figure 7-27
Output.doc dans Word



Les tableaux pivotants deviennent des tableaux Word et conservent l'intégralité de leur formatage d'origine (polices, couleurs, bordures, etc.).

Figure 7-28
Tableaux pivotants dans Word

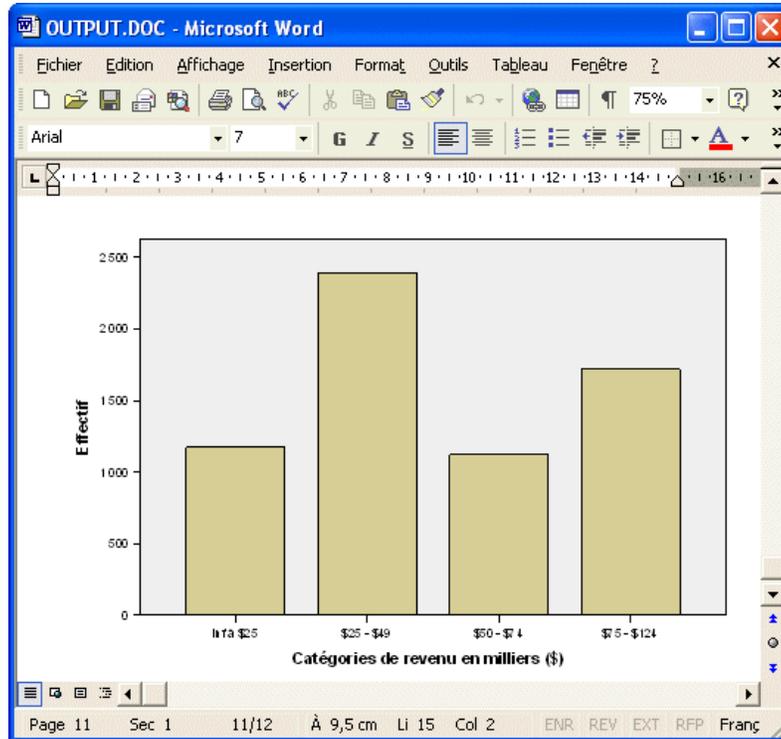
The screenshot shows a Microsoft Word window titled 'OUTPUT.DOC - Microsoft Word'. The window contains two pivot tables. The first table is titled 'sexe' and the second is titled 'Catégories de revenu en milliers (\$)'. Both tables show frequency, percentage, and cumulative percentage data.

		Fréquence	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	f	3179	49,7	49,7	49,7
	m	3221	50,3	50,3	100,0
Total		6400	100,0	100,0	

		Fréquence	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	1	1174	18,3	18,3	18,3
	2	2388	37,3	37,3	55,7
	3

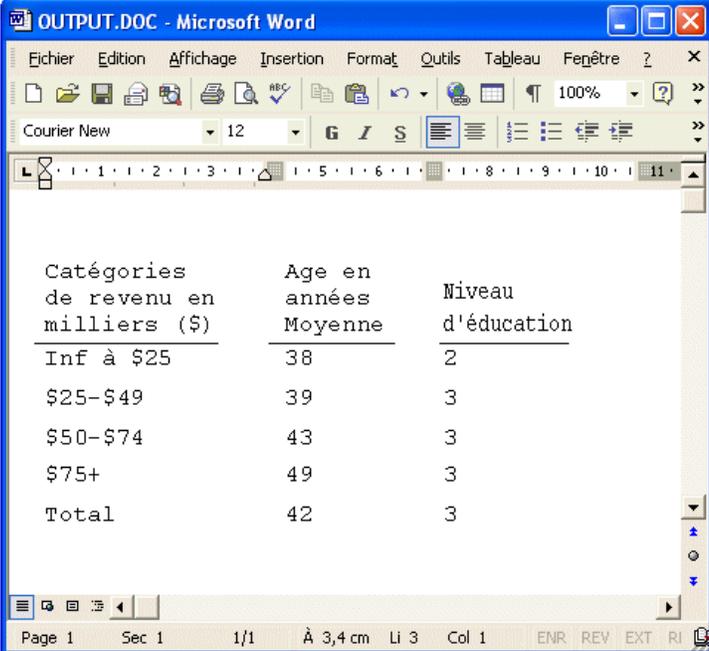
Les diagrammes sont inclus dans le document Word sous la forme d'images graphiques.

Figure 7-29
Diagrammes dans Word



Les résultats sous forme de texte sont affichés avec la même police que celle utilisée pour l'objet texte dans le Viewer. Tous les résultats de texte nécessitent une police à point fixe (espacement unique) pour un alignement correct.

Figure 7-30
Résultats texte dans Word



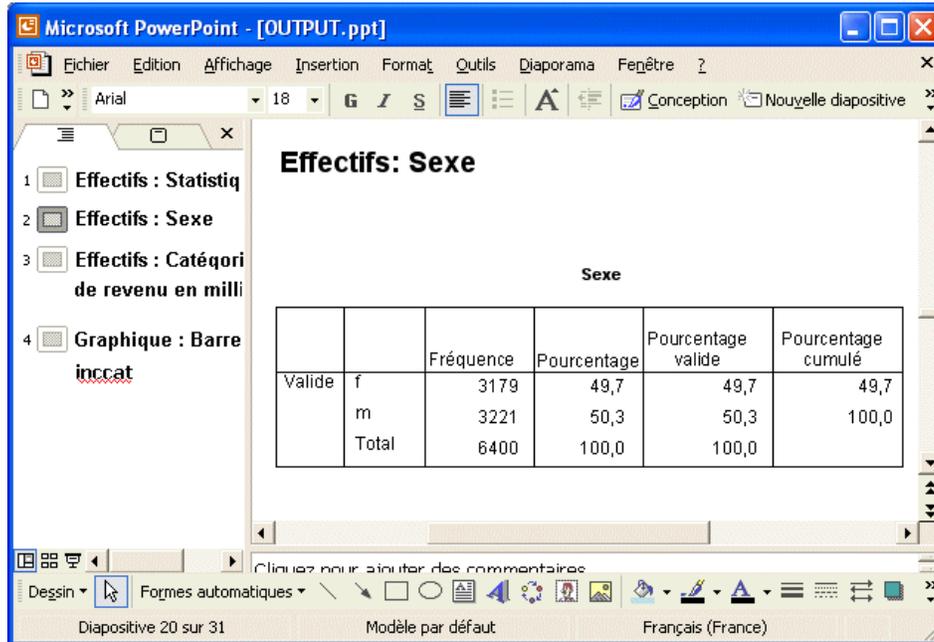
The screenshot shows a Microsoft Word window titled 'OUTPUT.DOC - Microsoft Word'. The window displays a table with three columns: 'Catégories de revenu en milliers (\$)', 'Age en années Moyenne', and 'Niveau d'éducation'. The table data is as follows:

Catégories de revenu en milliers (\$)	Age en années Moyenne	Niveau d'éducation
Inf à \$25	38	2
\$25-\$49	39	3
\$50-\$74	43	3
\$75+	49	3
Total	42	3

The status bar at the bottom of the window indicates 'Page 1', 'Sec 1', '1/1', 'À 3,4 cm', 'Li 3', 'Col 1', and 'ENR REV EXT RI'.

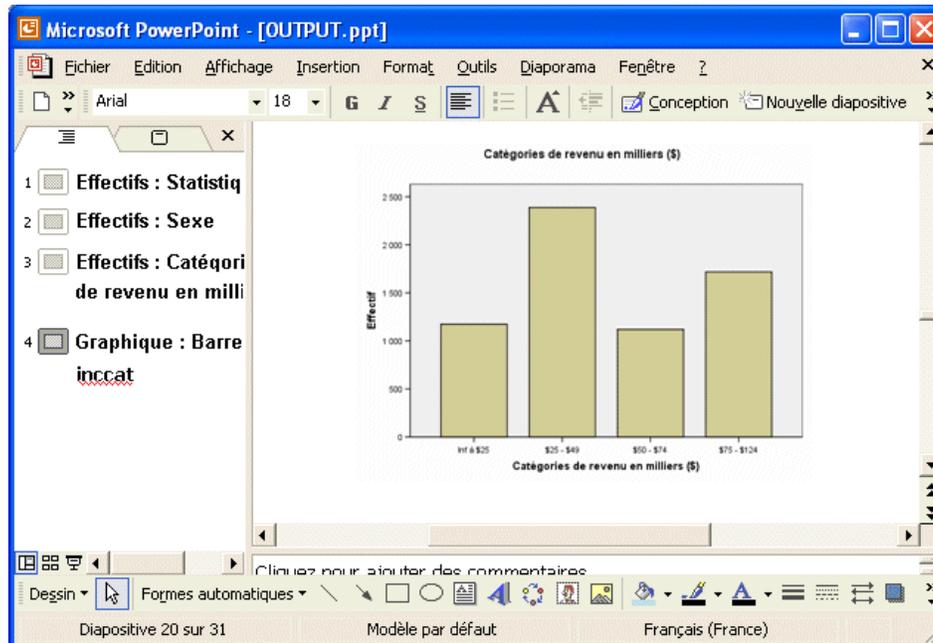
Si vous exportez des données vers un fichier PowerPoint, chaque élément exporté est placé sur une diapositive distincte. Les tableaux pivotants exportés dans PowerPoint deviennent des tableaux Word et conservent l'intégralité du formatage d'origine (polices, couleurs, bordures, etc.).

Figure 7-31
Tableaux pivotants dans PowerPoint



Les diagrammes sélectionnés pour l'exportation vers PowerPoint sont intégrés dans le fichier PowerPoint.

Figure 7-32
Diagrammes dans PowerPoint



Si vous exportez des résultats vers un fichier Excel, l'exportation est réalisée de manière différente.

Figure 7-33
Output.xls dans Excel

Sexe					
		Fréquence	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	F	3 179	49,7	49,7	49,7
	M	3 221	50,3	50,3	100,0
	Total	6 400	100,0	100,0	
Catégories de revenu en milliers (\$)					
		Fréquence	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	Inf à \$25	1 174	18,3	18,3	18,3
	\$25 - \$49	2 388	37,3	37,3	55,7
	\$50 - \$74	1 120	17,5	17,5	73,2
	\$75+	1 718	26,8	26,8	100,0
	Total	6 400	100,0	100,0	

Les lignes, les colonnes et les cellules des tableaux pivotants deviennent des lignes, des colonnes et des cellules Excel.

Chaque ligne du résultat texte est une ligne dans le fichier Excel, avec le contenu de toute la ligne dans une seule cellule.

Figure 7-34
Résultats texte dans Excel

The screenshot shows a Microsoft Excel window titled 'Microsoft Excel - OUTPUT.xls'. The spreadsheet contains a pivot table with the following data:

	A	B	C	D	E	F
68	Catégories					
69	de revenu	Age en				
70	en	années	Niveau			
71	milliers	Moyenne	d'éducation			
72						
73						
74	Inf à \$25	38	2			
75						
76	\$25 - \$49	39	3			
77						
78	\$50 - \$74	43	3			
79						
80	\$75+	49	3			
81						
82	Total	42	3			

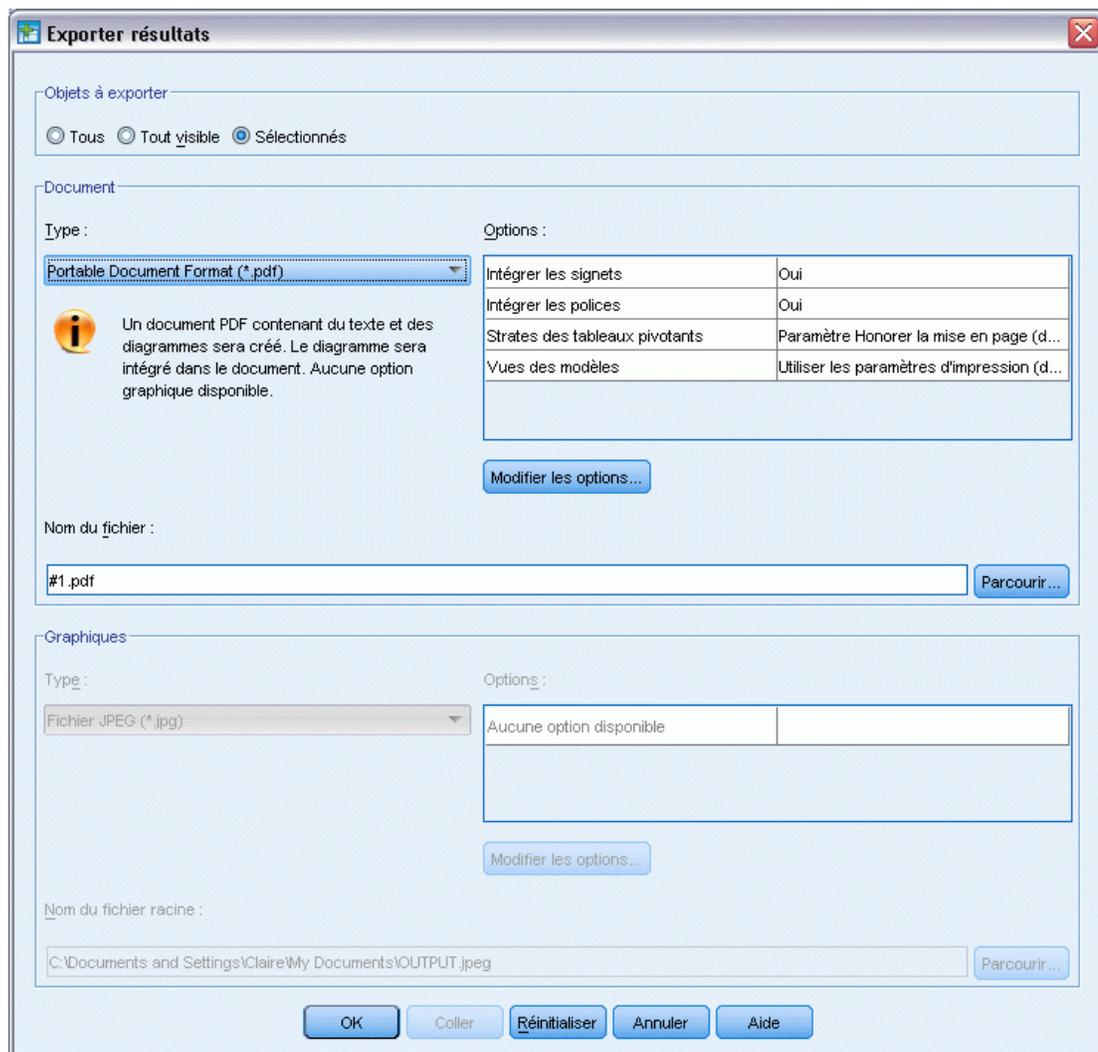
Exportation des résultats vers PDF

Vous pouvez exporter les éléments sélectionnés ou tous les éléments du Viewer dans un fichier PDF (Portable Document Format).

- A partir des menus de la fenêtre du Viewer contenant le résultat à exporter vers un fichier PDF, sélectionnez :
Fichier > Exporter...

- Dans la liste déroulante Type de fichier Format d'exportation de la boîte de dialogue Exporter résultats, sélectionnez Portable Document Format.

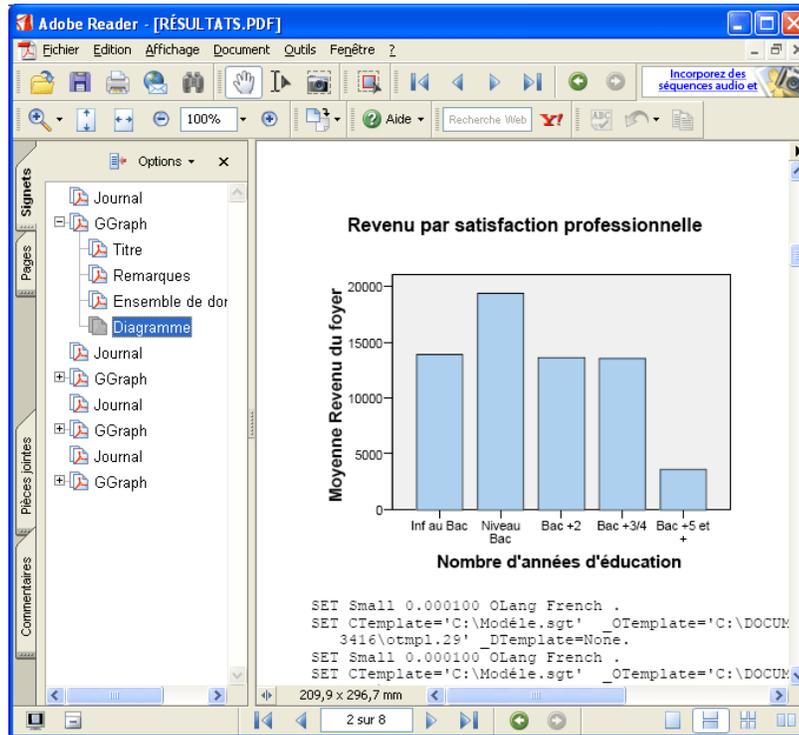
Figure 7-35
Boîte de dialogue Exporter résultats



- Le panneau de légende du document du Viewer est converti en signets dans le fichier PDF pour une navigation plus facile.
- La taille de la page et son orientation, les marges, le contenu et l'affichage des en-têtes et des pieds de page, ainsi que la taille des graphiques imprimés d'un document PDF sont gérés par les options de mise en page (menu Fichier, option Mise en page de la fenêtre du Viewer).
- La résolution (ppp) d'un document PDF est le paramètre de résolution courant de l'imprimante sélectionnée ou par défaut (vous pouvez changer l'imprimante via l'option Mise en page). La résolution maximale est de 1 200 ppp. Si le paramètre de résolution de l'imprimante est supérieur, la résolution du document PDF sera de 1 200 ppp. *Remarque* :

Une haute résolution peut générer des résultats médiocres si les documents sont imprimés sur des imprimantes de résolution inférieure.

Figure 7-36
fichier PDF avec signets



Exportation des résultats vers HTML

Vous pouvez également exporter des résultats vers HTML. Lorsque vous enregistrez les résultats au format HTML, tous les résultats non graphiques sont exportés dans un fichier HTML unique.

Figure 7-37
Output.htm dans un navigateur Web

The screenshot shows a web browser window displaying two tables of data. The first table is titled "Sexe" and the second is titled "Catégories de revenu en milliers". Both tables have columns for "Fréquence", "Pourcentage", "Pourcentage valide", and "Pourcentage cumulé".

		Fréquence	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	F	3179	49,7	49,7	49,7
	M	3221	50,3	50,3	100,0
	Total	6400	100,0	100,0	

		Fréquence	Pourcentage	Pourcentage valide	Pourcentage cumulé
Valide	Inf à \$25	1174	18,3	18,3	18,3
	\$25 - \$49	2388	37,3	37,3	55,7
	\$50 - \$74	1120	17,5	17,5	73,2
	\$75+	1718	26,8	26,8	100,0
	Total	6400	100,0	100,0	

Lorsque vous exportez les résultats vers le format HTML, vous pouvez également exporter des diagrammes, mais pas vers un fichier unique.

Chaque diagramme sera enregistré sous la forme d'un fichier au format que vous spécifiez, et les références à ces fichiers graphiques seront placées dans le document HTML. Une option vous permet également d'exporter tous les diagrammes (ou seulement certains) dans des fichiers graphiques distincts.

Utilisation de la syntaxe

Vous pouvez enregistrer et automatiser de nombreuses tâches courantes grâce au puissant langage de commande. Il fournit également des fonctionnalités qui ne se trouvent ni dans les menus ni dans les boîtes de dialogue. La plupart des commandes sont accessibles depuis les menus et boîtes de dialogue. Cependant, certaines options et commandes ne sont disponibles qu'en utilisant le langage de commande. Le langage de commande permet également d'enregistrer vos travaux dans un fichier de syntaxe afin de vous permettre de relancer votre analyse à une date ultérieure.

Un fichier de syntaxe de commande est un fichier texte simple contenant des commandes de syntaxe IBM® SPSS® Statistics. Vous pouvez ouvrir une fenêtre de syntaxe et y entrer des commandes directement, mais il est souvent plus simple de laisser la boîte de dialogue faire une partie ou tout le travail pour vous.

Les exemples de ce chapitre utilisent le fichier de données *demo.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.](#)

Remarque : La syntaxe de commande n'est pas disponible dans la version Student.

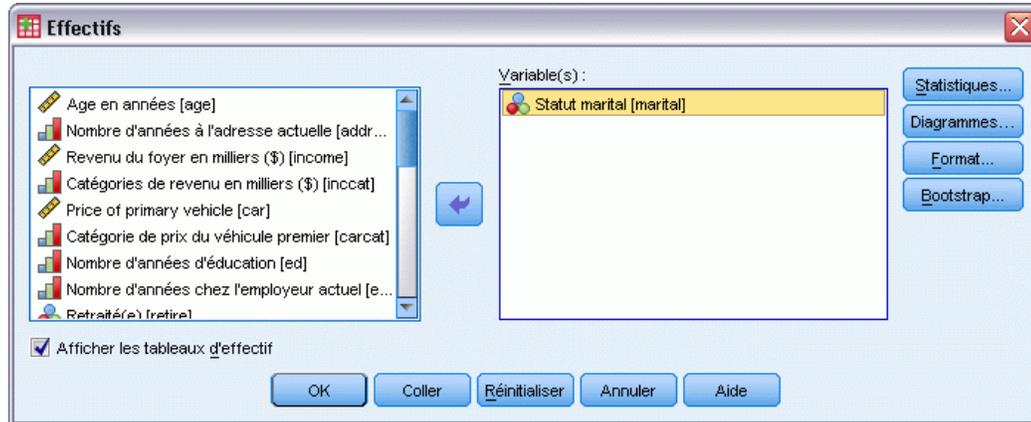
Collage d'une syntaxe

Le bouton Coller, affiché dans la plupart des boîtes de dialogue, constitue le moyen le plus simple de créer une syntaxe.

- ▶ Ouvrez le fichier de données *demo.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.](#)
- ▶ A partir des menus, sélectionnez :
Analyse > Statistiques descriptives > Effectifs

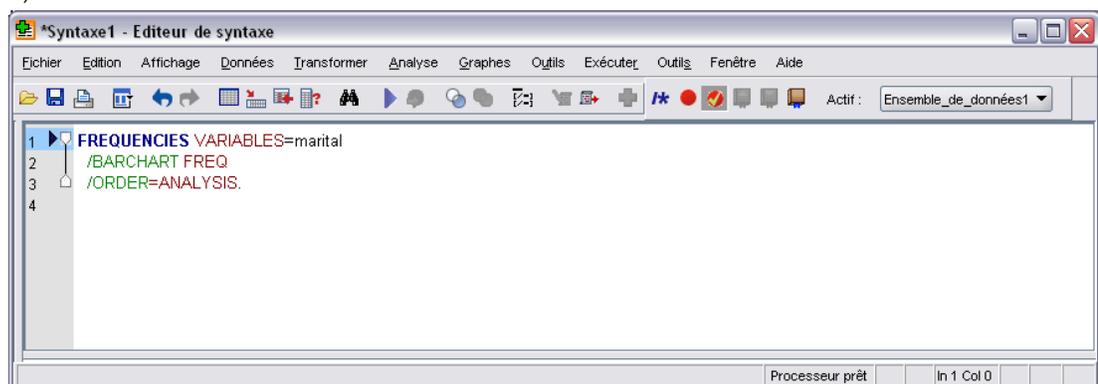
La boîte de dialogue Effectifs s'ouvre.

Figure 8-1
Boîte de dialogue Effectifs



- ▶ Sélectionnez la variable *Situation familiale [situation]* et déplacez-la vers la liste Variable(s).
- ▶ Cliquez sur Diagrammes.
- ▶ Dans la boîte de dialogue Diagrammes, sélectionnez Diagrammes en bâtons.
- ▶ Dans le groupe Valeurs du diagramme, sélectionnez Pourcentages.
- ▶ Cliquez sur Poursuivre.
- ▶ Cliquez sur Coller pour copier la syntaxe créée grâce aux sélections effectuées dans la boîte de dialogue de l'Editeur de syntaxe.

Figure 8-2
Syntaxe d'effectifs

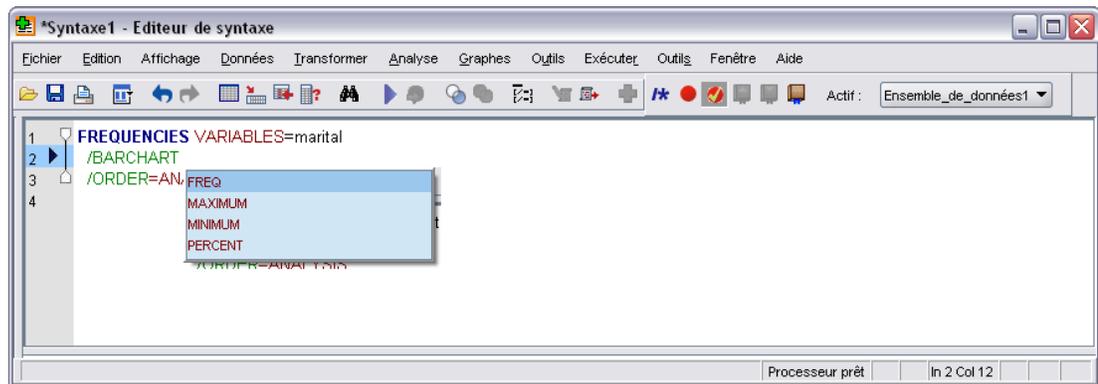


- ▶ Pour exécuter la syntaxe actuellement affichée, à partir des menus, sélectionnez :
Exécuter > Sélection

Modification d'une syntaxe

Dans la fenêtre de syntaxe, vous pouvez modifier la syntaxe. Par exemple, vous pouvez modifier la sous-commande `/BARCHART` pour afficher des effectifs à la place des pourcentages. (Une sous-commande est indiquée par une barre oblique.) Si vous connaissez le mot-clé pour afficher les effectifs, vous pouvez le saisir directement. Si vous ne le connaissez pas, vous pouvez obtenir une liste des mots-clés disponibles pour la sous-commande en positionnant le curseur n'importe où après le nom de la sous-commande et en appuyant simultanément sur `Ctrl+barre d'espace`. Ceci affiche le contrôle de saisie semi-automatique pour la sous-commande.

Figure 8-3
Contrôle de saisie semi-automatique affichant les mots-clés



- Cliquez sur l'élément étiqueté `FREQ` pour les effectifs. En cliquant sur un élément dans le contrôle de saisie semi-automatique, il sera inséré à l'emplacement actuel du curseur (le mot-clé initial `PERCENT` a été supprimé manuellement).

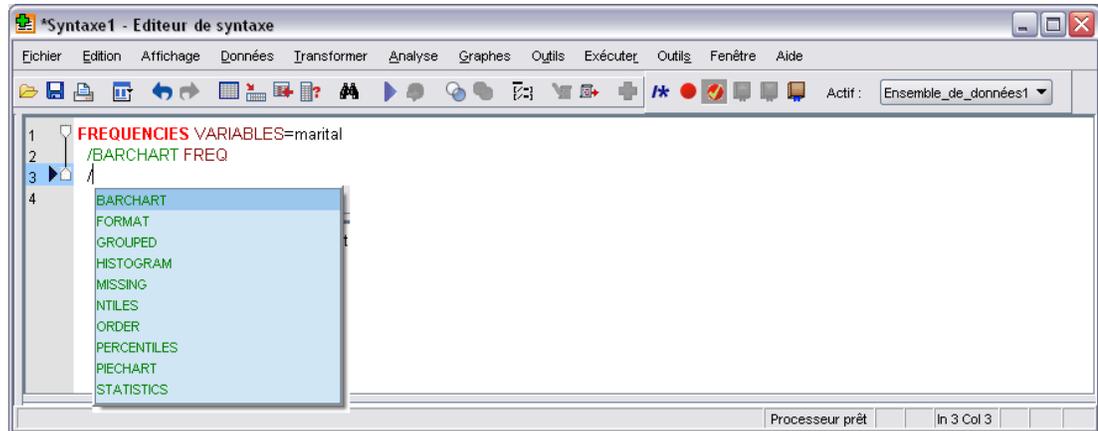
Par défaut, le contrôle de saisie semi-automatique vous proposera une liste des termes disponibles au fur et à mesure que vous saisissez. Supposons par exemple que vous souhaiteriez inclure un diagramme en secteurs avec un diagramme en bâtons. Le diagramme en secteurs est spécifié par une sous-commande distincte.

- Appuyez sur `Entrée` après le mot-clé `FREQ` et tapez une barre oblique pour indiquer le début de la sous-commande.

L'Éditeur de syntaxe vous propose une liste des sous-commandes pour la commande en cours.

Figure 8-4

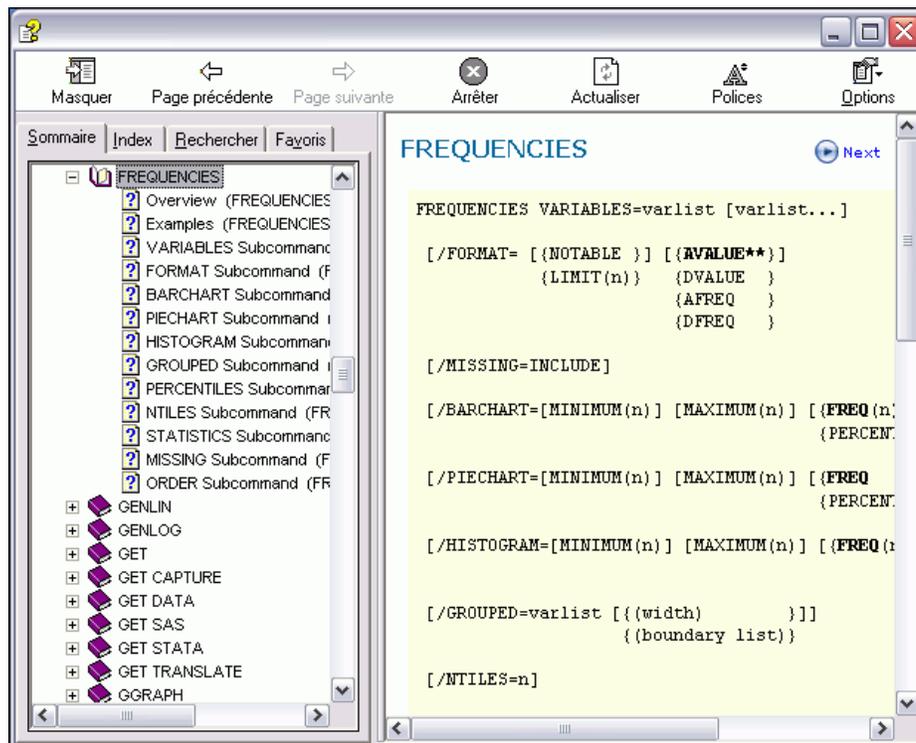
Contrôle de saisie semi-automatique affichant les sous-commandes



Pour obtenir plus d'aide sur la commande en cours, appuyez sur la touche F1. Vous êtes directement dirigé vers un guide de référence détaillé de la syntaxe de commande en cours.

Figure 8-5

Aide relative à la syntaxe de la commande FREQUENCIES



Vous avez peut-être remarqué que le texte affiché dans la fenêtre de syntaxe est en couleur. Le codage par couleur vous permet d'identifier rapidement les termes non reconnus, puisque que seuls les termes reconnus sont en couleur. Supposons par exemple que vous ayez mal orthographié

la sous-commande `FORMAT` en `FRMAT`. Les sous-commandes sont en vert par défaut, mais le texte `FRMAT` n'apparaîtra pas en couleur, puisqu'il n'est pas reconnu.

Ouverture et exécution d'un fichier de syntaxe

- Pour ouvrir un fichier de syntaxe enregistré, à partir du menu, sélectionnez :
Fichier > Ouvrir > Syntaxe...

Une boîte de dialogue standard d'ouverture de fichiers apparaît.

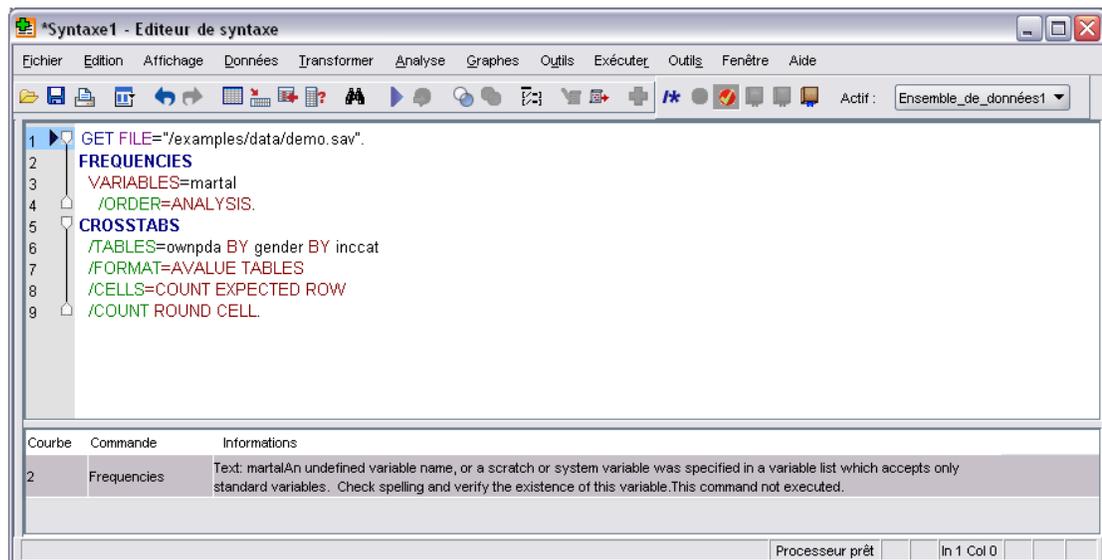
- Sélectionnez un fichier de syntaxe. Si aucun fichier de syntaxe n'apparaît, assurez-vous que l'option Syntaxe (*.sps) est sélectionnée en tant que type de fichier à afficher.
- Cliquez sur Ouvrir.
- Utilisez le menu Exécuter de la fenêtre de syntaxe pour exécuter les commandes.

Si les commandes s'appliquent à un fichier de données particulier, vous devez ouvrir ce dernier avant d'exécuter les commandes ou inclure une commande qui ouvre le fichier de données. Vous pouvez coller ce type de commande à partir des boîtes de dialogue permettant d'ouvrir les fichiers de données.

Présentation du Panneau d'erreur

Le panneau d'erreur affiche les erreurs d'exécution survenues lors de l'exécution la plus récente. Il contient les détails pour chaque erreur, ainsi que le numéro de la ligne de la commande dans laquelle l'erreur est survenue.

Figure 8-6
Panneau d'erreur affiché dans l'Editeur de syntaxe



Cliquez sur l'entrée d'une erreur positionne le curseur à la première ligne de la commande sur laquelle l'erreur est survenue.

Utilisation des points d'arrêt

Les points d'arrêt vous permettent d'interrompre l'exécution d'une syntaxe de commande à des points spécifiés à l'intérieur de la syntaxe et de poursuivre l'exécution lorsque vous êtes prêt. Ceci vous permet de visualiser les résultats ou les données à un moment intermédiaire de la tâche de syntaxe, ou d'exécuter la syntaxe de commande affichant les informations sur l'état actuel des données, telles que `FREQUENCIES`. Les points d'arrêt ne peuvent être réglés qu'au niveau d'une commande, et non sur les lignes spécifiques au sein d'une commande.

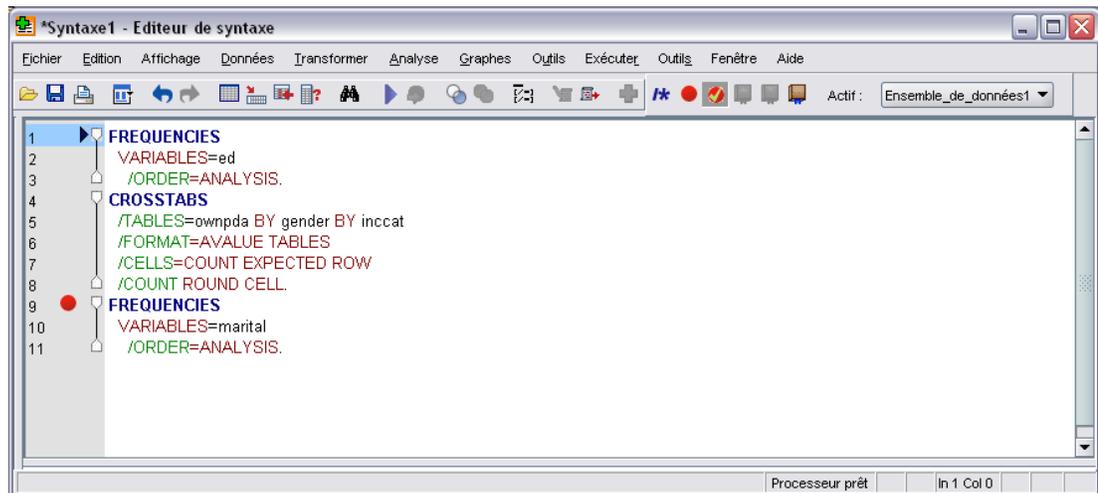
Pour insérer un point d'arrêt dans une commande :

- Cliquez n'importe où dans la zone à gauche du texte associé à la commande.

Le point d'arrêt est représenté par un cercle rouge dans la zone à gauche du texte de la commande et sur la même ligne que le nom de la commande, peu importe où vous cliquez.

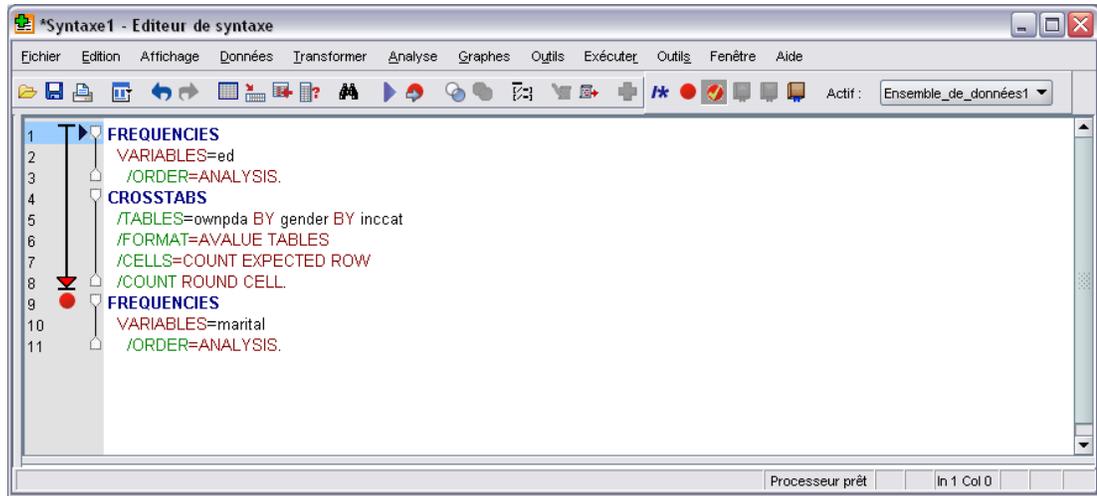
Figure 8-7

Point d'arrêt affiché dans la fenêtre de l'Editeur de syntaxe



Lorsque vous exécutez une syntaxe de commande contenant des points d'arrêt, l'exécution s'interrompt avant chaque commande contenant un point d'arrêt.

Figure 8-8
Exécution interrompue à un point d'arrêt



La flèche pointant vers le bas à gauche du texte de commande présente la progression de l'exécution de la syntaxe. Elle couvre la zone s'étendant de la première exécution de commande à la dernière exécution de commande. Cela s'avère très pratique lorsque l'exécution d'une syntaxe de commande contient des points d'arrêt.

Pour reprendre l'exécution après un point d'arrêt :

- Dans le menu de la fenêtre de l'Editeur de syntaxe, sélectionnez :
Exécuter > Poursuivre

Modification des valeurs de données

L'organisation des données de départ ne répond pas systématiquement à vos besoins en matière de création de rapport ou d'analyse. Par exemple, vous pouvez :

- Créer une variable qualitative à partir d'une variable d'échelle.
- Combiner plusieurs modalités de réponses en une seule modalité.
- Créer une nouvelle variable qui représente la différence calculée entre deux variables existantes.
- Calculer le temps écoulé entre deux dates.

Dans ce chapitre, nous utilisons le fichier de données *demo.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.](#)

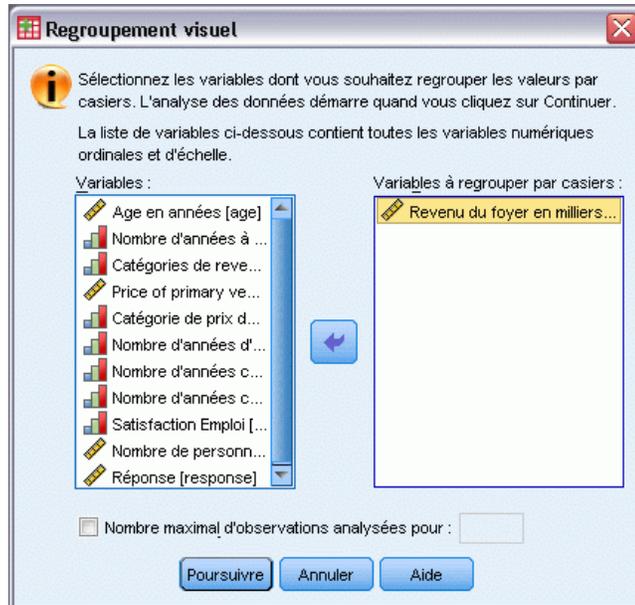
Création d'une variable qualitative à partir d'une variable d'échelle

Plusieurs variables catégorielles du fichier de données *demo.sav* sont dérivées de variables d'échelle de ce même fichier. Par exemple, la variable *inccat* est simplement *revenu*, réparti en quatre modalités. Cette variable qualitative utilise les valeurs entières comprises entre 1 et 4 pour représenter les modalités de revenus suivantes (en milliers) : inférieur à 25 €, 25 €–49 €, 50 €–74 € et 75 € ou plus.

Pour créer la variable qualitative *inccat* :

- ▶ A partir des menus de la fenêtre de l'éditeur de données, sélectionnez :
Transformer > Regroupement visuel...

Figure 9-1
Boîte de dialogue initiale Regroupement visuel



Dans la boîte de dialogue initiale Regroupement visuel, sélectionnez les variables d'échelle et/ou ordinales pour lesquelles vous souhaitez créer des variables regroupées. Le **regroupement** consiste à prendre plusieurs valeurs contiguës et à les regrouper dans une même modalité.

Etant donné que le regroupement visuel se base sur les valeurs réelles du fichier de données pour vous proposer les meilleurs choix de regroupement, il doit tout d'abord lire ce fichier. Si le fichier de données contient un nombre d'observations important, cette opération peut prendre un certain temps. Par conséquent, la boîte de dialogue initiale vous permet également de limiter le nombre d'observations à lire (« analyser »). Cette limitation est inutile avec notre fichier de données exemple. Il contient plus de 6 000 observations, mais l'analyse d'un tel nombre d'observations est rapide.

- Faites glisser *Revenu du foyer en milliers (revenu)* de la liste Variables vers la liste Variables à regrouper, puis cliquez sur Poursuivre.

Figure 9-2
Boîte de dialogue principale Regroupement visuel

Regroupement visuel

Liste des variables analysées :
 Revenu du foyer en milliers...

Nom : Variable actuelle : income Etiquette : Revenu du foyer en milliers (\$)
 Variable regroupée : inccat2 Etiquette : Revenu du foyer en milliers (\$) (Regroupé par casiers)
 Minimum : 9.00 Valeurs non manquantes Maximum : 1116.00

Entrez les divisions des intervalles ou cliquez sur Créer des divisions pour définir automatiquement des intervalles. La division 10, par exemple, définit un intervalle qui commence juste au-dessus de l'intervalle précédent et se termine à 10.

Grille :

	Valeur	Etiquette
1		ELEVÉE
2		

Observations analysées : 6400
 Valeurs manquantes : 0

Extrema supérie
 Inclus (<=)
 Exclus (<)

Créer des divisions...
 Créer des étiquettes
 Inverser l'échelle

OK Coller Réinitialiser Annuler Aide

- ▶ Dans la liste des variables analysées de la boîte de dialogue principale Regroupement visuel, sélectionnez *Revenu du foyer en milliers (revenu)*.
 Un histogramme affiche la distribution de la variable sélectionnée (très asymétrique dans le cas qui nous occupe).
- ▶ Affectez le nom *inccat2* à la nouvelle variable regroupée et sélectionnez l'étiquette de variable Modalité de revenu en milliers.
- ▶ Cliquez sur *Créer des divisions*.

Figure 9-3
Boîte de dialogue Divisions de regroupement visuel

Créer des divisions

Intervalles de longueur identique

Intervalles - Remplissez au moins deux champs

Emplacement de la première division : 25.00

Nombre de divisions : 3

Largeur : 25

Emplacement de la dernière division : 75.00

Centiles égaux fondés sur les observations analysées

Intervalles - Remplissez l'un des champs

Nombre de divisions :

Largeur (%) :

Divisions au niveau de la moyenne et des écarts-types sélectionnés, fondées sur les observations analysées

+/-1 écart-type

+/-2 écarts-types

+/-3 écarts-types

 L'option Appliquer remplace les définitions de division actuelles par cette spécification.
L'intervalle final inclut toutes les valeurs restantes : N divisions génèrent N+1 intervalles.

Appliquer Annuler Aide

- ▶ Sélectionnez Intervalles de longueur identique.
- ▶ Saisissez 25 pour l'emplacement de la première division, 3 pour le nombre de divisions et 25 pour la largeur.

Le nombre de modalités regroupées correspond au nombre de divisions, plus 1. Ainsi, dans cet exemple, la nouvelle variable regroupée comportera quatre modalités, les trois premières couvrant une plage de 25 (milliers) et la dernière toutes les valeurs supérieures à celle de la dernière division (75).

- ▶ Cliquez sur Appliquer.

Figure 9-4
Boîte de dialogue principale Regroupement visuel, dans laquelle des divisions ont été définies

Regroupement visuel

Liste des variables analysées :
 Revenu du foyer en milliers...

Nom :
 Variable actuelle : Etiquette :

Variable regroupée :

Minimum : Valeurs non manquantes Maximum :

Entrez les divisions des intervalles ou cliquez sur Créer des divisions pour définir automatiquement des intervalles. La division 10, par exemple, définit un intervalle qui commence juste au-dessus de l'intervalle précédent et se termine à 10.

Grille :

	Valeur	Etiquette
1	25.000	
2	50.000	
3	75.000	
4	ELEVEE	
5		

Observations analysées :

Valeurs manquantes :

Copier les casiers

Extrema supérieurs
 Inclus (<=)
 Exclus (<)

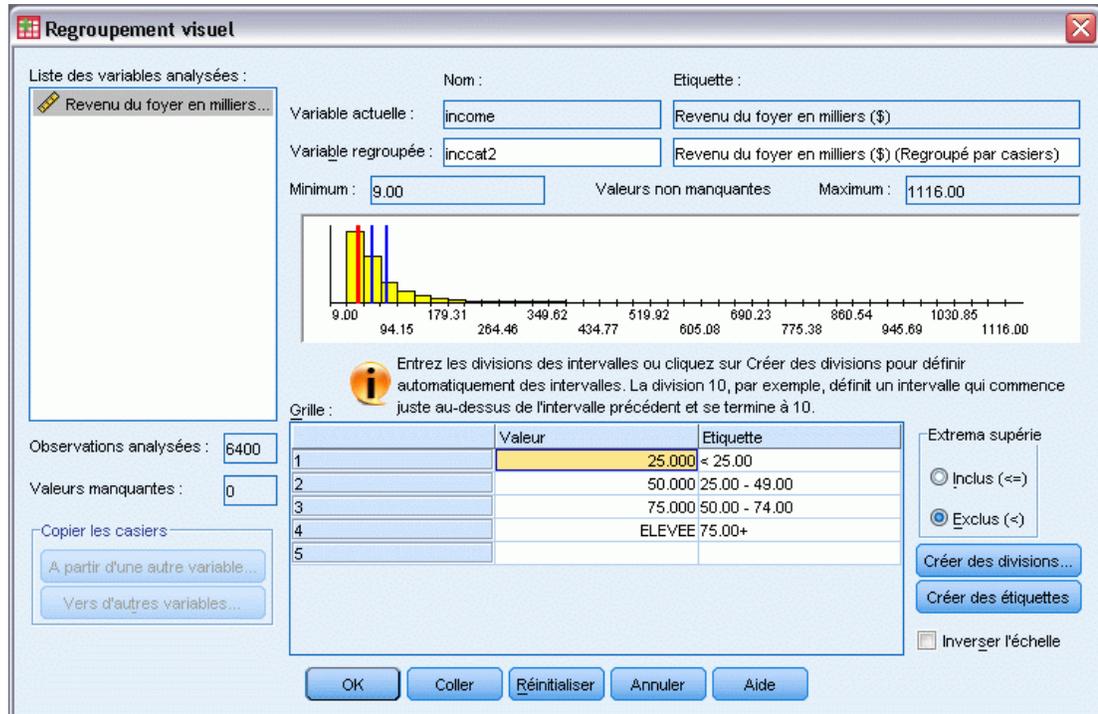
Inverser l'échelle

Les valeurs qui figurent à présent dans la grille représentent les divisions définies, à savoir les extrema supérieurs de chaque modalité. Par ailleurs, les lignes verticales de l'histogramme indiquent l'emplacement des divisions.

Par défaut, les valeurs de division sont incluses dans les modalités correspondantes. Par exemple, la première valeur (25) inclurait toutes les valeurs inférieures ou égales à 25. Dans cet exemple toutefois, nous voulons définir les modalités suivantes : inférieur à 25, 25–49, 50–74, et 75 ou plus.

- ▶ Dans le groupe Extrema supérieurs, sélectionnez Exclue (<).
- ▶ Cliquez ensuite sur Créer des étiquettes.

Figure 9-5
Étiquettes de valeur générées automatiquement



Cette opération permet de générer automatiquement des étiquettes de valeur descriptives pour chaque modalité. Étant donné que les valeurs réelles affectées à la nouvelle variable regroupée sont de simples entiers séquentiels (le premier étant 1), les étiquettes de valeur peuvent s'avérer très utiles.

Vous pouvez également procéder manuellement pour entrer ou modifier des divisions et des étiquettes dans la grille, pour modifier des emplacements de division en faisant glisser les lignes de division de l'histogramme, et pour supprimer des divisions en faisant glisser les lignes de division hors de l'histogramme.

- Cliquez sur OK pour créer la variable regroupée.

La nouvelle variable est affichée dans l'éditeur de données. Etant donné que la variable est ajoutée à la fin du fichier, elle apparaît dans la colonne la plus à droite de l'affichage Données et dans la dernière ligne de l'affichage Variable.

Figure 9-6
Nouvelle variable affichée dans l'éditeur de données

	ownpc	ownfax	news	response	incat2	var
1	Non	Non	Non	Non	50,00 - 74,00	
2	Non	Non	Non	Oui	75,00+	
3	Oui	Non	Oui	Non	25,00 - 49,00	
4	Oui	Oui	Oui	Non	25,00 - 49,00	
5	Non	Non	Oui	Non	< 25,00	
6	Oui	Non	Non	Non	75,00+	
7	Non	Non	Non	Non	25,00 - 49,00	
8	Oui	Non	Non	Non	50,00 - 74,00	
9	Non	Non	Oui	Non	< 25,00	
10	Non	Oui	Oui	Oui	75,00+	
11	Oui	Non	Oui	Non	50,00 - 74,00	
12	Oui	Oui	Oui	Non	< 25,00	
13	Non	Non	Oui	Non	25,00 - 49,00	
14	Oui	Non	Non	Oui	75,00+	

Calcul de nouvelles variables

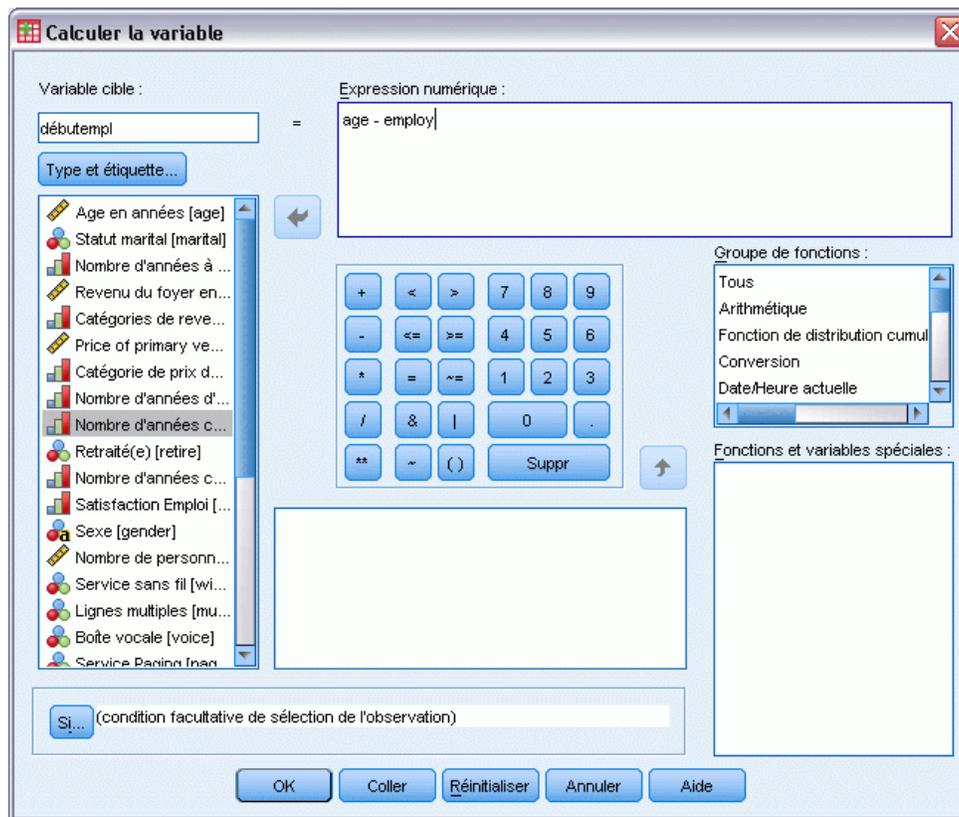
A l'aide d'une grande variété de fonctions mathématiques, vous pouvez calculer de nouvelles variables en fonction d'équations extrêmement complexes. Dans cet exemple, cependant, nous calculerons simplement une nouvelle variable qui représente la différence entre les valeurs de deux variables existantes.

Le fichier de données *demo.sav* contient une variable pour l'âge actuel du répondant et une variable pour le nombre d'années d'occupation du poste actuel. En revanche, il ne contient pas de variable indiquant l'âge du répondant au moment où il a été embauché à ce poste. Nous pouvons créer une nouvelle variable qui représenterait la différence calculée entre l'âge actuel et le nombre d'années d'occupation du poste actuel afin d'obtenir l'âge approximatif du répondant au moment de l'embauche.

- ▶ A partir des menus de la fenêtre de l'éditeur de données, sélectionnez :
Transformer > Calculer la variable...
- ▶ Pour Variable destination, saisissez débutempl.

- ▶ Sélectionnez la variable *Age en années [âge]* dans la liste des variables source et cliquez sur le bouton fléché pour la copier dans la zone Expression numérique.
- ▶ Cliquez sur le bouton moins (-) du clavier du pavé numérique de la boîte de dialogue (ou appuyez sur la touche moins du clavier).
- ▶ Sélectionnez *Nb d'années avec l'employeur actuel [emploi]* et cliquez sur le bouton fléché pour la copier dans l'expression.

Figure 9-7
Boîte de dialogue Calculer la variable



Remarque : Veillez à sélectionner la variable d'emploi appropriée. Il existe également une version qualitative recodée de la variable, mais ce n'est *pas* ce que vous recherchez. L'expression numérique doit être *âge-emploi* et non *âge-empl_dis*.

- ▶ Cliquez sur OK pour calculer la nouvelle variable.

La nouvelle variable est affichée dans l'éditeur de données. Etant donné que la variable est ajoutée à la fin du fichier, elle apparaît dans la colonne la plus à droite de l'affichage Données et dans la dernière ligne de l'affichage Variable.

Figure 9-8

Nouvelle variable affichée dans l'éditeur de données

	ownfax	news	response	inccat2	jobstart
1	Non	Non	Non	\$50 - \$74	52,00
2	Non	Non	Oui	\$75 - \$124	53,00
3	Non	Oui	Non	\$25 - \$49	27,00
4	Oui	Oui	Non	\$25 - \$49	23,00
5	Non	Oui	Non	Inf à \$25	23,00
6	Non	Non	Non	\$75 - \$124	43,00
7	Non	Non	Non	\$25 - \$49	40,00
8	Non	Non	Non	\$50 - \$74	34,00
9	Non	Oui	Non	Inf à \$25	44,00
10	Oui	Oui	Oui	\$75 - \$124	32,00
11	Non	Oui	Non	\$50 - \$74	54,00
12	Oui	Oui	Non	Inf à \$25	27,00
13	Non	Oui	Non	\$25 - \$49	30,00
14	Non	Non	Oui	\$75 - \$124	41,00

Utilisation de fonctions dans les expressions

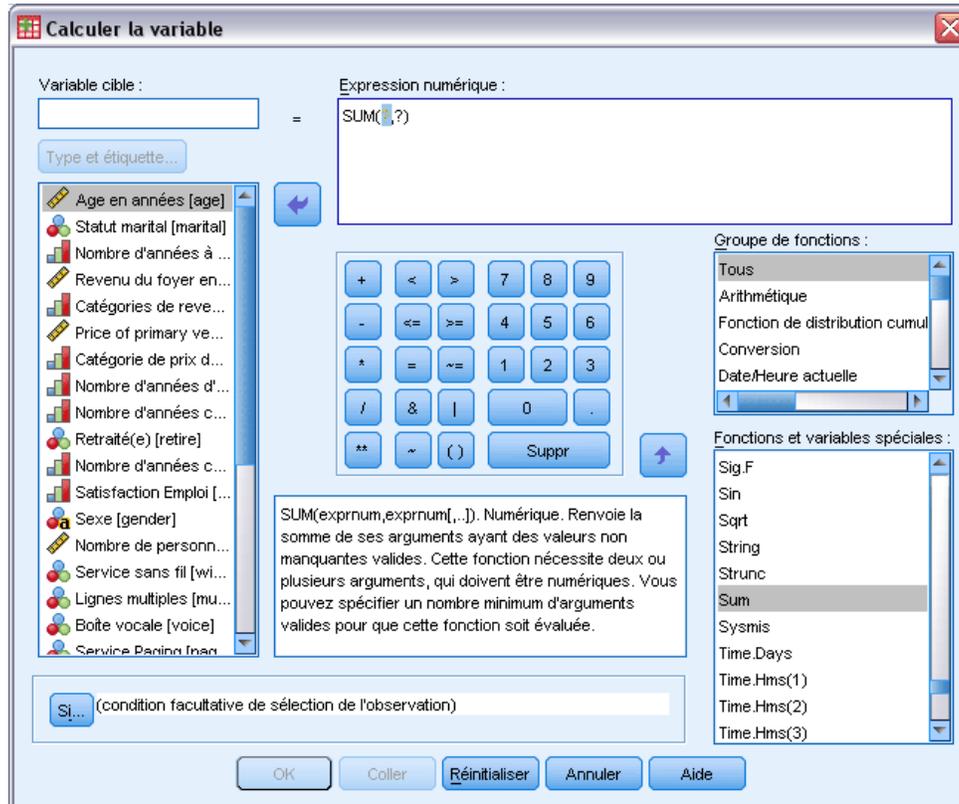
Vous pouvez également utiliser des fonctions prédéfinies dans les expressions. Vous disposez de plus de 70 fonctions intégrées, dont les suivantes :

- Fonctions arithmétiques
- Fonctions statistiques
- Fonctions de distribution
- Fonctions logiques
- Fonctions d'agrégation et d'extraction relatives à la date et à l'heure
- Fonctions de valeur manquante

- Fonctions d'observations croisées
- Fonctions sur chaînes

Figure 9-9

Boîte de dialogue Calculer la variable indiquant le regroupement des fonctions



Les fonctions sont organisées en groupes logiques, tels qu'un groupe pour les opérations arithmétiques et un autre pour le calcul des métriques statistiques. Dans un souci de commodité, une série de variables système couramment utilisées, telles que la variable *\$TIME* (date et heure actuelles), figurent également dans des groupes de fonctions appropriés. Une brève description de la variable système ou de la fonction actuellement sélectionnée (en l'occurrence, la fonction *SUM*) apparaît dans une zone particulière de la boîte de dialogue Calculer la variable.

Collage d'une fonction dans une expression

Pour coller une fonction dans une expression :

- ▶ Dans l'expression, positionnez le curseur là où vous souhaitez que la fonction apparaisse.
- ▶ Sélectionnez le groupe approprié dans la liste Groupe de fonctions. Le groupe étiqueté Tous répertorie toutes les fonctions et variables système disponibles.
- ▶ Double-cliquez sur la fonction dans la liste Fonctions et variables spéciales (ou sélectionnez-la, puis cliquez sur la flèche adjacente à la liste Groupe de fonctions).

La fonction est insérée dans l'expression. Si vous sélectionnez une partie de l'expression, puis insérez la fonction, la partie sélectionnée de l'expression est utilisée comme premier argument dans la fonction.

Modification d'une fonction dans une expression

La fonction collée n'est complète que lorsque vous avez saisi les arguments, symbolisés par des points d'interrogation. Le nombre de points d'interrogation indique le nombre minimum d'arguments requis pour compléter la fonction.

- ▶ Sélectionnez les points d'interrogation dans la fonction collée.
- ▶ Saisissez les arguments. Si les arguments sont des noms de variable, vous pouvez les coller à partir de la liste des variables.

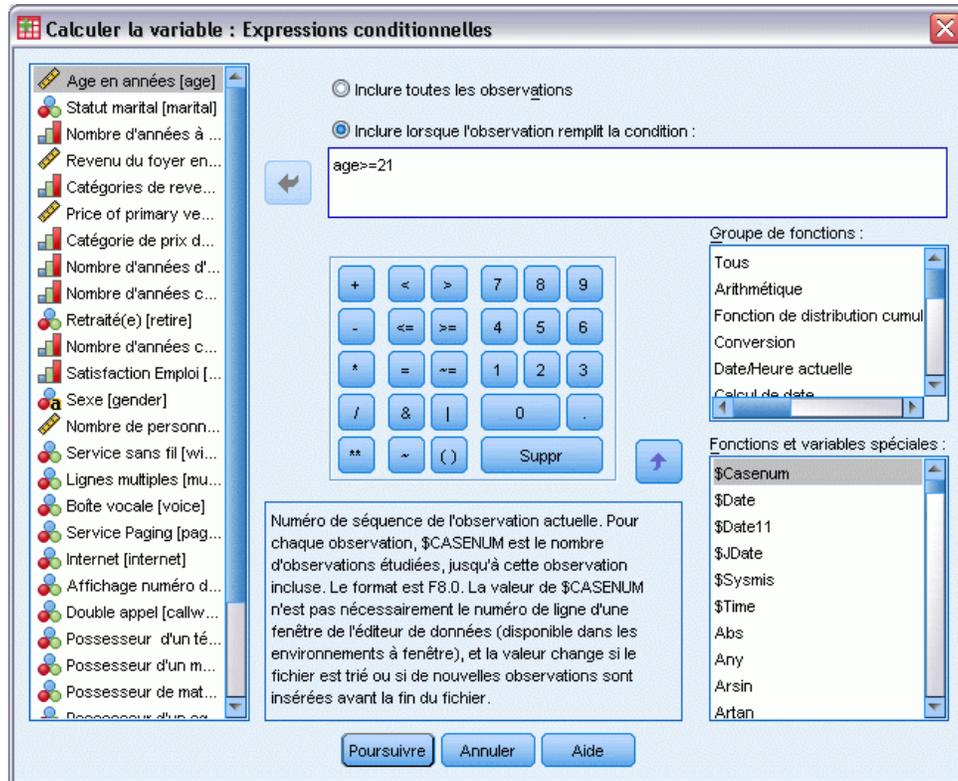
Utilisation d'expressions conditionnelles

Vous pouvez utiliser des expressions conditionnelles (également appelées expressions logiques) pour appliquer des transformations aux sous-groupes d'observations sélectionnés. Une expression conditionnelle renvoie la valeur vrai, faux ou manquant pour chaque observation. Si le résultat d'une expression conditionnelle est True (vrai), la transformation est appliquée à cette observation. Si le résultat est False (faux) ou Manquant, la transformation n'est pas appliquée à l'observation.

Pour indiquer une expression conditionnelle :

- ▶ Cliquez sur Si dans la boîte de dialogue Calculer la variable. Cette opération ouvre la boîte de dialogue Si les observations.

Figure 9-10
Boîte de dialogue Calculer la variable : si les observations



- ▶ Sélectionnez Inclure lorsque l'observation remplit la condition :
- ▶ Entrez l'expression conditionnelle.

La plupart des expressions conditionnelles contiennent au moins un opérateur relationnel, comme dans :

`âge>=21`

ou

`revenu*3<100`

Dans le premier exemple, seules les observations dont la valeur *Age [âge]* est supérieure ou égale à 21 sont sélectionnées. Dans le deuxième exemple, la valeur *Revenu du foyer en milliers [revenu]* multipliée par 3 doit être inférieure à 100 pour qu'une observation soit sélectionnée.

Vous pouvez aussi lier au moins deux expressions conditionnelles à l'aide d'opérateurs logiques, comme dans :

`âge>=21 | ed>=4`

ou

`revenu*3<100 & ed=5`

Dans le premier exemple, les observations qui remplissent la condition *Age [age]* ou *Nombre d'années d'éducation [ed]* sont sélectionnées. Dans le deuxième exemple, les observations doivent remplir les conditions *Revenu du foyer en milliers [revenu]* et *Nombre d'années d'éducation [ed]* pour être sélectionnées.

Utilisation des dates et heures

L'Assistant Date et heure vous permet de réaliser de nombreuses tâches courantes portant sur les dates et les heures. À l'aide de cet assistant, vous pouvez effectuer les opérations suivantes :

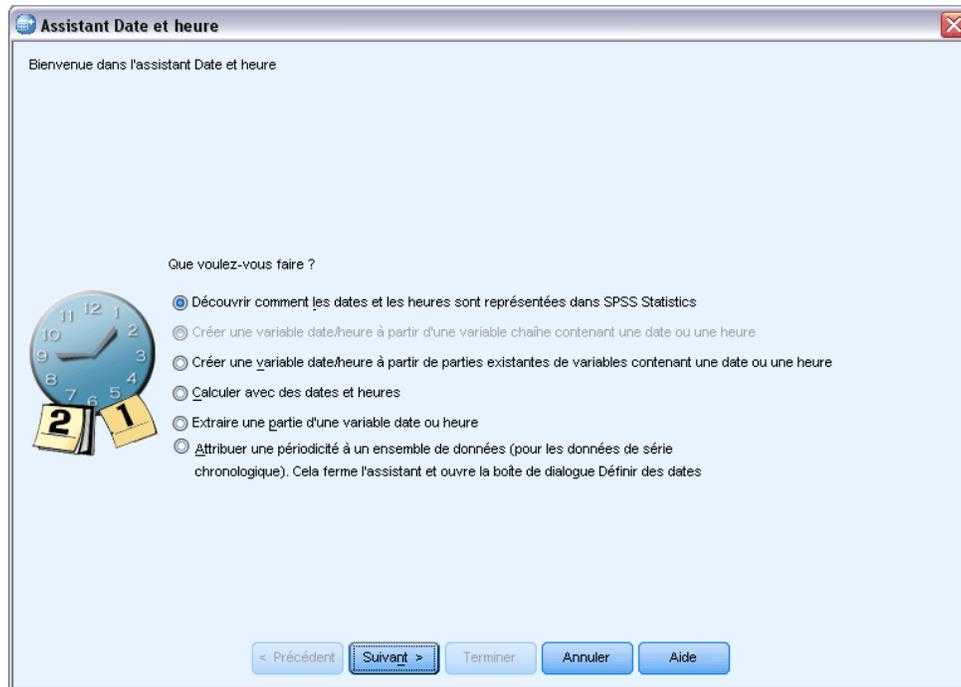
- Créer une variable date/heure à partir d'une variable caractères contenant une date ou une heure.
- Construire une variable date/heure en fusionnant des variables contenant différentes parties de la date ou de l'heure.
- Ajouter ou soustraire des valeurs aux variables date/heure, notamment ajouter ou soustraire deux variables date/heure.
- Extraire une partie d'une variable date ou heure, telle que le jour du mois d'une variable date/heure de format mm/jj/aaaa.

Les exemples de cette section utilisent le fichier de données *upgrade.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.](#)

Pour utiliser l'Assistant Date et heure :

- ▶ À partir des menus, sélectionnez :
Transformer > Assistant Date et heure...

Figure 9-11
Ecran de présentation de l'Assistant Date et heure



L'écran de présentation de l'Assistant Date et heure met à votre disposition une série de tâches générales. Les tâches inapplicables aux données utilisées sont désactivées. Par exemple, le fichier de données *upgrade.sav* ne contient aucune variable chaîne. Par conséquent, la tâche permettant de créer une variable date à partir d'une chaîne est désactivée.

Si vous ne savez pas comment manipuler les dates et les heures dans IBM® SPSS® Statistics, vous pouvez sélectionner l'option Connaître la représentation des dates et heures, puis cliquer sur le bouton Suivant. Cette opération fait apparaître un écran contenant une brève présentation des variables date/heure et un lien, par le biais du bouton Aide, vers des informations complémentaires.

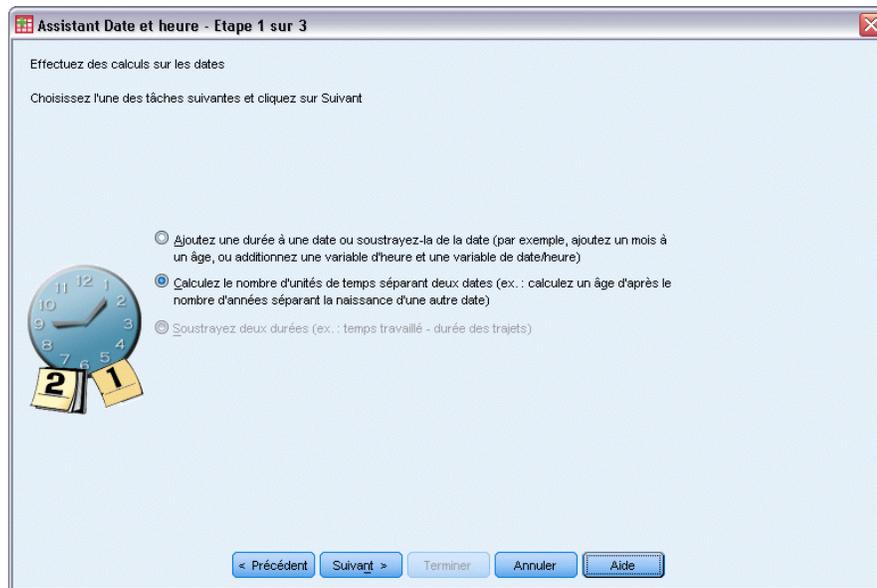
Calcul du temps écoulé entre deux dates

L'une des tâches les plus courantes portant sur les dates est le calcul du temps s'écoulant entre deux dates. A titre d'exemple, imaginez un éditeur de logiciels qui souhaite analyser les achats de licence de mise à niveau en déterminant le nombre d'années écoulées depuis le dernier achat d'une mise à niveau par chaque client. Le fichier de données *upgrade.sav* contient une variable pour la date du dernier achat d'une mise à niveau par client, mais pas pour le nombre d'années écoulées depuis cet achat. Une nouvelle variable représentant le temps écoulé, en années, entre la date de la dernière mise à niveau et la date de la version suivante du produit fournira une mesure de cette quantité.

Pour calculer le temps écoulé entre deux dates :

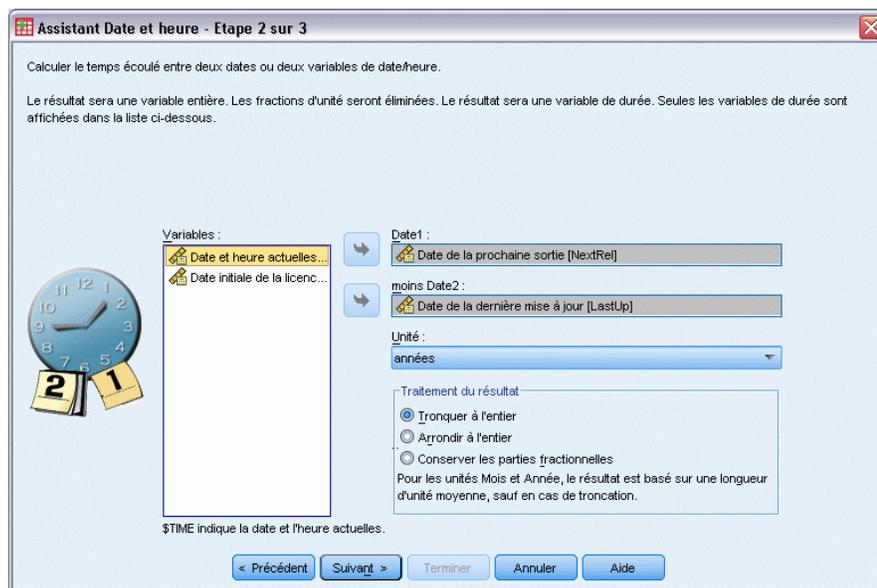
- ▶ Sélectionnez l'option Calculer avec des dates et heures dans l'écran de présentation de l'Assistant Date et heure, puis cliquez sur le bouton Suivant.

Figure 9-12
Calcul du temps écoulé entre deux dates : Etape 1



- Sélectionnez l'option *Calculez le nombre d'unités de temps séparant deux dates*, puis cliquez sur le bouton *Suivant*.

Figure 9-13
Calcul du temps écoulé entre deux dates : Etape 2



- Sélectionnez l'option *Date de la prochaine version* pour la variable *Date1*.
- Sélectionnez l'option *Date de la dernière mise à niveau* pour la variable *Date2*.

- ▶ Sélectionnez années pour l'unité et Tronquer à l'entier pour le traitement du résultat. (Ce sont des sélections par défaut.)
- ▶ Cliquez sur Suivant.

Figure 9-14

Calcul du temps écoulé entre deux dates : Etape 3

Assistant Date et heure - Etape 3 sur 3

Calcul : NextRel - LastUp

Variable de résultat : NbreAnsDepMAN Unités : années

Etiquette de variable : Nombre d'années écoulées depuis la dernière mise à niveau

Exécution

Créer la variable maintenant Coller la syntaxe dans la fenêtre de syntaxe

< Précédent Suivant > Terminer Annuler Aide

- ▶ Entrez *NbreAnsDepMAN* comme nom de la variable de résultat. Les variables de résultat ne peuvent pas porter le même nom qu'une variable existante.
- ▶ Entrez *Nombre d'années écoulées depuis la dernière mise à niveau* comme étiquette de la variable de résultat. Les étiquettes des variables de résultat sont facultatives.
- ▶ Conservez la sélection par défaut de l'option *Créer la variable maintenant*, puis cliquez sur le bouton *Terminer* pour créer la nouvelle variable.

La nouvelle variable *NbreAnsDepMAN*, affichée dans l'éditeur de données, est le nombre entier d'années entre les deux dates. Les parties fractionnelles d'une année ont été tronquées.

Figure 9-15
Nouvelle variable affichée dans l'éditeur de données

	PurDate	Support	LastUp	NextRel	NbreAnsDepMAN
1	12/30/1998	4	02/28/2002	06/01/04	2
2	06/28/2001	2	09/28/2002	06/01/04	1
3	08/27/1999	2	09/27/2001	06/01/04	2
4	02/22/2000	4	01/22/2003	06/01/04	1
5	01/26/2000	2	08/26/2001	06/01/04	2
6	07/10/1999	3	07/10/2003	06/01/04	0
7	01/24/2003	2	07/24/2003	06/01/04	0
8	06/15/1999	2	09/15/2003	06/01/04	0
9	01/18/2003	5	07/18/2003	06/01/04	0
10	12/02/2002	4	06/02/2003	06/01/04	0
11	08/10/2000	1	10/10/2002	06/01/04	1
12	05/27/1999	2	07/27/2000	06/01/04	3
13	02/28/1999	4	10/28/2002	06/01/04	1
14	01/02/2001	5	07/02/2001	06/01/04	2

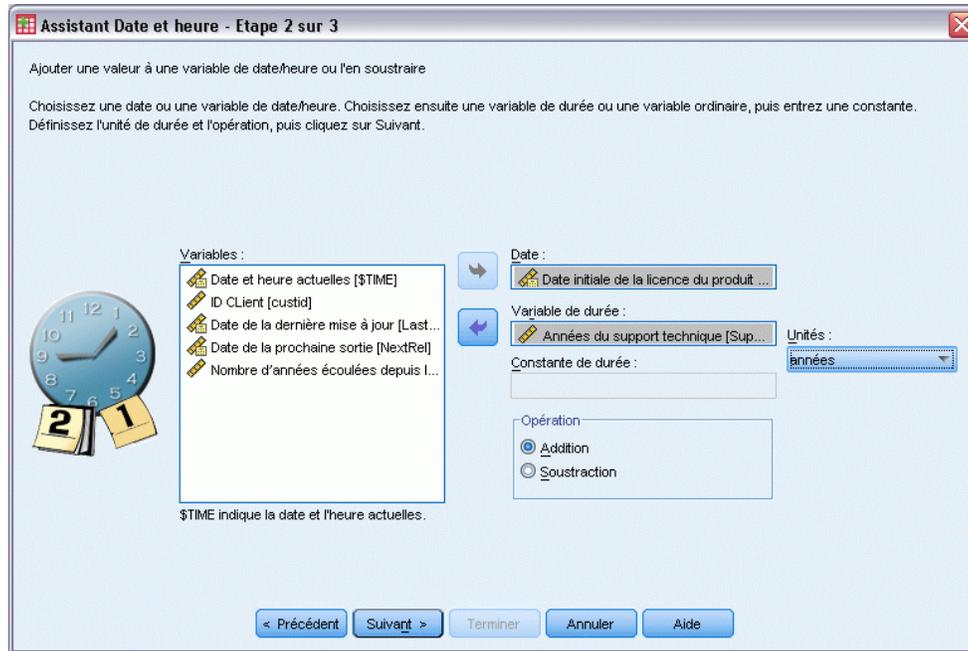
Ajout d'une durée à une date

Vous pouvez ajouter ou soustraire des durées, telles que 10 jours ou 12 mois, à une date. A partir de l'exemple de l'éditeur de logiciels utilisé dans la section précédente, vous pouvez envisager de déterminer la date à laquelle expire le contrat d'assistance technique d'origine de chaque client. Le fichier de données *upgrade.sav* contient une variable pour le nombre d'années d'assistance sous contrat et une variable pour la date d'achat initiale. Vous pouvez donc déterminer la date d'expiration de l'assistance initiale en ajoutant le nombre d'années d'assistance à la date achat.

Pour ajouter une durée à une date :

- ▶ Sélectionnez l'option Calculer avec des dates et heures dans l'écran de présentation de l'Assistant Date et heure, puis cliquez sur le bouton Suivant.
- ▶ Sélectionnez l'option Ajout d'une durée à une date ou Suppression d'une durée d'une date, puis cliquez sur le bouton Suivant.

Figure 9-16
Ajout d'une durée à une date : Etape 2



- ▶ Sélectionnez l'option *Date de la première licence produit* pour la variable Date.
- ▶ Sélectionnez *Années d'assistance technique* pour la variable de la durée.

Dans la mesure où la variable *Années d'assistance technique* est uniquement une variable numérique, vous devez indiquer l'unité dans laquelle cette variable doit être ajoutée en tant que durée.

- ▶ Sélectionnez *Années* dans la liste déroulante Unités.
- ▶ Cliquez sur *Suivant*.

Figure 9-17
Ajout d'une durée à une date : Etape 3

Assistant Date et heure - Etape 3 sur 3

Entrez un nom (et, éventuellement, une étiquette) pour la variable de résultat. Le format de la variable d'entrée de date/heure sera appliqué à la variable de résultat.

Calcul :
PurDate + Support (années)

Variable de résultat :
DateExpAss

Etiquette de variable :
Date d'expiration de l'assistance

Exécution
 Créer la variable maintenant Coller la syntaxe dans la fenêtre de syntaxe

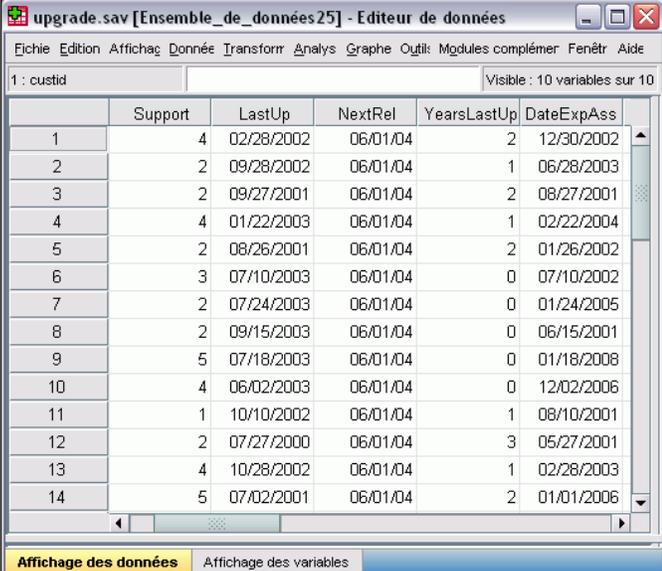
< Précédent Suivant > Terminer Annuler Aide

- ▶ Entrez *DateExpAss* comme nom de la variable résultat. Les variables de résultat ne peuvent pas porter le même nom qu'une variable existante.
- ▶ Entrez *Date d'expiration de l'assistance* comme étiquette de la variable résultat. Les étiquettes des variables de résultat sont facultatives.
- ▶ Cliquez sur le bouton Terminer pour créer la nouvelle variable.

La nouvelle variable est affichée dans l'éditeur de données.

Figure 9-18

Nouvelle variable affichée dans l'éditeur de données



upgrade.sav [Ensemble_de_données25] - Editeur de données

Fichier Edition Affichage Donnée Transferr Analys Graphe Outil Modules complément Fenêtr Aide

1 : custid Visible : 10 variables sur 10

	Support	LastUp	NextRel	YearsLastUp	DateExpAss
1	4	02/28/2002	06/01/04	2	12/30/2002
2	2	09/28/2002	06/01/04	1	06/28/2003
3	2	09/27/2001	06/01/04	2	08/27/2001
4	4	01/22/2003	06/01/04	1	02/22/2004
5	2	08/26/2001	06/01/04	2	01/26/2002
6	3	07/10/2003	06/01/04	0	07/10/2002
7	2	07/24/2003	06/01/04	0	01/24/2005
8	2	09/15/2003	06/01/04	0	06/15/2001
9	5	07/18/2003	06/01/04	0	01/18/2008
10	4	06/02/2003	06/01/04	0	12/02/2006
11	1	10/10/2002	06/01/04	1	08/10/2001
12	2	07/27/2000	06/01/04	3	05/27/2001
13	4	10/28/2002	06/01/04	1	02/28/2003
14	5	07/02/2001	06/01/04	2	01/01/2006

Affichage des données Affichage des variables

Tri et sélection de données

L'organisation des fichiers de données ne répond pas toujours idéalement à vos besoins. Pour préparer des données pour une analyse, vous avez le choix entre un grand nombre de transformations de fichiers. Vous pouvez effectuer les opérations suivantes :

- **Trier les données** : Vous pouvez trier les observations en fonction de la valeur de certaines variables.
- **Sélectionner des sous-ensembles d'observations** : Vous pouvez limiter votre analyse à un sous-ensemble d'observations ou effectuer des analyses simultanées sur différents sous-ensembles.

Les exemples de ce chapitre utilisent le fichier de données *demo.sav*. [Pour plus d'informations, reportez-vous à la section Fichiers d'exemple dans l'annexe A sur p. 156.](#)

Tri des données

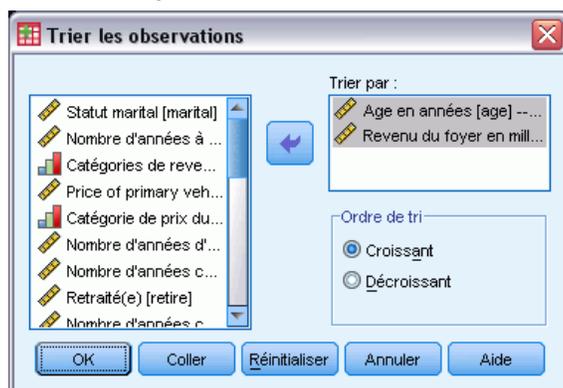
Il est souvent utile (et parfois nécessaire) de trier les observations (tri des lignes du fichier de données) pour certains types d'analyse.

Pour réorganiser la séquence des observations dans le fichier de données en fonction de la valeur de certaines variables de tri, procédez comme suit :

- A partir des menus, sélectionnez :
Données > Trier les observations...

La boîte de dialogue Trier les observations est affichée.

Figure 10-1
Boîte de dialogue Trier les observations



- Ajoutez les variables *Age en années [âge]* et *Revenu du ménage en milliers [revenu]* à la liste Trier par.

Si vous sélectionnez plusieurs variables de tri, l'ordre dans lequel elles apparaissent dans la liste Trier par détermine celui dans lequel les observations sont triées. Dans cet exemple, basé sur les entrées de la liste Trier par, les observations seront triées en fonction de la valeur de la variable *Revenu du ménage en milliers [revenu]* dans les modalités de la variable *Age en années [âge]*. Pour les variables chaîne, les majuscules précèdent leurs équivalents minuscules dans l'ordre de tri (par exemple, la variable chaîne *Oui* précède la variable chaîne *oui* dans l'ordre de tri).

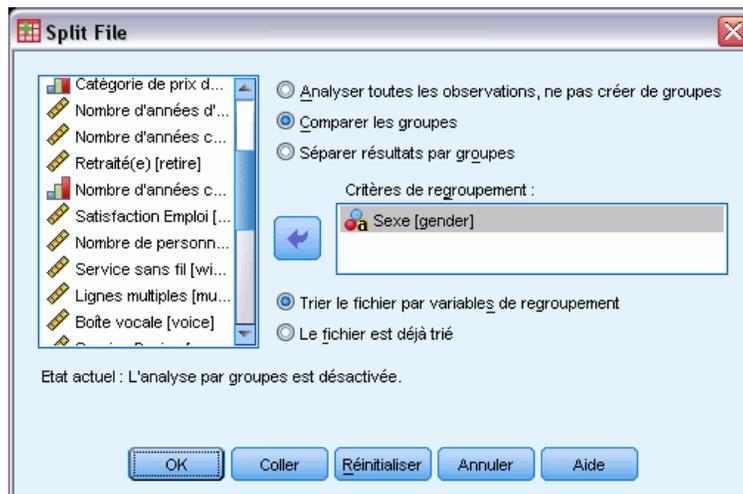
Traitement d'un fichier scindé

Pour scinder le fichier de données en groupes distincts pour l'analyse :

- A partir des menus, sélectionnez :
Données > Scinder un fichier

La boîte de dialogue Scinder un fichier est affichée.

Figure 10-2
Boîte de dialogue Scinder un fichier



- Sélectionnez *Comparer les groupes* ou *Séparer résultats par groupes*. (Les exemples qui suivent ces étapes montrent les différences entre les deux options.)
- Sélectionnez *Sexe [sexe]* afin de scinder le fichier en groupes distincts pour ces variables.

Vous pouvez utiliser des variables numériques, des variables chaîne courte et des variables chaîne longue comme variables de regroupement. Une analyse distincte est effectuée pour chaque sous-groupe défini par les variables de regroupement. Si vous sélectionnez plusieurs variables de regroupement, l'ordre dans lequel elles apparaissent dans la liste Critères de regroupement détermine la façon dont les observations sont regroupées.

Si vous sélectionnez Comparer les groupes, les résultats de tous les groupes du fichier scindé seront inclus dans le(s) même(s) tableau(x), comme dans le tableau des statistiques récapitulatives suivant généré par la procédure Fréquences.

Figure 10-3
Sortie Scinder un fichier contenant un seul tableau pivotant

Statistiques			
Revenu du ménage en milliers (\$)			
Femme	N	Valide	3179
		Manquante	0
		Moyenne	68,7798
		Médiane	44,0000
		Ecart-type	75,73510
Homme	N	Valide	3221
		Manquante	0
		Moyenne	70,1608
		Médiane	45,0000
		Ecart-type	81,56216

Si vous sélectionnez Séparer résultats par groupes et que vous exécutez la procédure Fréquences, deux tableaux pivotants sont créés : un tableau pour les femmes et un pour les hommes.

Figure 10-4
Sortie Scinder un fichier contenant un tableau pivotant pour les femmes

Statistiques ^a		
Revenu du ménage en milliers (\$)		
N	Valide	3179
	Manquante	0
	Moyenne	68,7798
	Médiane	44,0000
	Ecart-type	75,73510

a. Sexe = Femme

Figure 10-5
Sortie Scinder un fichier contenant un tableau pivotant pour les personnes de sexe masculin

Statistiques ^a		
Revenu du ménage en milliers (\$)		
N	Valide	3221
	Manquante	0
	Moyenne	70,1608
	Médiane	45,0000
	Ecart-type	81,56216

a. Sexe = Homme

Tri d'observations pour le traitement d'un fichier scindé

La procédure Scinder un fichier crée un nouveau sous-groupe chaque fois qu'elle rencontre une valeur différente pour l'une des variables de regroupement. Par conséquent, il est important de trier les observations en fonction des valeurs des variables de regroupement avant de lancer le traitement du fichier scindé.

Par défaut, la procédure Scinder un fichier trie automatiquement le fichier de données en fonction des valeurs des variables de regroupement. Si le fichier est déjà trié dans l'ordre correct, vous pouvez gagner du temps lors du traitement en sélectionnant Le fichier est déjà trié.

Activation et désactivation du traitement d'un fichier scindé

Une fois le traitement d'un fichier scindé lancé, il reste actif pour le reste de la session, à moins que vous ne le désactiviez.

- **Analyser toutes les observations.** Cette option désactive le traitement du fichier scindé.
- **Comparer les groupes et Séparer résultats par groupes.** Cette option active le traitement du fichier scindé.

Si le traitement du fichier scindé est activé, le message Séparer fichier actif apparaît sur la barre d'état située au bas de la fenêtre de l'application.

Sélection de sous-groupes d'observations

Vous pouvez limiter votre analyse à un sous-groupe donné en fonction de critères contenant des variables et des expressions complexes. Vous pouvez également sélectionner un échantillon aléatoire d'observations. Les critères utilisés pour définir un sous-groupe comprennent :

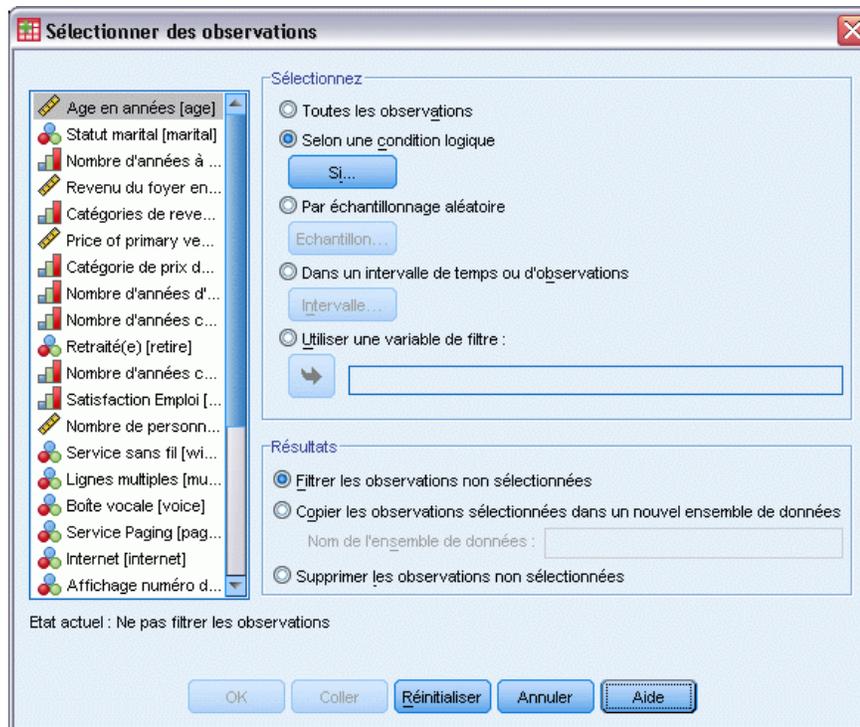
- Plages et valeurs de variables
- Plages de dates et d'heures
- Nombres d'observations (lignes)
- Expressions arithmétiques
- Expressions logiques
- Fonctions

Pour sélectionner un sous-ensemble d'observations à analyser, procédez comme suit :

- ▶ A partir des menus, sélectionnez :
Données > Sélectionner des observations

Cette opération ouvre la boîte de dialogue Sélectionner des observations.

Figure 10-6
Boîte de dialogue Sélectionner des observations



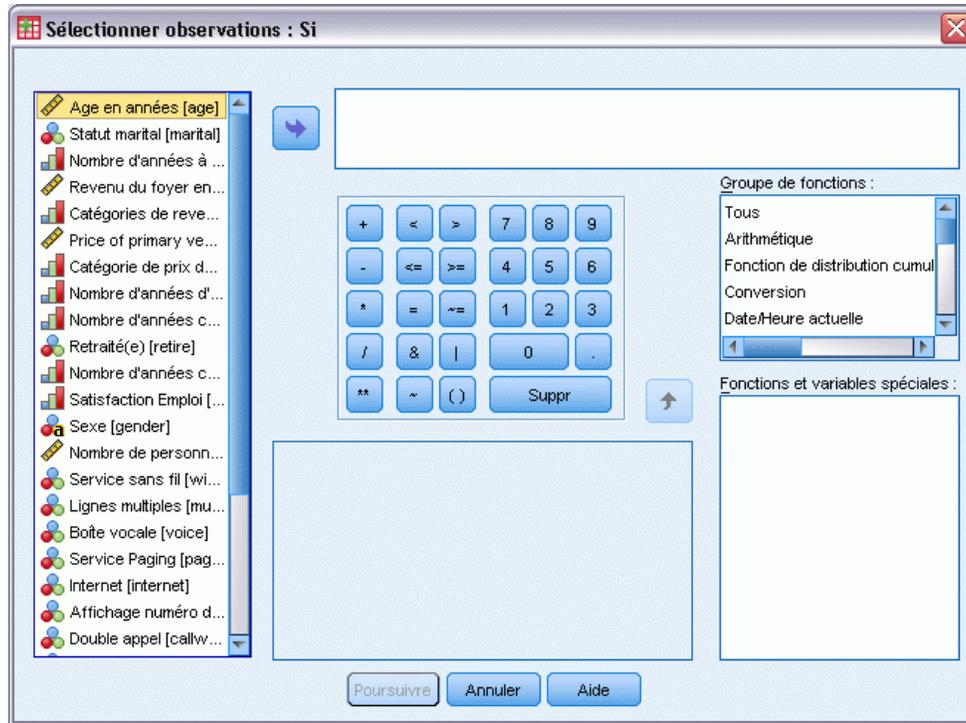
Sélectionner des sous-ensembles d'observations à l'aide d'une expression conditionnelle

Pour sélectionner des observations sur la base d'une expression conditionnelle, procédez comme suit :

- Sélectionnez Selon une condition logique, puis cliquez sur Si dans la boîte de dialogue Sélectionner des observations.

Cette opération ouvre la boîte de dialogue Sélectionner des observations - Si.

Figure 10-7
Boîte de dialogue Sélectionner des observations : Si



L'expression conditionnelle peut utiliser des noms de variables existantes, des constantes, des opérateurs arithmétiques, des opérateurs logiques, des opérateurs relationnels et des fonctions. Vous pouvez saisir et modifier l'expression dans la zone de texte comme n'importe quel texte dans une fenêtre de résultats. Vous pouvez également utiliser le pavé numérique de la boîte de dialogue, la liste des variables et la liste des fonctions pour coller des éléments dans l'expression. [Pour plus d'informations, reportez-vous à la section Utilisation d'expressions conditionnelles dans le chapitre 9 sur p. 137.](#)

Sélectionner un échantillon aléatoire

Pour obtenir un échantillon aléatoire, procédez comme suit :

- ▶ Sélectionnez Par échantillonnage aléatoire dans la boîte de dialogue Sélectionner des observations.
- ▶ Cliquez sur Echantillon.

Cette opération ouvre la boîte de dialogue Sélectionner observations : Echantillon aléatoire.

Figure 10-8
Boîte de dialogue Sélectionner observations : Echantillon aléatoire



Pour la taille de l'échantillon, deux options sont disponibles :

- **Environ** : Pourcentage défini par l'utilisateur. Cette option génère un échantillon aléatoire d'observations dont le nombre correspond approximativement au pourcentage indiqué.
- **Exactement** : Nombre d'observations spécifié par l'utilisateur. Vous devez également indiquer le nombre d'observations à partir duquel l'échantillon sera généré. Ce deuxième nombre doit être inférieur ou égal au nombre total d'observations dans le fichier de données. Si ce nombre dépasse le nombre total d'observations dans le fichier de données, l'échantillon contiendra proportionnellement moins d'observations que le nombre demandé.

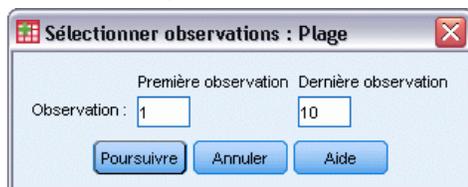
Sélection d'un intervalle de temps ou d'observations

Pour sélectionner un intervalle d'observations sur la base de dates, d'heures ou de numéros (lignes) d'observations, procédez comme suit :

- Sélectionnez Dans un intervalle de temps ou d'observations, puis cliquez sur Plage dans la boîte de dialogue Sélectionner des observations.

Cette opération ouvre la boîte de dialogue Sélectionner des observations : Intervalle, dans laquelle vous pouvez sélectionner un intervalle de numéros (lignes) d'observations.

Figure 10-9
Boîtes de dialogue Sélectionner des observations : Intervalle



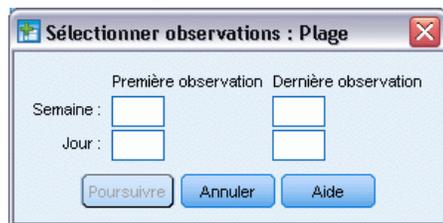
- **Première**. Entrez la date de début et/ou les valeurs de temps de l'intervalle. Si aucune variable de date n'est définie, entrez le numéro d'observation de départ (numéro de ligne dans l'éditeur de données, sauf si l'option Scinder un fichier est activée). Si vous n'indiquez aucune valeur pour la zone Dernière, toutes les observations à partir de la date/l'heure de début jusqu'à la fin de la série chronologique sont sélectionnées.
- **Dernière**. Entrez la date de fin et/ou les valeurs de temps pour l'intervalle. Si aucune variable de date n'est définie, entrez le numéro d'observation de fin (numéro de ligne dans l'éditeur de données, sauf si l'option Scinder un fichier est activée). Si vous n'indiquez aucune valeur

pour la zone Première, toutes les observations à partir du début de la série chronologique jusqu'à la date/l'heure de fin sont sélectionnées.

Pour les données de séries chronologiques comportant des variables de date définies, vous pouvez sélectionner un intervalle de dates et/ou d'heures sur la base des variables de date définies. Chaque observation représente des constatations effectuées à différents moments et le fichier est trié dans l'ordre chronologique.

Figure 10-10

Boîte de dialogue Sélectionner des observations : Intervalle (séries chronologiques)



Pour générer des variables de date pour des données de séries chronologiques, procédez comme suit :

- A partir des menus, sélectionnez :
Données > Définir des dates

Traitement des observations exclues

Vous pouvez choisir l'une des options suivantes pour traiter les observations exclues :

- **Filtrez les observations exclues.** Les observations exclues ne sont pas incluses dans l'analyse, mais restent dans l'ensemble de données. Vous pouvez utiliser les observations exclues ultérieurement dans la session si vous désactivez le filtrage. Si vous sélectionnez un échantillon aléatoire ou si vous sélectionnez des observations sur la base d'une expression conditionnelle, une variable nommée *filter_\$* est générée ; elle comporte la valeur 1 pour les observations sélectionnées et la valeur 0 pour les observations exclues.
- **Copiez les observations sélectionnées dans l'ensemble de données.** Les observations sélectionnées sont copiées dans un nouvel ensemble de données ; l'ensemble de données d'origine reste inchangé. Les observations exclues ne sont pas incluses dans le nouvel ensemble de données et sont conservées dans leur état d'origine dans l'ensemble de données d'origine.
- **Supprimez les observations exclues.** Les observations exclues sont supprimées de l'ensemble de données. Vous pouvez récupérer les observations supprimées uniquement en fermant le fichier sans enregistrer les modifications et en l'ouvrant à nouveau. La suppression des observations est définitive si vous enregistrez les modifications apportées au fichier de données.

Remarque : Si vous supprimez des observations exclues et enregistrez le fichier, vous ne pouvez pas récupérer ces observations.

Etat de la sélection d'une observation

Si vous avez sélectionné un sous-ensemble d'observations mais n'avez pas écarté les observations non sélectionnées, celles-ci sont identifiées dans l'éditeur de données par une ligne verticale dans le numéro de ligne.

Figure 10-11

Etat de la sélection d'une observation

*demo.sav [Ensemble_de_données2] - Editeur de données

Fichier Edition Affichage Données Transformer Analyse Graphes Outils Modules complémentaires Fenêtre Aide

1 : age 55 Visible : 30 variables sur 30

	age	marital	address	income	inccat	car	carcat
1	55	Marrié(e)	12	72,00	\$50 - \$74	36,20	Luxe
2	56	Non marri...	29	153,00	\$75 - \$1...	76,90	Luxe
3	28	Marrié(e)	9	28,00	\$25 - \$49	13,70	Economique
4	24	Marrié(e)	4	26,00	\$25 - \$49	12,50	Economique
5	25	Non marri...	2	23,00	Inf à \$25	11,30	Economique
6	45	Marrié(e)	9	76,00	\$75 - \$1...	37,20	Luxe
7	42	Non marri...	19	40,00	\$25 - \$49	19,80	Standard
8	35	Non marri...	15	57,00	\$50 - \$74	28,20	Standard
9	46	Non marri...	26	24,00	Inf à \$25	12,20	Economique
10	34	Marrié(e)	0	89,00	\$75 - \$1...	46,10	Luxe
11	55	Marrié(e)	17	72,00	\$50 - \$74	36,50	Luxe
12	28	Non marri...	3	24,00	Inf à \$25	11,80	Economique
13	31	Marrié(e)	9	40,00	\$25 - \$49	21,30	Standard

Affichage des données Affichage des variables

Processeur prêt Filtre actif

Fichiers d'exemple

Les fichiers d'exemple installés avec le produit figurent dans le sous-répertoire *Echantillons* du répertoire d'installation. Il existe un dossier distinct au sein du sous-répertoire *Echantillons* pour chacune des langues suivantes : Anglais, Français, Allemand, Italien, Japonais, Coréen, Polonais, Russe, Chinois simplifié, Espagnol et Chinois traditionnel.

Seuls quelques fichiers d'exemples sont disponibles dans toutes les langues. Si un fichier d'exemple n'est pas disponible dans une langue, le dossier de langue contient la version anglaise du fichier d'exemple.

Descriptions

Voici de brèves descriptions des fichiers d'exemple utilisés dans divers exemples à travers la documentation.

- **accidents.sav.** Ce fichier de données d'hypothèse concerne une société d'assurance qui étudie les facteurs de risque liés à l'âge et au sexe dans les accidents de la route survenant dans une région donnée. Chaque observation correspond à une classification croisée de la catégorie d'âge et du sexe.
- **adl.sav.** Ce fichier de données d'hypothèse concerne les mesures entreprises pour identifier les avantages d'un type de thérapie proposé aux patients qui ont subi une attaque cardiaque. Les médecins ont assigné de manière aléatoire les patients du sexe féminin ayant subi une attaque cardiaque à un groupe parmi deux groupes possibles. Le premier groupe a fait l'objet de la thérapie standard tandis que le second a bénéficié en plus d'une thérapie émotionnelle. Trois mois après les traitements, les capacités de chaque patient à effectuer les tâches ordinaires de la vie quotidienne ont été notées en tant que variables ordinales.
- **advert.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un détaillant pour examiner la relation existant entre l'argent dépensé dans la publicité et les ventes résultantes. Pour ce faire, il collecte les chiffres des ventes passées et les coûts associés à la publicité.
- **aflatoxin.sav.** Ce fichier de données d'hypothèse concerne le test de l'aflatoxine dans des récoltes de maïs. La concentration de ce poison varie largement d'une récolte à l'autre et au sein de chaque récolte. Un processeur de grain a reçu 16 échantillons issus de 8 récoltes de maïs et a mesuré les niveaux d'aflatoxine en parties par milliard (PPB).
- **aflatoxin20.sav.** Ce fichier de données contient les mesures d'aflatoxine de chacun des 16 échantillons des récoltes 4 et 8 du fichier de données *aflatoxin.sav*.
- **anorectic.sav.** En cherchant à développer une symptomatologie standardisée du comportement anorexique/boulimique, des chercheurs ont examiné 55 adolescents souffrant de troubles alimentaires. Chaque patient a été observé quatre fois sur une période de quatre années, soit

un total de 220 observations. A chaque observation, les patients ont été notés pour chacun des 16 symptômes. En raison de l'absence de scores de symptôme pour le patient 71/visite 2, le patient 76/visite 2 et le patient 47/visite 3, le nombre d'observations valides est de 217.

- **autoaccidents.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un analyste en assurances pour modéliser le nombre d'accidents de la route par conducteur tout en prenant en compte l'âge et le sexe du conducteur. Chaque observation représente un conducteur distinct et enregistre son sexe, son âge et le nombre d'accidents de la route au cours des cinq dernières années.
- **band.sav.** Ce fichier de données contient les chiffres de ventes hebdomadaires hypothétiques de CD musicaux d'un groupe. Les données relatives à trois variables explicatives possibles sont également incluses.
- **bankloan.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend une banque pour réduire le taux de défaut de paiement. Il contient des informations financières et démographiques sur 850 clients existants et éventuels. Les premières 700 observations concernent des clients auxquels des prêts ont été octroyés. Les 150 dernières observations correspondent aux clients éventuels que la banque doit classer comme bons ou mauvais risques de crédit.
- **bankloan_binning.sav.** Ce fichier de données d'hypothèse concerne des informations financières et démographiques sur 5 000 clients existants.
- **behavior.sav.** Dans un exemple classique, on a demandé à 52 étudiants de noter les combinaisons établies à partir de 15 situations et de 15 comportements sur une échelle de 0 à 9, où 0 = « extrêmement approprié » et 9 = « extrêmement inapproprié ». En effectuant la moyenne des résultats de l'ensemble des individus, on constate une certaine différence entre les valeurs.
- **behavior_ini.sav.** Ce fichier de données contient la configuration initiale d'une solution bidimensionnelle pour *behavior.sav*.
- **brakes.sav.** Ce fichier de données d'hypothèse concerne le contrôle qualité effectué dans une usine qui fabrique des freins à disque pour des voitures haut de gamme. Le fichier de données contient les mesures de diamètre de 16 disques de 8 machines de production. Le diamètre cible des freins est de 322 millimètres.
- **breakfast.sav.** Au cours d'une étude classique, on a demandé à 21 étudiants en MBA (Master of Business Administration) de l'école de Wharton et à leurs conjoints de classer 15 aliments du petit-déjeuner selon leurs préférences, de 1 = « aliment préféré » à 15 = « aliment le moins apprécié ». Leurs préférences ont été enregistrées dans six scénarios différents, allant de « Préférence générale » à « En-cas avec boisson uniquement ».
- **breakfast-overall.sav.** Ce fichier de données contient les préférences de petit-déjeuner du premier scénario uniquement, « Préférence générale ».
- **broadband_1.sav.** Ce fichier de données d'hypothèse concerne le nombre d'abonnés, par région, à un service haut débit. Le fichier de données contient le nombre d'abonnés mensuels de 85 régions sur une période de quatre ans.
- **broadband_2.sav.** Ce fichier de données est identique au fichier *broadband_1.sav* mais contient les données relatives à trois mois supplémentaires.

- **car_insurance_claims.sav.** Il s'agit d'un ensemble de données présenté et analysé ailleurs qui concerne des actions en indemnisation pour des voitures. Le montant d'action en indemnisation moyen peut être modélisé comme présentant une distribution gamma, à l'aide d'une fonction de lien inverse pour associer la moyenne de la variable dépendante à une combinaison linéaire de l'âge de l'assuré, du type de véhicule et de l'âge du véhicule. Le nombre d'actions entreprises peut être utilisé comme pondération de positionnement.
- **car_sales.sav.** Ce fichier de données contient des estimations de ventes hypothétiques, des barèmes de prix et des spécifications physiques concernant divers modèles et marques de véhicule. Les barèmes de prix et les spécifications physiques proviennent tour à tour de *edmunds.com* et des sites des constructeurs.
- **car_sales_uprepared.sav.** Il s'agit d'une version modifiée de *car_sales.sav* qui n'inclut aucune version transformée des champs.
- **carpet.sav.** Dans un exemple courant, une société intéressée par la commercialisation d'un nouveau nettoyeur de tapis souhaite examiner l'influence de cinq critères sur la préférence du consommateur : la conception du conditionnement, la marque, le prix, une étiquette *Economique* et une garantie satisfait ou remboursé. Il existe trois niveaux de critère pour la conception du conditionnement, suivant l'emplacement de l'applicateur, trois marques (*K2R*, *Glory* et *Bissell*), trois niveaux de prix et deux niveaux (non ou oui) pour chacun des deux derniers critères. Dix consommateurs classent 22 profils définis par ces critères. La variable *Préférence* indique le classement des rangs moyens de chaque profil. Un rang faible correspond à une préférence élevée. Cette variable reflète une mesure globale de préférence pour chaque profil.
- **carpet_prefs.sav.** Ce fichier de données repose sur le même exemple que celui décrit pour *carpet.sav*, mais contient les classements réels issus de chacun des 10 clients. On a demandé aux consommateurs de classer les 22 profils de produits, du préféré au moins intéressant. Les variables *PREF1* à *PREF22* contiennent les identificateurs des profils associés, tels qu'ils sont définis dans *carpet_plan.sav*.
- **catalog.sav.** Ce fichier de données contient des chiffres de ventes mensuelles hypothétiques relatifs à trois produits vendus par une entreprise de vente par correspondance. Les données relatives à cinq variables explicatives possibles sont également incluses.
- **catalog_seasfac.sav.** Ce fichier de données est identique à *catalog.sav* mais contient en plus un ensemble de facteurs saisonniers calculés à partir de la procédure de désaisonnalisation, ainsi que les variables de date correspondantes.
- **cellular.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un opérateur téléphonique pour réduire les taux de désabonnement. Des scores de propension au désabonnement sont attribués aux comptes, de 0 à 100. Les comptes ayant une note égale ou supérieure à 50 sont susceptibles de changer de fournisseur.
- **ceramics.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un fabricant pour déterminer si un nouvel alliage haute qualité résiste mieux à la chaleur qu'un alliage standard. Chaque observation représente un test séparé de l'un des deux alliages ; le degré de chaleur auquel l'alliage ne résiste pas est enregistré.
- **cereal.sav.** Ce fichier de données d'hypothèse concerne un sondage de 880 personnes interrogées sur leurs préférences de petit-déjeuner et sur leur âge, leur sexe, leur situation familiale et leur mode de vie (actif ou non actif, selon qu'elles pratiquent une activité physique au moins deux fois par semaine). Chaque observation correspond à un répondant distinct.

- **clothing_defects.sav.** Ce fichier de données d'hypothèse concerne le processus de contrôle qualité observé dans une usine de textile. Dans chaque lot produit à l'usine, les inspecteurs prélèvent un échantillon de vêtements et comptent le nombre de vêtements qui ne sont pas acceptables.
- **coffee.sav.** Ce fichier de données concerne l'image perçue de six marques de café frappé . Pour chacun des 23 attributs d'image de café frappé, les personnes sollicitées ont sélectionné toutes les marques décrites par l'attribut. Les six marques sont appelées AA, BB, CC, DD, EE et FF à des fins de confidentialité.
- **contacts.sav.** Ce fichier de données d'hypothèse concerne les listes de contacts d'un groupe de représentants en informatique d'entreprise. Chaque contact est classé selon le service de l'entreprise où il travaille et le classement de son entreprise. Sont également enregistrés le montant de la dernière vente effectuée, le temps passé depuis la dernière vente et la taille de l'entreprise du contact.
- **creditpromo.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprind un grand magasin pour évaluer l'efficacité d'une promotion récente de carte de crédit. A cette fin, 500 détenteurs de carte ont été sélectionnés au hasard. La moitié a reçu une publicité faisant la promotion d'un taux d'intérêt réduit sur les achats effectués dans les trois mois à venir. L'autre moitié a reçu une publicité saisonnière standard.
- **customer_dbase.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprind une société pour utiliser les informations figurant dans sa banque de données et proposer des offres spéciales aux clients susceptibles d'être intéressés. Un sous-groupe de la base de clients a été sélectionné au hasard et a reçu des offres spéciales. Les réponses des clients ont été enregistrées.
- **customer_information.sav.** Un fichier de données d'hypothèse qui contient les informations postales du client, telles que le nom et l'adresse.
- **customer_subset.sav.** Un sous-ensemble de 80 observations de *customer_dbase.sav*.
- **customers_model.sav.** Ce fichier de données d'hypothèse concerne les personnes ciblées par une campagne de marketing. Ces données incluent des informations démographiques, un récapitulatif de l'historique d'achat et indiquent si chaque personne a répondu ou non à la campagne. Chaque observation représente une personne distincte.
- **customers_new.sav.** Ce fichier de données d'hypothèse concerne les personnes constituant des cibles potentielles pour une campagne de marketing. Ces données incluent des informations démographiques et un récapitulatif de l'historique d'achat pour chaque personne. Chaque observation représente une personne distincte.
- **debate.sav.** Ce fichier de données d'hypothèse concerne des réponses appariées à une enquête donnée aux participants à un débat politique avant et après le débat. Chaque observation représente un répondant distinct.
- **debate_aggregate.sav.** Il s'agit d'un fichier de données d'hypothèse qui rassemble les réponses dans le fichier *debate.sav*. Chaque observation correspond à une classification croisée de préférence avant et après le débat.
- **demo.sav.** Ce fichier de données d'hypothèse concerne une base de données clients achetée en vue de diffuser des offres mensuelles. Les données indiquent si le client a répondu ou non à l'offre et contiennent diverses informations démographiques.

- **demo_cs_1.sav.** Ce fichier de données d'hypothèse concerne la première mesure entreprise par une société pour compiler une base de données contenant des informations d'enquête. Chaque observation correspond à une ville différente. La région, la province, le quartier et la ville sont enregistrés.
- **demo_cs_2.sav.** Ce fichier de données d'hypothèse concerne la seconde mesure entreprise par une société pour compiler une base de données contenant des informations d'enquête. Chaque observation correspond à un ménage différent issu des villes sélectionnées à la première étape. La région, la province, le quartier, la ville, la sous-division et l'identification sont enregistrés. Les informations d'échantillonnage des deux premières étapes de la conception sont également incluses.
- **demo_cs.sav.** Ce fichier de données d'hypothèse concerne des informations d'enquête collectées via une méthode complexe d'échantillonnage. Chaque observation correspond à un ménage différent et diverses informations géographiques et d'échantillonnage sont enregistrées.
- **dmdata.sav.** Ceci est un fichier de données d'hypothèse qui contient des informations démographiques et des informations concernant les achats pour une entreprise de marketing direct. *dmdata2.sav* contient les informations pour un sous-ensemble de contacts qui ont reçu un envoi d'essai, et *dmdata3.sav* contient des informations sur les contacts restants qui n'ont pas reçu l'envoi d'essai.
- **dietstudy.sav.** Ce fichier de données d'hypothèse contient les résultats d'une étude portant sur le régime de Stillman. Chaque observation correspond à un sujet distinct et enregistre son poids en livres avant et après le régime, ainsi que ses niveaux de triglycérides en mg/100 ml.
- **dvdplayer.sav.** Ce fichier de données d'hypothèse concerne le développement d'un nouveau lecteur DVD. À l'aide d'un prototype, l'équipe de marketing a collecté des données de groupes spécifiques. Chaque observation correspond à un utilisateur interrogé et enregistre des informations démographiques sur cet utilisateur, ainsi que ses réponses aux questions portant sur le prototype.
- **german_credit.sav.** Ce fichier de données provient de l'ensemble de données « German credit » figurant dans le référentiel Machine Learning Databases de l'université de Californie, Irvine.
- **grocery_1month.sav.** Ce fichier de données d'hypothèse est le fichier de données *grocery_coupons.sav* dans lequel les achats hebdomadaires sont organisés par client distinct. Certaines variables qui changeaient toutes les semaines disparaissent. En outre, le montant dépensé enregistré est à présent la somme des montants dépensés au cours des quatre semaines de l'enquête.
- **grocery_coupons.sav.** Il s'agit d'un fichier de données d'hypothèse qui contient des données d'enquête collectées par une chaîne de magasins d'alimentation qui cherche à déterminer les habitudes de consommation de ses clients. Chaque client est suivi pendant quatre semaines et chaque observation correspond à une semaine distincte. Les informations enregistrées concernent les endroits où le client effectue ses achats, la manière dont il les effectue, ainsi que les sommes dépensées en provisions au cours de cette semaine.
- **guttman.sav.** Bell a présenté un tableau pour illustrer les groupes sociaux possibles. Guttman a utilisé une partie de ce tableau, dans lequel cinq variables décrivant des éléments tels que l'interaction sociale, le sentiment d'appartenance à un groupe, la proximité physique des membres et la formalité de la relation, ont été croisées avec sept groupes sociaux théoriques, dont les foules (par exemple, le public d'un match de football), l'audience (par exemple, au

cinéma ou dans une salle de classe), le public (par exemple, les journaux ou la télévision), les bandes (proche d'une foule, mais qui serait caractérisée par une interaction beaucoup plus intense), les groupes primaires (intimes), les groupes secondaires (volontaires) et la communauté moderne (groupement lâche issu d'une forte proximité physique et d'un besoin de services spécialisés).

- **health_funding.sav.** Ce fichier de données d'hypothèse concerne des données sur le financement des soins de santé (montant par groupe de 100 individus), les taux de maladie (taux par groupe de 10 000 individus) et les visites chez les prestataires de soins de santé (taux par groupe de 10 000 individus). Chaque observation représente une ville différente.
- **hivassay.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un laboratoire pharmaceutique pour développer une analyse rapide de détection d'infection HIV. L'analyse a pour résultat huit nuances de rouge, les nuances les plus marquées indiquant une plus forte probabilité d'infection. Un test en laboratoire a été effectué sur 2 000 échantillons de sang, la moitié de ces échantillons étant infectée par le virus HIV et l'autre moitié étant saine.
- **hourlywagedata.sav.** Ce fichier de données d'hypothèse concerne les salaires horaires d'infirmières occupant des postes administratifs et dans les services de soins, et affichant divers niveaux d'expérience.
- **insurance_claims.sav.** Il s'agit d'un fichier de données hypothétiques qui concerne une compagnie d'assurance souhaitant développer un modèle pour signaler des réclamations suspectes, potentiellement frauduleuses. Chaque observation correspond à une réclamation distincte.
- **insure.sav.** Ce fichier de données d'hypothèse concerne une compagnie d'assurance qui étudie les facteurs de risque indiquant si un client sera amené à déclarer un incident au cours d'un contrat d'assurance vie d'une durée de 10 ans. Chaque observation figurant dans le fichier de données représente deux contrats, l'un ayant enregistré une réclamation et l'autre non, appariés par âge et sexe.
- **judges.sav.** Ce fichier de données d'hypothèse concerne les scores attribués par des juges expérimentés (plus un juge enthousiaste) à 300 performances de gymnastique. Chaque ligne représente une performance distincte ; les juges ont examiné les mêmes performances.
- **kinship_dat.sav.** Rosenberg et Kim se sont lancés dans l'analyse de 15 termes de parenté (cousin/cousine, fille, fils, frère, grand-mère, grand-père, mère, neveu, nièce, oncle, père, petite-fille, petit-fils, sœur, tante). Ils ont demandé à quatre groupes d'étudiants (deux groupes de femmes et deux groupes d'hommes) de trier ces termes en fonction des similarités. Deux groupes (un groupe de femmes et un groupe d'hommes) ont été invités à effectuer deux tris, en basant le second sur un autre critère que le premier. Ainsi, un total de six "sources" a été obtenu. Chaque source correspond à une matrice de proximité 15×15 , dont le nombre de cellules est égal au nombre de personnes dans une source moins le nombre de fois où les objets ont été partitionnés dans cette source.
- **kinship_ini.sav.** Ce fichier de données contient une configuration initiale d'une solution tridimensionnelle pour *kinship_dat.sav*.
- **kinship_var.sav.** Ce fichier de données contient les variables indépendantes *sexe*, *génér(ation)* et *degré* (de séparation) permettant d'interpréter les dimensions d'une solution pour *kinship_dat.sav*. Elles permettent en particulier de réduire l'espace de la solution à une combinaison linéaire de ces variables.

- **marketvalues.sav.** Ce fichier de données concerne les ventes de maisons dans un nouvel ensemble à Algonquin (Illinois) au cours des années 1999–2000. Ces ventes relèvent des archives publiques.
- **nhis2000_subset.sav.** Le NHIS (National Health Interview Survey) est une enquête de grande envergure concernant la population des États-Unis. Des entretiens ont lieu avec un échantillon de ménages représentatifs de la population américaine. Des informations démographiques et des observations sur l'état de santé et le comportement sanitaire sont recueillies auprès des membres de chaque ménage. Ce fichier de données contient un sous-groupe d'informations issues de l'enquête de 2000. National Center for Health Statistics. National Health Interview Survey, 2000. Fichier de données et documentation d'usage public. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/. Accès en 2003.
- **ozone.sav.** Les données incluent 330 observations portant sur six variables météorologiques pour prévoir la concentration d'ozone à partir des variables restantes. Des chercheurs précédents , , ont décelé parmi ces variables des non-linéarités qui pénalisent les approches standard de la régression.
- **pain_medication.sav.** Ce fichier de données d'hypothèse contient les résultats d'un essai clinique d'un remède anti-inflammatoire traitant les douleurs de l'arthrite chronique. On cherche notamment à déterminer le temps nécessaire au médicament pour agir et les résultats qu'il permet d'obtenir par rapport à un médicament existant.
- **patient_los.sav.** Ce fichier de données d'hypothèse contient les dossiers médicaux de patients admis à l'hôpital pour suspicion d'infarctus du myocarde suspecté (ou « attaque cardiaque »). Chaque observation correspond à un patient distinct et enregistre de nombreuses variables liées à son séjour à l'hôpital.
- **patlos_sample.sav.** Ce fichier de données d'hypothèse contient les dossiers médicaux d'un échantillon de patients sous traitement thrombolytique après un infarctus du myocarde. Chaque observation correspond à un patient distinct et enregistre de nombreuses variables liées à son séjour à l'hôpital.
- **polishing.sav.** Il s'agit du fichier de données du « Nambeware Polishing Times » de la Data and Story Library. Il concerne les mesures qu'entreprend un fabricant de vaisselle en métal (Nambe Mills, Santa Fe, Nouveau-Mexique) pour planifier sa production. Chaque observation représente un article différent de la gamme de produits. Le diamètre, le temps de polissage, le prix et le type de produit sont enregistrés pour chaque article.
- **poll_cs.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un enquêteur pour déterminer le niveau de soutien du public pour un projet de loi avant législature. Les observations correspondent à des électeurs enregistrés. Chaque observation enregistre le comté, la ville et le quartier où habite l'électeur.
- **poll_cs_sample.sav.** Ce fichier de données d'hypothèse contient un échantillon des électeurs répertoriés dans le fichier *poll_cs.sav*. L'échantillon a été prélevé selon le plan spécifié dans le fichier de plan *poll.csplan* et ce fichier de données enregistre les probabilités d'inclusion et les pondérations d'échantillon. Toutefois, ce plan faisant appel à une méthode d'échantillonnage de probabilité proportionnelle à la taille (PPS – Probability-Proportional-to-Size), il existe également un fichier contenant les probabilités de sélection conjointes (*poll_jointprob.sav*). Les variables supplémentaires correspondant à la répartition démographique des électeurs et à leur opinion sur le projet de loi proposé ont été collectées et ajoutées au fichier de données une fois l'échantillon prélevé.

- **property_assess.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un contrôleur au niveau du comté pour maintenir les évaluations de valeur de propriété à jour sur des ressources limitées. Les observations correspondent à des propriétés vendues dans le comté au cours de l'année précédente. Chaque observation du fichier de données enregistre la ville où se trouve la propriété, l'évaluateur ayant visité la propriété pour la dernière fois, le temps écoulé depuis cette évaluation, l'évaluation effectuée à ce moment-là et la valeur de vente de la propriété.
- **property_assess_cs.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un contrôleur du gouvernement pour maintenir les évaluations de valeur de propriété à jour sur des ressources limitées. Les observations correspondent à des propriétés de l'état. Chaque observation du fichier de données enregistre le comté, la ville et le quartier où se trouve la propriété, le temps écoulé depuis la dernière évaluation et l'évaluation alors effectuée.
- **property_assess_cs_sample.sav.** Ce fichier de données d'hypothèse contient un échantillon des propriétés répertoriées dans le fichier *property_assess_cs.sav*. L'échantillon a été prélevé selon le plan spécifié dans le fichier de plan *property_assess_csplan* et ce fichier de données enregistre les probabilités d'inclusion et les pondérations d'échantillon. La variable supplémentaire *Valeur courante* a été collectée et ajoutée au fichier de données une fois l'échantillon prélevé.
- **recidivism.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend une agence administrative d'application de la loi pour interpréter les taux de récidive dans la juridiction. Chaque observation correspond à un récidiviste et enregistre les informations démographiques qui lui sont propres, certains détails sur le premier délit commis, ainsi que le temps écoulé jusqu'à la seconde arrestation si elle s'est produite dans les deux années suivant la première.
- **recidivism_cs_sample.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend une agence administrative d'application de la loi pour interpréter les taux de récidive dans la juridiction. Chaque observation correspond à un récidiviste libéré suite à la première arrestation en juin 2003 et enregistre les informations démographiques qui lui sont propres, certains détails sur le premier délit commis et les données relatives à la seconde arrestation, si elle a eu lieu avant fin juin 2006. Les récidivistes ont été choisis dans plusieurs départements échantillonnés conformément au plan d'échantillonnage spécifié dans *recidivism_cs_csplan*. Ce plan faisant appel à une méthode d'échantillonnage de probabilité proportionnelle à la taille (PPS - Probability proportional to size), il existe également un fichier contenant les probabilités de sélection conjointes (*recidivism_cs_jointprob.sav*).
- **rfm_transactions.sav.** Un fichier de données d'hypothèse qui contient les données de transaction d'achat, y compris la date d'achat, le/les élément(s) acheté(s) et le montant monétaire pour chaque transaction.
- **salesperformance.sav.** Ce fichier de données d'hypothèse concerne l'évaluation de deux nouveaux cours de formation en vente. Soixante employés, divisés en trois groupes, reçoivent chacun une formation standard. En outre, le groupe 2 suit une formation technique et le groupe 3 un didacticiel pratique. À l'issue du cours de formation, chaque employé est testé et sa note enregistrée. Chaque observation du fichier de données représente un stagiaire distinct et enregistre le groupe auquel il a été assigné et la note qu'il a obtenue au test.
- **satisf.sav.** Il s'agit d'un fichier de données d'hypothèse portant sur une enquête de satisfaction effectuée par une société de vente au détail au niveau de quatre magasins. Un total de 582 clients ont été interrogés et chaque observation représente la réponse d'un seul client.

- **screws.sav.** Ce fichier de données contient des informations sur les descriptives des vis, des boulons, des écrous et des clous..
- **shampoo_ph.sav.** Ce fichier de données d'hypothèse concerne le processus de contrôle qualité observé dans une usine de produits capillaires. A intervalles réguliers, six lots de sortie distincts sont mesurés et leur pH enregistré. La plage cible est 4,5–5,5.
- **ships.sav.** Il s'agit d'un ensemble de données présenté et analysé ailleurs et concernant les dommages causés à des cargos par les vagues. Les effectifs d'incidents peuvent être modélisés comme des incidents se produisant selon un taux de Poisson en fonction du type de navire, de la période de construction et de la période de service. Les mois de service totalisés pour chaque cellule du tableau formé par la classification croisée des facteurs fournissent les valeurs d'exposition au risque.
- **site.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend une société pour choisir de nouveaux sites pour le développement de ses activités. L'entreprise a fait appel à deux consultants pour évaluer séparément les sites. Ces consultants, en plus de fournir un rapport approfondi, ont classé chaque site comme constituant une éventualité « bonne », « moyenne » ou « faible ».
- **smokers.sav.** Ce fichier de données est extrait de l'étude National Household Survey of Drug Abuse de 1998 et constitue un échantillon de probabilité des ménages américains. (<http://dx.doi.org/10.3886/ICPSR02934>) Ainsi, la première étape dans l'analyse de ce fichier doit consister à pondérer les données pour refléter les tendances de population.
- **stroke_clean.sav.** Ce fichier de données d'hypothèse concerne l'état d'une base de données médicales une fois celle-ci purgée via des procédures de l'option Validation de données.
- **stroke_invalid.sav.** Ce fichier de données d'hypothèse concerne l'état initial d'une base de données médicales et comporte plusieurs erreurs de saisie de données.
- **stroke_survival.** Ce fichier de données d'hypothèse concerne les temps de survie de patients qui quittent un programme de rééducation à la suite d'un accident ischémique et rencontrent un certain nombre de problèmes. Après l'attaque, l'occurrence d'infarctus du myocarde, d'accidents ischémiques ou hémorragiques est signalée, et le moment de l'événement enregistré. L'échantillon est tronqué à gauche car il n'inclut que les patients ayant survécu durant le programme de rééducation mis en place suite à une attaque.
- **stroke_valid.sav.** Ce fichier de données d'hypothèse concerne l'état d'une base de données médicales une fois les valeurs vérifiées via la procédure Validation de données. Elle contient encore des observations anormales potentielles.
- **survey_sample.sav.** Ce fichier de données concerne des informations d'enquête dont des données démographiques et des mesures comportementales. Il est basé sur un sous-ensemble de variables de la 1998 NORC General Social Survey, bien que certaines valeurs de données aient été modifiées et que des variables supplémentaires fictives aient été ajoutées à titre de démonstration.
- **telco.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend une société de télécommunications pour réduire les taux de désabonnement de sa base de clients. Chaque observation correspond à un client distinct et enregistre diverses informations démographiques et d'utilisation de service.

- **telco_extra.sav.** Ce fichier de données est semblable au fichier de données *telco.sav* mais les variables de permanence et de dépenses des consommateurs transformées log ont été supprimées et remplacées par des variables de dépenses des consommateurs transformées log standardisées.
- **telco_missing.sav.** Ce fichier de données est un sous-ensemble du fichier de données *telco.sav* mais certaines des valeurs de données démographiques ont été remplacées par des valeurs manquantes.
- **testmarket.sav.** Ce fichier de données d'hypothèse concerne une chaîne de fast foods et ses plans marketing visant à ajouter un nouveau plat à son menu. Trois campagnes étant possibles pour promouvoir le nouveau produit, le nouveau plat est introduit sur des sites sur plusieurs marchés sélectionnés au hasard. Une promotion différente est effectuée sur chaque site et les ventes hebdomadaires du nouveau plat sont enregistrées pour les quatre premières semaines. Chaque observation correspond à un site-semaine distinct.
- **testmarket_1month.sav.** Ce fichier de données d'hypothèse est le fichier de données *testmarket.sav* dans lequel les ventes hebdomadaires sont organisées par site distinct. Certaines variables qui changeaient toutes les semaines disparaissent. En outre, les ventes enregistrées sont à présent la somme des ventes réalisées au cours des quatre semaines de l'enquête.
- **tree_car.sav.** Ce fichier de données d'hypothèse concerne des données démographiques et de prix d'achat de véhicule.
- **tree_credit.sav.** Ce fichier de données d'hypothèse concerne des données démographiques et d'historique de prêt bancaire.
- **tree_missing_data.sav** Ce fichier de données d'hypothèse concerne des données démographiques et d'historique de prêt bancaire avec un grand nombre de valeurs manquantes.
- **tree_score_car.sav.** Ce fichier de données d'hypothèse concerne des données démographiques et de prix d'achat de véhicule.
- **tree_textdata.sav.** Ce fichier de données simples ne comporte que deux variables et vise essentiellement à indiquer l'état par défaut des variables avant affectation du niveau de mesure et des étiquettes de valeurs.
- **tv-survey.sav.** Ce fichier de données d'hypothèse concerne une enquête menée par un studio de télévision qui envisage de prolonger la diffusion d'un programme ou de l'arrêter. On a demandé à 906 personnes si elles regarderaient le programme dans diverses situations. Chaque ligne représente un répondant distinct et chaque colonne une situation distincte.
- **ulcer_recurrence.sav.** Ce fichier contient des informations partielles d'une enquête visant à comparer l'efficacité de deux thérapies de prévention de la récurrence des ulcères. Il fournit un bon exemple de données censurées par intervalle et a été présenté et analysé ailleurs .
- **ulcer_recurrence_recoded.sav.** Ce fichier réorganise les informations figurant dans le fichier *ulcer_recurrence.sav* pour que vous puissiez modéliser la probabilité d'événement pour chaque intervalle de l'enquête plutôt que la probabilité d'événement de fin d'enquête. Il a été présenté et analysé ailleurs .
- **verd1985.sav.** Ce fichier de données concerne une enquête . Les réponses de 15 sujets à 8 variables ont été enregistrées. Les variables présentant un intérêt sont divisées en trois ensembles. Le groupe 1 comprend l'âge et la *situation familiale*, le groupe 2 les *animaux domestiques* et la *presse*, et le groupe 3 la *musique* et l'*habitat*. A la variable *animal*

domestique est appliqué un codage nominal multiple et à *âge*, un codage ordinal ; toutes les autres variables ont un codage nominal simple.

- **virus.sav.** Ce fichier de données d'hypothèse concerne les mesures qu'entreprend un fournisseur de services Internet pour déterminer les effets d'un virus sur ses réseaux. Il a suivi le pourcentage (approximatif) de trafic de messages électroniques infectés par un virus sur ses réseaux sur la durée, de la découverte à la circonscription de la menace.
- **wheeze_steubenville.sav.** Il s'agit d'un sous-ensemble d'une enquête longitudinale des effets de la pollution de l'air sur la santé des enfants . Les données contiennent des mesures binaires répétées de l'état asthmatique d'enfants de la ville de Steubenville (Ohio), âgés de 7, 8, 9 et 10 ans, et indiquent si la mère fumait au cours de la première année de l'enquête.
- **workprog.sav.** Ce fichier de données d'hypothèse concerne un programme de l'administration visant à proposer de meilleurs postes aux personnes défavorisées. Un échantillon de participants potentiels au programme a ensuite été prélevé. Certains de ces participants ont été sélectionnés au hasard pour participer au programme. Chaque observation représente un participant au programme distinct.

Notices

Licensed Materials – Property of SPSS Inc., an IBM Company. © Copyright SPSS Inc. 1989, 2010.

Patent No. 7,023,453

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: SPSS INC., AN IBM COMPANY, PROVIDES THIS PUBLICATION “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. SPSS Inc. may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-SPSS and non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this SPSS Inc. product and use of those Web sites is at your own risk.

When you send information to IBM or SPSS, you grant IBM and SPSS a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

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

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to SPSS Inc., for the purposes of developing,

using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. SPSS Inc., therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided “AS IS”, without warranty of any kind. SPSS Inc. shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and [ibm.com](http://www.ibm.com) are trademarks of IBM Corporation, registered in many jurisdictions worldwide. A current list of IBM trademarks is available on the Web at <http://www.ibm.com/legal/copytrade.shtml>.

SPSS is a trademark of SPSS Inc., an IBM Company, registered in many jurisdictions worldwide.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Java and all Java-based trademarks and logos are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

This product uses WinWrap Basic, Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com>.

Other product and service names might be trademarks of IBM, SPSS, or other companies.

Adobe product screenshot(s) reprinted with permission from Adobe Systems Incorporated.

Microsoft product screenshot(s) reprinted with permission from Microsoft Corporation.



- Access (Microsoft), 14
- Assistant Date et heure, 139
- Assistant de base de données, 14
- Assistant d'importation de texte, 20
- attribution d'un nouveau nom aux ensembles de données, 55
- Attributs de variable
 - Réutilisation, 42

- Calcul de nouvelles variables, 133
- Collage d'une syntaxe
 - Boîte de dialogue, 120
- copie d'attributs de variables, 42
- Création d'étiquettes de variable, 33

- Data Editor
 - plusieurs fichiers de données ouverts, 51
- déplacement
 - Eléments dans des tableaux pivotants, 90
 - Eléments dans le Viewer, 87
- diagrammes
 - Bâtons, 58, 64
 - création de diagrammes, 64
 - Edition de diagrammes, 70
 - Histogrammes, 62
 - Modèles, 78
 - Options de diagramme, 83
- Diagrammes en bâtons, 58
- Données chaîne
 - Saisie de données, 31
- Données continues, 56
- Données de type ratio, 56
- Données d'échelle, 56
- Données d'intervalle, 56
- Données nominales, 56
- Données numériques, 28
- Données ordinales, 56
- Données qualitatives, 56
 - Mesures récapitulatives, 56
- Données quantitatives, 56

- Editeur de données
 - Saisie de données non numériques, 31
 - Saisie de données numériques, 28
- Edition de tableaux pivotants, 94
- effectif
 - Tableaux d'effectifs, 56
- ensembles de données
 - renommer, 55
- Etiquettes de valeurs
 - Affectation, 34, 36
 - Contrôle de l'affichage dans le Viewer, 34, 36
 - Variables non numériques, 36
 - Variables numériques, 34
- Etiquettes de variable
 - Création, 33
- Excel (Microsoft)
 - exportation des résultats, 108
- exportation des résultats
 - dans Excel, 108
 - dans PowerPoint, 108
 - dans Word, 108
 - HTML, 119
- Expressions conditionnelles, 137

- Fenêtres de syntaxe
 - codage par couleur, 123
 - Collage de commandes, 120
 - Exécution de commandes, 121, 124
 - Modification de commandes, 122
 - points d'arrêt, 125
 - saisie semi-automatique, 122
- Fichiers de base de données
 - Lecture, 14
- Fichiers de données
 - plusieurs fichiers de données ouverts, 51
- Fichiers de données texte
 - Lecture, 20
- Fichiers de feuille de calcul
 - Lecture, 12
 - Lecture des noms de variable, 12
- Fichiers de syntaxe
 - Ouverture, 124
- fichiers d'exemple
 - emplacement, 156
- Fichiers Excel
 - Lecture, 12
- fonctions dans les expressions, 135

- graphiques
 - Bâtons, 64
 - création de diagrammes, 64
 - Modèles, 78
 - modification de diagrammes, 70
 - Options de diagramme, 83

- Histogrammes, 62
- HTML
 - exportation des résultats, 119

- legal notices, 167

-
- masquage des lignes et colonnes dans des tableaux pivotants, 95
 - Mesures récapitulatives
 - Données qualitatives, 56
 - Variables d'échelle, 60
 - niveau de mesure, 56
 - Niveau de mesure, 56
 - Observations
 - Sélection, 150
 - Tri, 147, 150
 - Outil Aide syntaxe, 123
 - plusieurs fichiers de données ouverts, 51
 - suppression, 55
 - PowerPoint (Microsoft)
 - exportation des résultats, 108
 - Recodage de valeurs, 127
 - Saisie de données, 28, 31
 - Non numérique, 31
 - numérique, 28
 - Sélection d'observations, 150
 - Sous-groupes d'observations
 - Dates et heures, 153
 - Echantillon aléatoire, 152
 - Expressions conditionnelles, 151
 - Filtrage des observations exclues, 154
 - Sélection, 150
 - Selon une condition logique, 151
 - Suppression d'observations exclues, 154
 - strates
 - Création dans des tableaux pivotants, 92
 - Syntaxe, 120
 - tableaux de fréquences, 56
 - Tableaux pivotants
 - Accès aux définitions, 89
 - Formatage, 94
 - Formats de cellule, 95
 - Masquage de lignes et de colonnes, 95
 - Masquage du séparateur décimal, 95
 - modification, 94
 - Strates, 92
 - Structure pivotante, 90
 - Transposition de lignes et de colonnes, 90
 - Types de données des cellules, 95
 - trademarks, 168
 - traitement d'un fichier scindé, 148
 - Transposition de lignes et de colonnes dans des tableaux pivotants, 90
 - Tri d'observations, 147
 - Types de données
 - Pour les variables, 34
 - Valeurs manquantes
 - manquantes par défaut, 38
 - Variables non numériques, 41
 - Variables numériques, 39
 - Valeurs manquantes par défaut , 38
 - Variables, 28
 - Etiquettes, 33
 - Types de données, 34
 - Variables de date et d'heure, 139
 - Variables d'échelle
 - Mesures récapitulatives, 60
 - Viewer
 - Déplacement du résultat, 87
 - Masquage et affichage du résultat, 87
 - Word (Microsoft)
 - exportation des résultats, 108