
Introduction au logiciel SPSS version 11 pour Windows

**Patrick Gosselin
Hans Ivers
SERVICE D'AIDE À LA RECHERCHE**

École de Psychologie



**UNIVERSITÉ
LAVAL**

© Patrick Gosselin & Hans Ivers, 2003

Tous droits réservés

Le contenu de ce guide a été développé dans le cadre des activités du Service d'Aide à la Recherche (SAR) grâce au support financier de l'École de psychologie de l'Université Laval. Il ne peut être vendu ou utilisé, en tout ou en partie, à des fins commerciales. Les auteurs demeurent les seuls responsables de son contenu.

Tout commentaire ou suggestion peut être signalé au SAR, organisme responsable de la mise à jour du guide, à l'adresse électronique suivante : sar@psy.ulaval.ca
ou sur l'Internet : <http://www.psy.ulaval.ca/~sar/>

Version 1.2

Mise à jour le 5 septembre 2003

AVANT PROPOS

Ce document constitue une introduction pratique à l'utilisation du logiciel SPSS (*Statistical Package for the Social Sciences*) version 11 pour Windows. Il s'adresse principalement aux personnes qui désirent s'initier aux rudiments de cet outil informatique et plus particulièrement aux étudiants et professeurs de l'École de psychologie de l'Université Laval. Ce guide est notamment destiné aux étudiants des cours *Méthodes Quantitatives I* (MAT-12239), *Méthodes Quantitatives II* (MAT-12243), *Recherche dirigée I* (PSY-17297) et *Recherche dirigée II* (PSY-17298) du baccalauréat spécialisé en psychologie, ainsi qu'aux étudiants des cours *Analyse des Plans d'Expériences* (PSY-62558) et *Analyses Multivariées* (PSY-62559) du doctorat en psychologie. Nous espérons qu'il permettra aux étudiants réalisant une recherche dirigée ou une thèse d'effectuer de façon autonome et efficace leur démarche de saisie et d'analyse statistique de données.

Plusieurs personnes ont contribué de près ou de loin à la réalisation de ce guide. Nous tenons à remercier tout spécialement le Dr François Doré, directeur de l'École de psychologie, pour avoir mis à notre disposition les ressources nécessaires au développement de ce manuel et pour supporter et encourager les opérations du *Service d'Aide à la Recherche* (SAR). Son support contribue grandement à parfaire les connaissances statistiques et méthodologiques des étudiants de l'École de psychologie de l'Université Laval.

Patrick Gosselin
Candidat au doctorat
École de psychologie

Hans Ivers
Coordonnateur
Service d'Aide à la Recherche

TABLE DES MATIÈRES

Avant propos.....	i
Table des matières	iii
1. Introduction	1
Comment se procurer le logiciel SPSS ?.....	2
2. Présentation du logiciel SPSS	3
Le démarrage du logiciel	3
Les principales fenêtres du logiciel.....	4
3. L'Éditeur de données SPSS	7
Description générale.....	7
Les menus et les sous-menus de l'Éditeur de données	7
Le parcours des différents menus	9
4. La gestion des données.....	11
La saisie directe des données	11
Les trois étapes de la saisie des données.....	12
Comment obtenir un sommaire des variables ?	14
L'ouverture d'un fichier SPSS.....	15
L'importation des données dans SPSS	15
La manipulation des données.....	25
La transformation des données	34
5. L'analyse statistique des données	40
L'utilisation des fenêtres de dialogue.....	40
Les sous-menus du menu Analyse.....	41
Exemple d'une analyse exploratoire des données	44
Exemple d'un test t de Student pour échantillons indépendants	46
Exemple du calcul d'un coefficient de corrélation de Pearson.....	47
Exemple du calcul d'un coefficient du khi-deux.....	49
6. La représentation graphique des données.....	52
Illustration de la dispersion des données.....	52
7. Le Navigateur SPSS	56
Description générale.....	56
L'identification et la transformation du contenu	57
La sauvegarde des résultats SPSS	57
L'impression des résultats SPSS.....	57
L'exportation des résultats SPSS.....	58

8. La programmation en mode syntaxe	60
L'Éditeur de syntaxe.....	60
Exemple d'une transformation de données en mode syntaxe.....	61
Exemple d'une analyse statistique en mode syntaxe.....	63
Liste des commandes de syntaxe utilisées dans ce manuel	65
9. Les ressources d'aide du logiciel SPSS	66
Les ressources d'aide internes au logiciel SPSS	66
Les ressources d'aide externes au logiciel SPSS	69
10. Références	72

1. INTRODUCTION

L'analyse des données représente une étape importante dans la démarche d'une recherche scientifique. L'avancement des connaissances et des technologies en matière d'analyses statistiques a fait en sorte que plusieurs logiciels informatiques sont maintenant disponibles pour aider le chercheur à organiser, transformer et analyser ses données. Les plus connus sont les logiciels SAS, EXCEL, SPSS, STATVIEW et BMDP.

Le logiciel SPSS constitue l'un des logiciels les plus utilisés en sciences sociales. Il permet entre autres d'emmagasiner les données d'une recherche, de les transformer, de les analyser et de représenter les résultats d'analyses à l'aide de tableaux et de graphiques détaillés. En raison de son interface graphique et ses menus clairs et compréhensibles, SPSS représente un logiciel des plus faciles à utiliser, permettant ainsi aux chercheurs en formation d'effectuer de façon autonome leurs analyses statistiques.

Ce document constitue un guide d'introduction au logiciel SPSS pour Windows, version 11. Il a comme principal objectif de fournir au lecteur des informations pertinentes sur l'utilisation et le fonctionnement du logiciel. Il est rédigé particulièrement pour les utilisateurs débutant l'analyse des données et ne possédant pas de connaissances avancées en statistiques et en informatique. Bien que certains exemples d'analyses statistiques vous soient présentés pour illustrer le fonctionnement du logiciel, ce guide ne passera pas en revue chacune des méthodes d'analyse statistique. De même, il n'a pas pour objectif de vous familiariser avec l'interprétation des résultats d'analyse.

Afin de maîtriser le contenu de ce guide, nous vous conseillons de vous asseoir face à un ordinateur et de suivre les étapes directement sur le logiciel SPSS. Les illustrations présentées sont directement tirées de l'écran du logiciel et devraient correspondre au contenu de votre écran. Une connaissance minimale de l'environnement Windows est nécessaire à la compréhension de ce guide. Si vous utilisez Windows pour la première fois, consultez l'aide du système d'exploitation (disponible dans le menu **DÉMARRER** du système d'exploitation Windows) avant d'utiliser le logiciel SPSS.

Comment se procurer le logiciel SPSS ?

Le logiciel SPSS est développé par la compagnie *SPSS Corporation* en 1968 et actuellement utilisé par près de 250 000 personnes. Il est disponible en versions Windows (9X, ME, 2000 et XP), Macintosh et UNIX, dans les langues anglaise et française. Pour plus de détails, vous pouvez consulter le site américain de cette compagnie aux références à la fin du guide.

Le logiciel SPSS n'est pas un produit unique. Il est composé d'une série de modules ajoutés au module de base **SPSS BASE** pour augmenter les fonctionnalités de SPSS. Ces modules peuvent répondre à des besoins très variés, comme l'interface de saisie, la préparation (gestion des données manquantes, analyses de puissance), l'analyse (l'ensemble des techniques statistiques, les analyses confirmatoires, l'analyse qualitative, les réseaux de neurones), la présentation (production de rapports, scénarios de décision) et la publication des données dans différentes bases de données et sur l'Internet.

Il est possible d'acheter directement de la compagnie les principaux modules d'analyse statistique, soit **SPSS BASE**, **REGRESSION STATS** et **ADVANCED STATS** au prix « académique » d'environ \$2400. Toutefois, les étudiants impliqués dans un laboratoire de la Faculté des Sciences Sociales, de la Faculté des Sciences de l'Éducation ou de la Faculté des Sciences Administratives sont encouragés à commander une licence institutionnelle à prix modique pour une durée limitée (3 à 5 ans). Notez bien que la licence est généralement réservée aux ordinateurs enregistrés à l'Université Laval, soit les ordinateurs des professeurs, du personnel de soutien, des professionnels de recherche, des laboratoires et des programmes de support institutionnels (e.g., Ulysse). Pour de plus amples informations concernant les licences, vous pouvez contacter le responsable informatique de votre département ou faculté. Pour connaître cette personne ressource, consultez la rubrique Assistance en informatique à l'adresse Internet suivante :

<http://www.sit.ulaval.ca/public/dsd/information/personnel/>

Les étudiants qui désirent installer SPSS sur un ordinateur personnel peuvent se procurer une version étudiante complète du logiciel, le **SPSS 11 GRADUATE PACK**. Cette version est en vente à la librairie universitaire Zone Laval au coût d'environ \$315 + taxes. Elle contient les trois principaux modules du logiciel SPSS cités au paragraphe précédent. Pour de plus amples informations, vous pouvez contacter la section informatique de la librairie Zone Université.

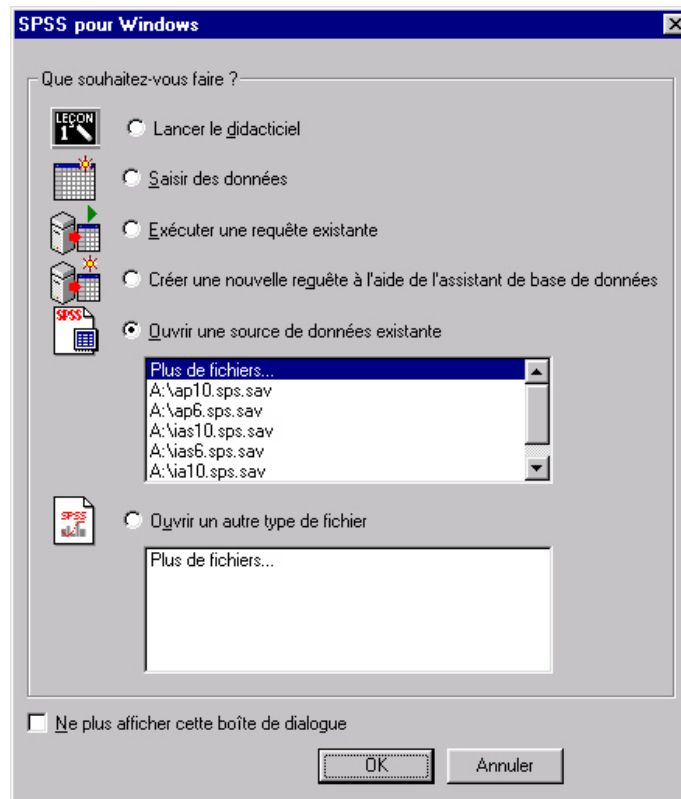
2. PRÉSENTATION DU LOGICIEL SPSS

Le démarrage du logiciel

Pour démarrer le logiciel SPSS, repérez l'icône suivant sur le bureau de votre poste de travail. Double-cliquez sur celui-ci à l'aide de votre souris d'ordinateur¹.



À l'ouverture du logiciel, une fenêtre apparaît automatiquement.



¹ Si l'icône SPSS n'apparaît pas sur l'écran de votre ordinateur, cliquez sur **DÉMARRER** de la barre de tâches Windows puis déplacez-vous à **PROGRAMMES**. Si SPSS n'y apparaît pas, repérez l'endroit où le logiciel est installé, ouvrez le répertoire d'installation et double-cliquez sur le fichier « spsswin.exe ».

Cette fenêtre vous permet de choisir ce que vous désirez faire parmi cinq options de lancement : **LANCER LE DIDACTICIEL**, **SAISIR DES DONNÉES**, **EXÉCUTER UNE REQUÊTE EXISTANTE**, **CRÉER UNE NOUVELLE REQUÊTE** ou encore **OUVRIR UNE SOURCE DE DONNÉES EXISTANTE**, tel qu'un fichier de données ou un fichier de résultats SPSS. Si vous ouvrez SPSS pour la première fois et que vous désirez simplement explorer son interface, choisissez **SAISIR DES DONNÉES**.

L'ouverture automatique de cette fenêtre au démarrage du logiciel peut être désactivée en tout temps en cochant la case apparaissant dans la partie inférieure gauche. Si cette fenêtre ne s'affiche pas lors du démarrage de SPSS, c'est qu'elle a tout simplement été désactivée par un autre utilisateur.

Avant de commencer l'exploration du logiciel SPSS, il importe d'aborder certaines notions importantes sur son mode de fonctionnement général.

Les principales fenêtres du logiciel

Le logiciel SPSS interagit avec l'utilisateur principalement à partir de deux fenêtres : la fenêtre **ÉDITEUR DE DONNÉES** et la fenêtre **NAVIGATEUR SPSS**. Chacune de ces fenêtres possède sa propre barre de menus².

L'**ÉDITEUR DE DONNÉES** est la fenêtre qui s'affiche en arrière plan lorsque vous démarrez une session SPSS (voir Figure 2.1). Les données que vous souhaitez analyser seront affichées dans cette fenêtre. La section suivante présente en détails les différentes fonctions et menus de cette fenêtre.

² Les menus **ANALYSE** et **GRAPHES** sont disponibles dans les deux fenêtres du logiciel afin de pouvoir effectuer une analyse ou un graphique sans avoir à retourner dans l'**ÉDITEUR DE DONNÉES**

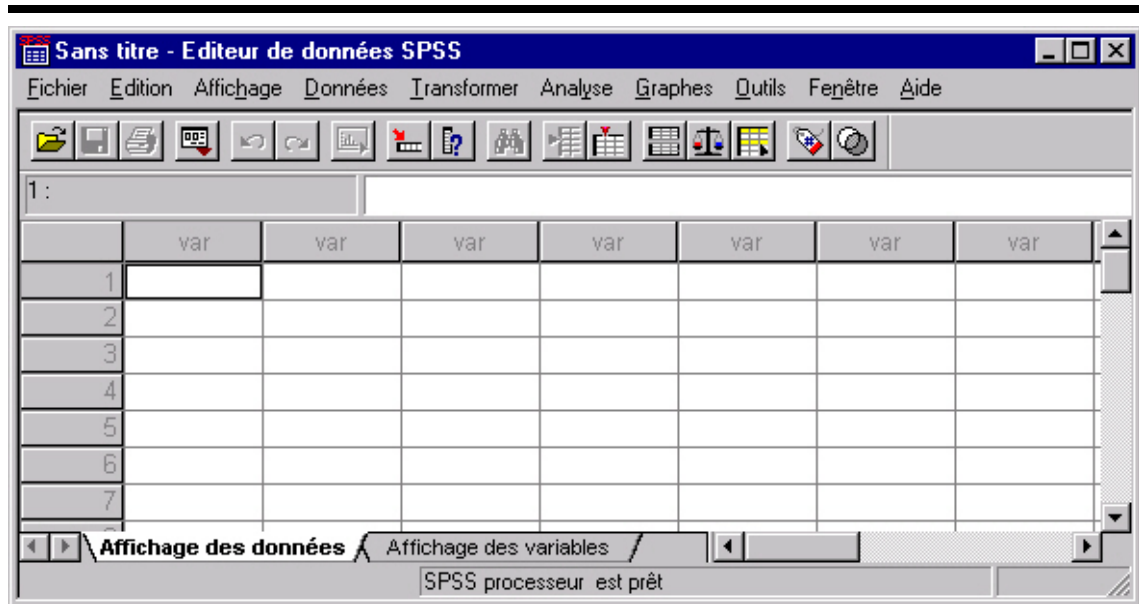


Figure 2.1. Éditeur de données

Le **NAVIGATEUR SPSS** s'ouvre automatiquement lorsque vous réalisez une procédure qui génère des résultats (voir Figure 2.2). Ainsi, que ce soit les résultats d'une analyse statistique ou encore un graphique que vous désirez créer, le résultat final s'affiche dans cette fenêtre. Notez que le **NAVIGATEUR SPSS** est également appelé « Fenêtre de résultats » ou « Fenêtre de sortie (*output*) ». La section intitulée « Navigateur SPSS » présente en détails les fonctionnalités de cette fenêtre.

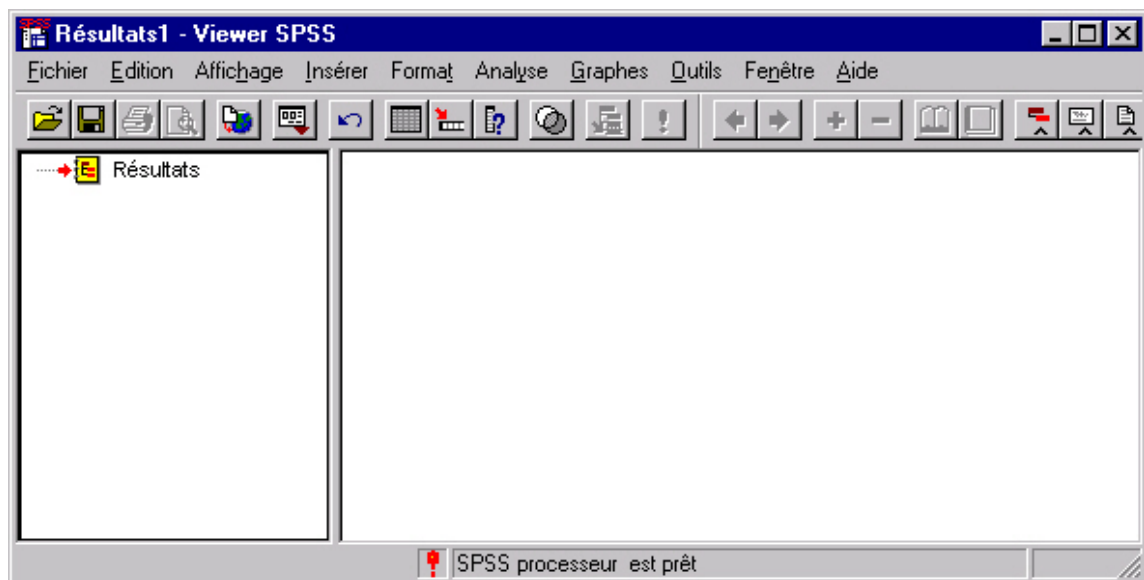
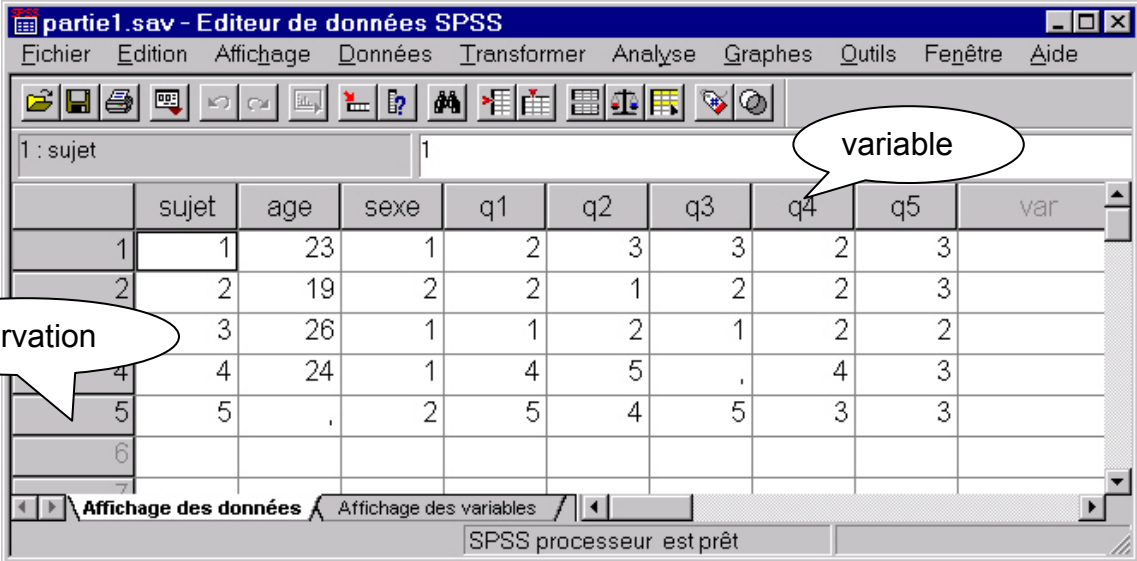


Figure 2.2. Navigateur SPSS

3. L'ÉDITEUR DE DONNÉES SPSS

Description générale

L'ÉDITEUR DE DONNÉES est la fenêtre de travail du logiciel SPSS. C'est dans cette fenêtre que les données sont saisies et affichées. Chacune des lignes de cette fenêtre correspond à une observation, tandis que chaque colonne représente une variable du fichier de données (voir Figure 3.1).



The screenshot shows the SPSS Data Editor window titled 'partie1.sav - Editeur de données SPSS'. The menu bar includes Fichier, Edition, Affichage, Données, Transformer, Analyse, Graphes, Outils, Fenêtre, and Aide. The toolbar contains various icons for file operations and data manipulation. The data grid has columns labeled 'sujet', 'age', 'sexe', 'q1', 'q2', 'q3', 'q4', 'q5', and 'var'. The first column is labeled '1 : sujet' in the header. The first row is labeled '1' in the header. A callout bubble labeled 'observation' points to the first row, and another callout bubble labeled 'variable' points to the 'q4' column. The status bar at the bottom indicates 'Affichage des données' and 'SPSS processeur est prêt'.

	sujet	age	sexe	q1	q2	q3	q4	q5	var
1	1	23	1	2	3	3	2	3	
2	2	19	2	2	1	2	2	3	
3	3	26	1	1	2	1	2	2	
4	4	24	1	4	5	.	4	3	
5	5	.	2	5	4	5	3	3	
6									
7									

Figure 3.1. Fichier de données SPSS

Dans cet exemple, vous pouvez apercevoir huit variables différentes, soit le numéro du participant (*sujet*), son âge (*age*), son sexe (*sexe*) et ses résultats à un inventaire de stress (*q1*–*q5*). Chaque ligne représente les résultats d'un participant. Ainsi, le participant #1 est un homme de 23 ans qui a répondu 3 à la question #2.

Les menus et les sous-menus de l'Éditeur de données

La fenêtre ÉDITEUR DE DONNÉES comporte dix principaux menus. Chacun de ces menus contient différentes fonctions du logiciel SPSS reliées à la saisie et la gestion de données. Le Tableau 3.1 présente les menus de l'ÉDITEUR DE DONNÉES avec une description des principales fonctions offertes par chacun d'eux.

Tableau 3.1. Menus et fonctions de l'Éditeur de données

Menus	Fonctions
FICHIER	Contient les commandes reliées à la gestion des fichiers (données, syntaxe, script) : créer un nouveau fichier, ouvrir un fichier existant, enregistrer un fichier, imprimer un fichier, quitter SPSS.
ÉDITION	Contient les fonctions permettant de faire l'édition des travaux sous SPSS (opérations de traitement de texte) : annuler la dernière action effectuée, rétablir l'action annulée, copier/coller des observations ou des variables, rechercher une expression, modifier les options du logiciel.
AFFICHAGE	Comprend les options permettant de définir les paramètres d'affichage du logiciel : afficher ou masquer différents menus, modifier les options d'affichage du logiciel (polices utilisées, etc.).
DONNÉES	Contient les fonctions traitant de la gestion du fichier de données : insérer une variable ou une observation, trier les données, fusionner, agréger ou scinder des fichiers, restructurer une variable en observation et vice-versa.
TRANSFORMER	Contient les fonctions permettant de transformer les variables d'un fichier de données : créer une nouvelle variable à partir d'un calcul, générer des nombres aléatoires, recoder une variable, trier les observations, remplacer les données manquantes.
ANALYSE	Contient toutes les procédures d'analyses statistiques pouvant être effectuées par SPSS.
GRAPHES	Contient les fonctions de présentation graphique des données.
OUTILS	Permet d'obtenir des informations sur les variables du fichier de données, de définir et utiliser des groupes de variable, d'exécuter un script et de personnaliser les menus.
FENÊTRE	Affiche les fenêtres ouvertes de SPSS. Permet de gérer les fenêtres du logiciel (e.g., passer d'une fenêtre à l'autre).
AIDE	Contient les rubriques d'aide du logiciel, le didacticiel, les manuels de syntaxe fournis avec le logiciel et l'assistant statistique.

Le parcours des différents menus

L'accès aux menus et fenêtres du logiciel se fait à l'aide de la souris ou du clavier d'ordinateur. Pour fureter dans les menus du logiciel à l'aide de la souris, il suffit simplement de cliquer sur le menu désiré pour faire apparaître les commandes ou sous-commandes disponibles. Si une commande apparaît en gris pâle, cela indique qu'elle n'est pas disponible lors de l'opération effectuée. Elle apparaît à nouveau en noir lorsqu'elle est applicable.

L'utilisateur qui préfère utiliser le clavier au lieu de la souris pour naviguer à l'intérieur du logiciel doit tenir enfoncée la touche [ALT] de son clavier, suivie de la lettre correspondant au menu désiré. La lettre associée à un menu est celle qui apparaît soulignée dans la barre des menus. Par exemple, pour accéder aux fonctions du menu **F**ICHIER, la lettre [F] doit être utilisée. Ensuite, pour accéder aux sous menus, vous pouvez utiliser la lettre correspondante ou encore les flèches de direction du clavier.

4. LA GESTION DES DONNÉES

La gestion des données constitue généralement la première étape reliée à l'utilisation du logiciel SPSS et probablement l'étape la plus importante d'une démarche statistique puisqu'elle détermine la façon dont les données pourront être analysées ultérieurement. Chacune des fonctions reliées à la gestion des données doit donc être effectuée de manière rigoureuse. Que se soit dans la saisie, l'ouverture d'un fichier, l'importation ou encore la transformation des données, vous devez vous assurer de bien vérifier vos résultats à chacune des étapes. Plus ces étapes auront été réalisées de manière rigoureuse, plus vos résultats d'analyse seront fiables et valides. Ainsi, même si le logiciel possède des fonctions automatisées permettant de transformer rapidement vos données, il est important de vérifier *manuellement* les résultats de vos transformations (i.e., regarder les résultats pour quelques observations) afin de vous assurer que ce qui a été fait par le logiciel correspond à vos attentes. Rappelez-vous que l'utilisation d'un ordinateur et d'un logiciel statistique a pour unique but de vous faciliter certaines tâches et non de remplacer la démarche rigoureuse du chercheur.

Cette section du guide aborde en cinq parties les procédures du logiciel SPSS reliées à la saisie, la manipulation et la transformation des données. Les trois premières parties présentent les façons d'obtenir des données dans SPSS, soit (a) la saisie des données directement dans la fenêtre **ÉDITEUR DE DONNÉES**, (b) l'ouverture ou le rappel d'un fichier de données SPSS déjà construit et (c) l'importation ou le transfert des données saisies dans un autre logiciel (e.g., Excel, Access, Filemaker, Lotus, dBASE, Word). La quatrième partie illustre la manipulation des données alors que la dernière partie présente les procédures reliées à la transformation des données, telle que la gestion des données manquantes, la recodification d'une variable et le calcul d'une somme.

La saisie directe des données

Tel que mentionné précédemment, la saisie des données d'une recherche s'effectue dans l'**ÉDITEUR DE DONNÉES**. Si vos données n'ont pas encore été saisies, il est préférable de compléter l'entrée de données directement dans SPSS afin d'éviter certains problèmes de conversion ou d'importation de données.

Après l'ouverture du logiciel, choisissez la commande **ENTRER DES DONNÉES** dans les options de lancement ou encore sélectionnez le menu **FICHIER**, puis **NOUVEAU** puis le sous-menu **DONNÉES**. Vous obtiendrez alors une feuille de données vide prête pour la saisie.

Les trois étapes de la saisie des données

Définir les variables. La première étape consiste à définir les propriétés de chacune des variables pour lesquelles vous désirez saisir des informations. Ceci nous amène à introduire une particularité de la fenêtre **ÉDITEUR DE DONNÉES** que nous avons volontairement évitée d'aborder jusqu'ici. Au bas de la fenêtre, vous apercevrez l'onglet **AFFICHAGE DES VARIABLES**³. Cet onglet vous permet d'accéder à une sous-fenêtre de l'**ÉDITEUR DE DONNÉES** où vous pourrez définir différents paramètres des variables que vous désirez saisir (voir Figure 4.1) :

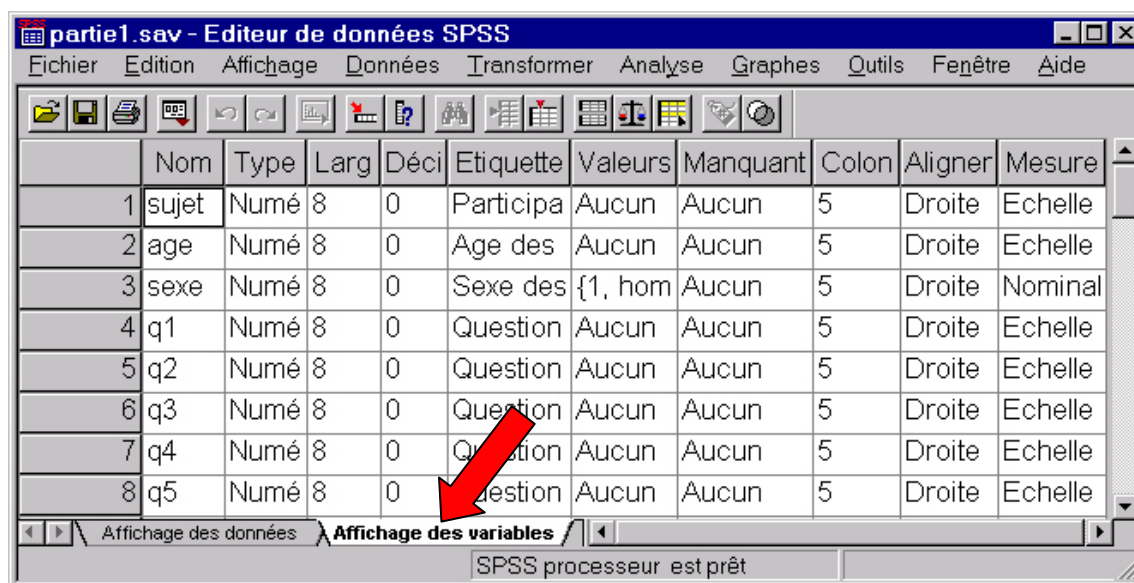


Figure 4.1. Affichage des variables

le nom de la variable : le nom de vos variables est celui qui apparaîtra dans la partie supérieure d'une colonne de données. Celui-ci doit être d'une longueur maximale de huit caractères (chiffres et lettres).

³ La feuille **AFFICHAGE DES VARIABLES** constitue une nouveauté depuis la version 10 de SPSS. Pour les utilisateurs d'une version antérieure du logiciel, les fonctions associées à cette nouvelle fenêtre se retrouvent dans le menu **DONNÉES** de la barre de menus.

le type de variable : le type de variable permet de définir le format de vos données (e.g., numérique, chaîne, dollar, date).

la largeur de la variable : cette option permet de choisir la longueur d'une variable du fichier de données, qui détermine le nombre de caractères exportés lors d'une sauvegarde du fichier de données en format texte.

le nombre de décimales de la variable : lorsque les données d'une variable sont sous forme de chiffres, il est possible de définir le nombre de décimales à afficher après la virgule.

l'étiquette de la variable : l'étiquette d'une variable permet de préciser son contenu par un énoncé pouvant être supérieur à huit caractères. Cette fonction est notamment utile pour préciser les variables pour lesquelles le nom révèle peu d'informations sur son contenu. Par exemple, pour une variable appelée `IDB1`, il peut être utile de produire l'étiquette `Inventaire de Dépression de Beck, question 1`, ce qui fournira à l'utilisateur des informations pertinentes sur la nature des variables. Ainsi, si vous devez utiliser votre fichier de données plusieurs années après la réalisation de votre étude, il vous sera possible de connaître précisément la nature de chaque variable.

les étiquettes de valeur : les étiquettes de valeur permettent d'attribuer une information qualitative à une valeur saisie dans les champs de votre variable. Ceci est notamment utile lorsque la personne codifie ses données nominales en chiffres. Par exemple, il est d'usage de codifier la variable `sexe` en chiffre (e.g., 1 et 2) plutôt que d'inscrire `homme` ou `femme` à répétition. À l'aide de la fonction **ÉTIQUETTES DE VALEUR**, vous pouvez préciser que « 1 » représente les hommes et « 2 » représente les femmes.

les valeurs manquantes : cette option permet de définir certaines valeurs d'une variable comme valeurs manquantes. Par exemple, vous pouvez distinguer les données manquantes associées à un refus de répondre de la part d'un participant et les données manquantes obtenues parce que la question ne s'applique pas au participant. Les valeurs définies comme manquantes sont repérées par un indicateur en vue d'un traitement spécial et sont exclues de la plupart des analyses.

la largeur de la colonne : cette option permet de définir la largeur d'affichage d'une colonne dans l'**ÉDITEUR DE DONNÉES**.

la fonction d'alignement : l'option **ALIGNER** vous permet de choisir l'alignement des données d'une variable (à gauche, à droite ou centré).

le type de mesure : Finalement, l'option **MESURE** permet de choisir le type d'échelle de mesure associée aux données d'une variable (i.e., nominale, ordinale ou échelle).

Saisir les données. Après avoir défini les propriétés de chacune des variables pour lesquelles vous désirez saisir de l'information, vous passez à la deuxième étape de la saisie qui consiste à revenir dans la fenêtre **AFFICHAGE DES DONNÉES** et à saisir *directement* les données dans le chiffrier. Rappelons que chacune des colonnes représente une variable et que chacune des rangées représente une observation (i.e., un participant). La fenêtre de saisie des données est semblable à l'écran d'une feuille d'un chiffrier électronique, tel que Excel, Lotus ou Quattro Pro. À cette étape, vous remarquerez que le nom des variables que vous avez spécifié précédemment apparaît dans la partie supérieure de chacune des colonnes. Pour faire apparaître l'étiquette de votre variable, déplacez le curseur de votre souris sur le nom de votre variable et attendez une seconde. Pour saisir vos données, entrez chaque information relative à une variable dans la case correspondante. Utilisez les flèches de déplacement de votre clavier pour passer d'une case à une autre.

Sauvegarder les informations saisies. Finalement, la troisième étape consiste à *sauvegarder* les informations saisies précédemment. Choisissez le menu **FICHIER**, puis **ENREGISTRER SOUS**. Entrez un nom pour votre fichier puis cliquez sur le bouton **SAUVEGARDER**. Vous noterez que le fichier est sauvegardé en format *.sav, soit un fichier de données SPSS.

Comment obtenir un sommaire des variables ?

Nous terminons cette partie sur la saisie directe des données dans SPSS en introduisant deux fonctions du menu **OUTILS** de l'**ÉDITEUR DE DONNÉES**, reliées directement à la première étape de saisie. Il est possible d'obtenir un sommaire des informations concernant une ou plusieurs variables grâce aux commandes **VARIABLES** et **INFORMATION SUR LE FICHIER**. La commande **VARIABLES** permet d'obtenir un résumé des informations sur une ou plusieurs variables de votre fichier de données que vous choisissez à l'aide d'une fenêtre de dialogue. Par exemple, vous pouvez savoir à l'aide de cette fonction que la variable `sexe` est une variable nominale dont l'étiquette correspond à `sexe` du participant. La commande **INFORMATION SUR LE FICHIER** permet d'obtenir des informations sur le fichier de données entier ainsi que sur chacune de ses variables. Lorsque vous choisissez cette option, le **NAVIGATEUR SPSS** est activé afin d'afficher les informations sur votre fichier. Cette commande est très utile notamment lorsque vous désirez obtenir un résumé des informations qui définissent les variables de votre fichier, également connu sous le nom de « codeur de données ».

L'ouverture d'un fichier SPSS

Pour ouvrir un fichier SPSS déjà construit, choisissez la commande **OUVRI**R du menu **FICHIER** puis sélectionnez le type de fichier que vous désirez ouvrir, soit un fichier de données (*.sav), un fichier de syntaxe (*.sps), un fichier de résultats (*.spo) ou tout autre type de fichier supporté par le logiciel (voir Figure 4.2). À l'ouverture de la fenêtre de dialogue, spécifiez l'endroit où se trouve votre fichier, sélectionnez-le puis cliquez sur le bouton **OUVRI**R.

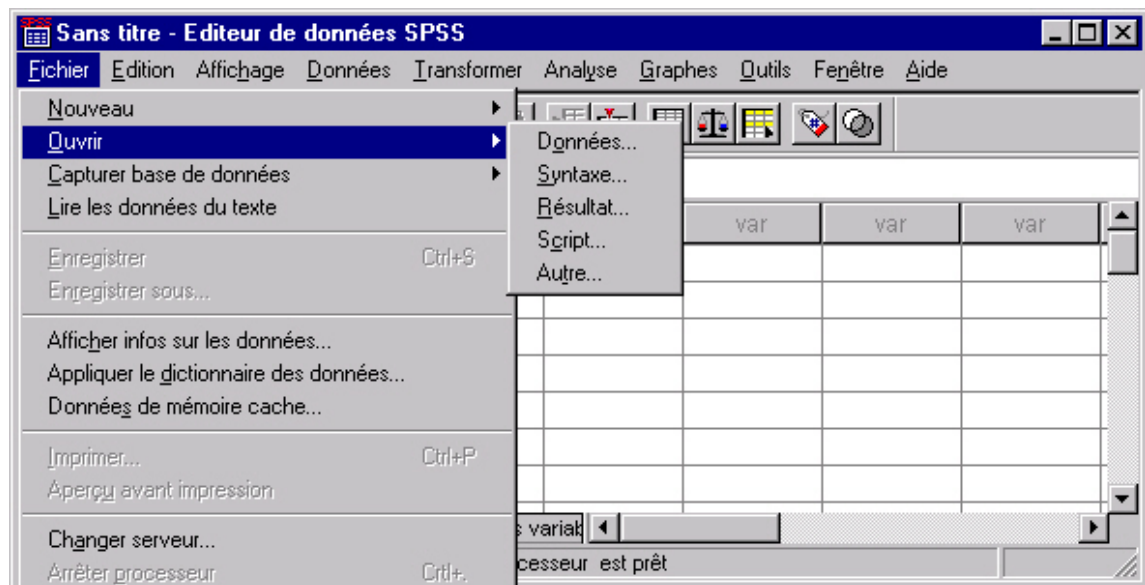


Figure 4.2. Ouverture d'un fichier existant

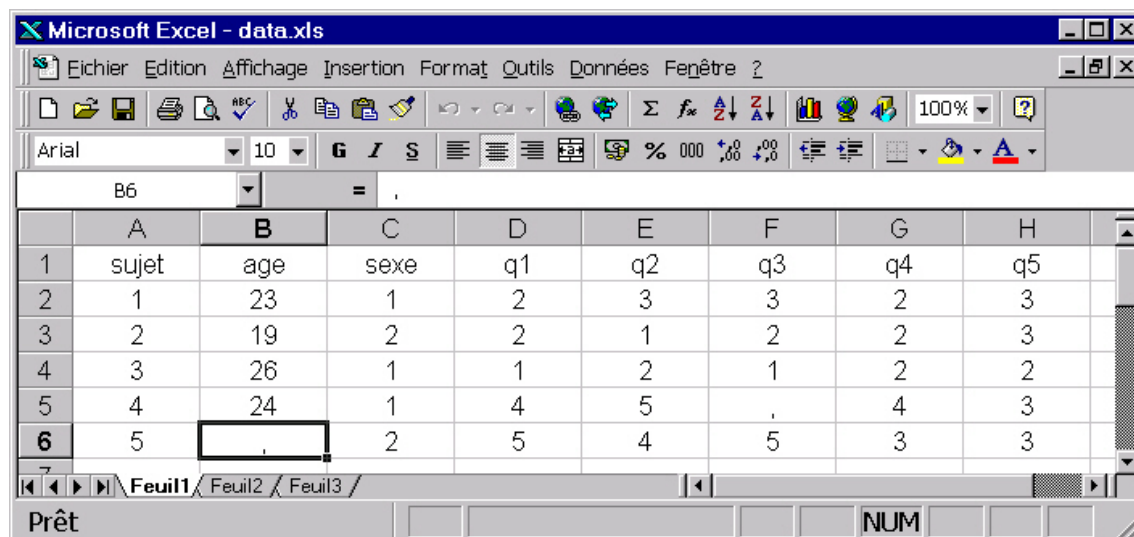
L'importation des données dans SPSS

Il est fréquent que les données à analyser dans SPSS aient été saisies dans un autre logiciel, tel qu'un gestionnaire de bases de données (e.g., Access, Paradox, Filemaker), un chiffrier électronique (e.g., dbase, Excel, Quattro Pro), un autre logiciel d'analyse statistique (e.g., SAS) ou un logiciel de traitement de texte (e.g., Word, WordPerfect). Le logiciel SPSS vous permet d'importer et de convertir facilement tous vos fichiers de données afin d'en faire l'analyse.

Le transfert peut s'effectuer selon deux procédures. Le choix de la procédure à utiliser dépend de la nature du fichier de données à importer. La première procédure concerne les fichiers de données créés à l'aide d'un gestionnaire de bases de données ou un chiffrier électronique. La deuxième procédure permet

d'importer un fichier de texte. Ces deux procédures se retrouvent dans le menu **FICHIER** de SPSS.

Transfert d'un fichier provenant d'un chiffrier électronique. Cette procédure décrit les étapes à effectuer pour lire et importer les données saisies dans un chiffrier électronique de type Excel (voir Figure 4.3).



	A	B	C	D	E	F	G	H
1	sujet	age	sexe	q1	q2	q3	q4	q5
2	1	23	1	2	3	3	2	3
3	2	19	2	2	1	2	2	3
4	3	26	1	1	2	1	2	2
5	4	24	1	4	5	.	4	3
6	5	.	2	5	4	5	3	3

Figure 4.3. Données brutes dans le logiciel Excel

Étape #1. Entrez dans le menu **FICHIER** du logiciel SPSS puis choisissez l'option **OUVRI** suivie de **DONNÉES** (voir Figure 4.4).

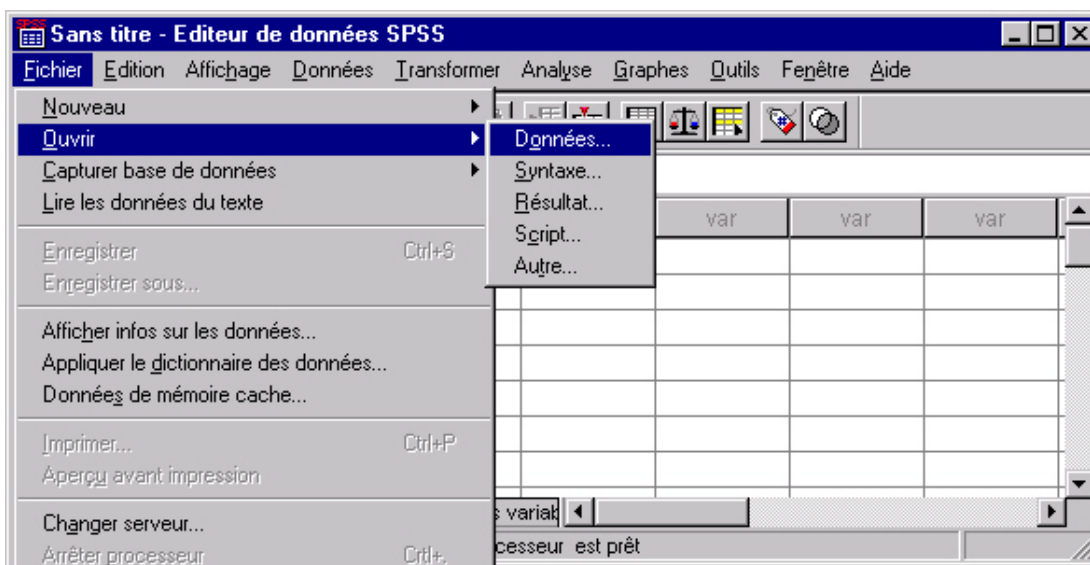


Figure 4.4. Ouvrir un fichier

Étape #2. Sélectionnez le type de fichier que vous désirez transférer dans SPSS à l'aide du menu se situant au bas de la fenêtre de dialogue (voir Figure 4.5).

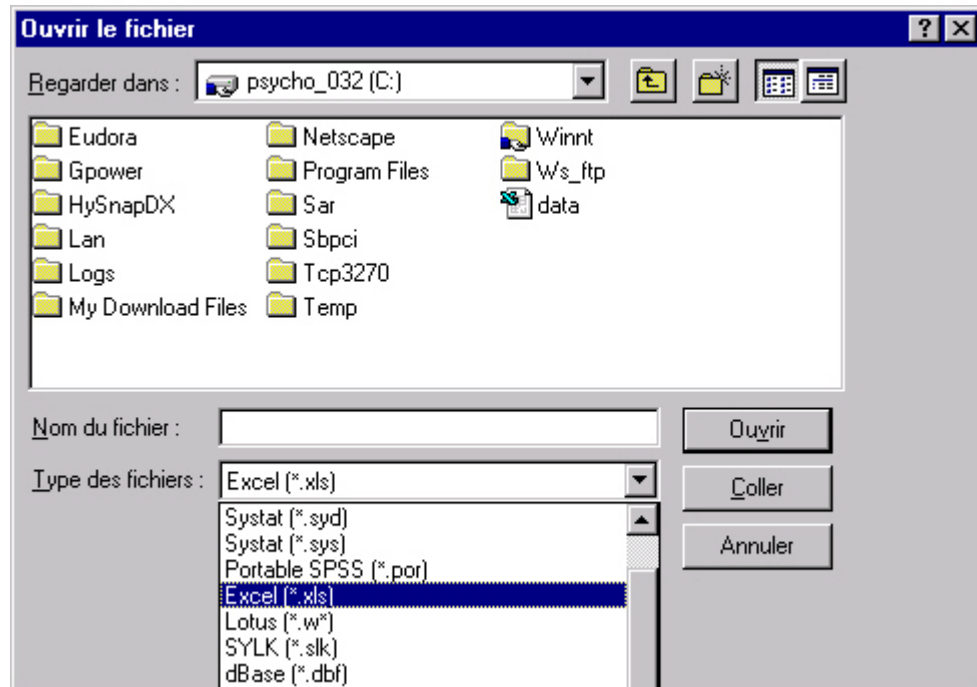


Figure 4.5. Sélectionner le type de fichier (selon l'extension)

Étape #3. Sélectionnez le fichier que vous désirez transférer. Si celui-ci ne se retrouve pas dans la liste des fichiers disponibles, explorez le contenu de votre ordinateur à l'aide de la fonction **REGARDER DANS** apparaissant au haut de la fenêtre de dialogue afin de repérer le bon fichier. Cliquez ensuite sur **OUVRIR**. Si vous obtenez un message de violation de partage de fichier, vous devez fermer ce fichier dans le logiciel Excel avant de réaliser l'importation dans SPSS.

Étape #4. Si la première ligne de votre fichier contient le nom des variables, cochez la case **LIRE LE NOM DES VARIABLES À PARTIR DE LA PREMIÈRE LIGNE DES DONNÉES**. Dans le cas de fichiers Excel récents (version 5 et plus), choisissez la feuille de calcul à utiliser de même que la plage de données à capturer (voir Figure 4.6). Par défaut, SPSS sélectionne toutes les données de la première feuille du chiffrier.

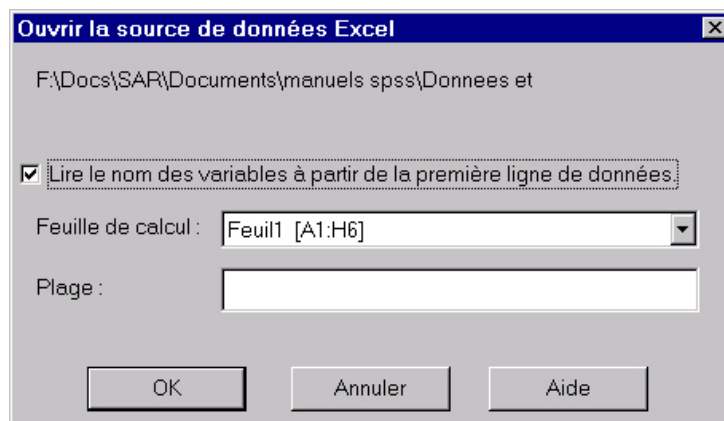


Figure 4.6. Sélection de la plage de données du fichier Excel

Cliquez ensuite sur **OK** pour confirmer le transfert. Vous obtiendrez alors votre fichier de données SPSS (voir Figure 4.7). N'oubliez pas de *sauvegarder* ce fichier à l'aide de la commande **FICHIER / ENREGISTRER SOUS**.

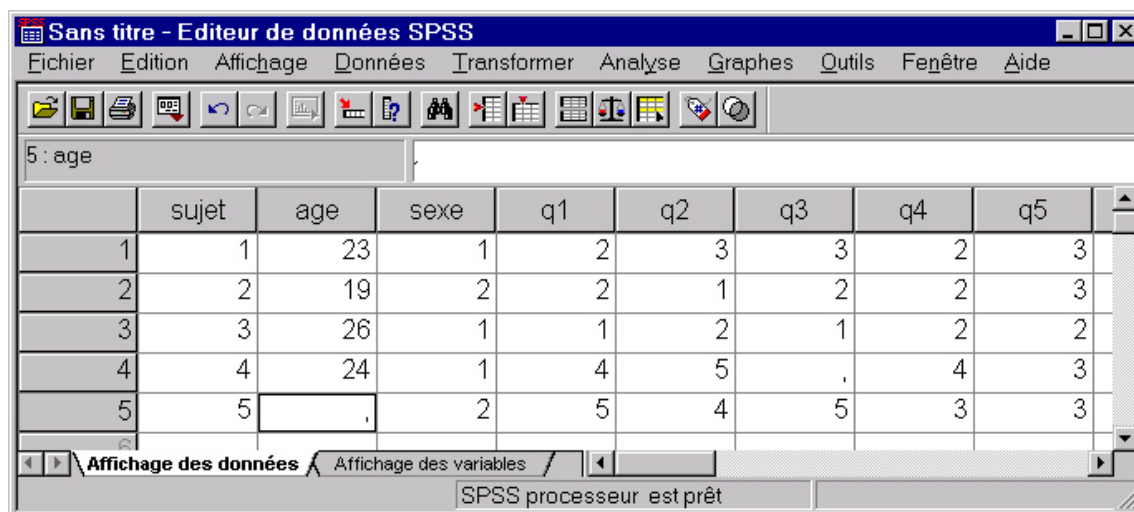


Figure 4.7. Données importées dans le logiciel SPSS

Notez bien que si une cellule contient des données en format mixte (e.g., 500\$), elles sont lues comme une chaîne de données (i.e., fichiers Excel 5.0 ou plus). Pour les versions antérieures du logiciel Excel, les données sont transformées au format du type de la première valeur des données. Les autres types sont transformés en données manquantes. Si le nom d'une variable est remplacé par un numéro, c'est qu'il contient plus de huit caractères ou encore des caractères non acceptés par le logiciel SPSS.

Le transfert des fichiers en format texte. La procédure suivante illustre les étapes nécessaires à la lecture d'un fichier de texte brut. Il est essentiel de bien connaître le fichier de données avant de procéder au transfert, afin de pouvoir identifier clairement la nature et la position de chacune des variables.

Les données de l'exemple précédent (voir Figure 4.3) sont utilisées afin d'illustrer les étapes du transfert de fichier en format texte. Elles proviennent d'un fichier texte (`data.txt`) illustré à la Figure 4.8. Chacune des étapes d'importation doit être effectuée avec soin pour éviter toute perte ou distorsion des données. Notez bien que vous pourrez sauvegarder vos données seulement après avoir complété toutes les étapes.

```
sujet;age;sexe;q1;q2;q3;q4;q5  
1;23;1;2;3;3;2;3  
2;19;2;2;1;2;2;3  
3;26;1;1;2;1;2;2  
4;24;1;4;5;.;4;3  
5;.;2;5;4;5;3;3
```

Figure 4.8. Données en format texte (séparateur point virgule)

Étape #1. Dans le menu **FICHIER** de SPSS, sélectionnez la commande **LIRE LES DONNÉES DU TEXTE**.

Étape #2. Sélectionnez ensuite le fichier en format texte que vous désirez transférer. Si celui-ci ne se retrouve pas dans la liste des fichiers disponibles, explorez le contenu de votre ordinateur à l'aide de la fonction **REGARDER DANS** apparaissant au haut de la fenêtre de dialogue pour repérer le bon fichier. Cliquez ensuite sur **OUVRI**R.

Étape #3. L'**ASSISTANT D'IMPORTATION DE TEXTE** devrait apparaître à cette étape pour vous guider dans le transfert d'un fichier texte. La partie inférieure de l'**ASSISTANT D'IMPORTATION** présente le contenu de votre fichier de données (voir Figure 4.9). Vous devez indiquer si votre texte correspond à un format prédéfini. Si vous avez déjà ouvert ce type de fichier avec SPSS et que vous avez enregistré sa configuration auparavant, vous pouvez utiliser le format prédéfini. Si c'est la première fois que vous ouvrez votre fichier, cochez **NON**. Cliquez ensuite sur **SUIVANT**.

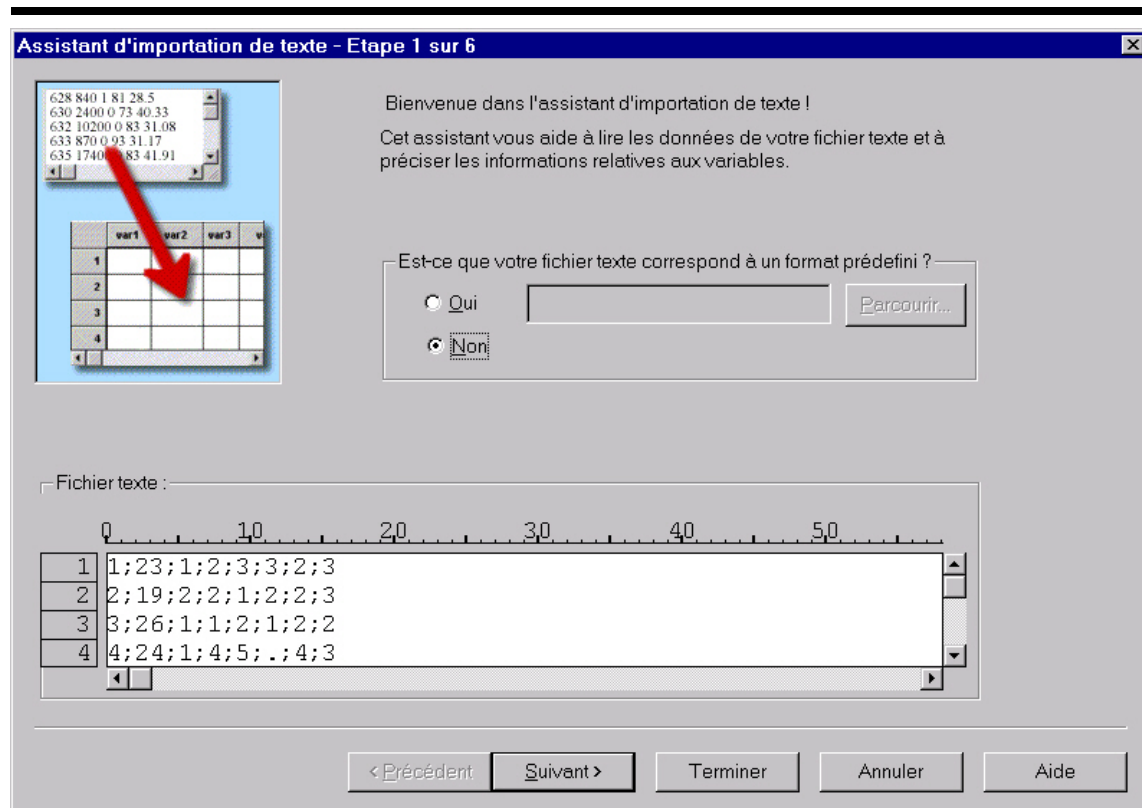


Figure 4.9. Assistant d'importation des données en format texte

Étape #4. Vous devez maintenant indiquer si vos variables sont disposées selon un format *délimité* ou *fixe* (voir Figure 4.10). Le format *délimité* s'applique si vos données sont séparées par un caractère particulier comme une virgule, une tabulation ou un point virgule. Le format *fixe* s'applique lorsque les données sont clairement définies par des colonnes et qu'elles se retrouvent toujours à la même position pour chaque observation. Les variables de l'exemple sont en format *délimité* puisqu'elles sont séparées par un point virgule. Une fois le format choisi, précisez si le nom de vos variables apparaît dans la partie supérieure de vos données et cliquez sur **SUIVANT**.

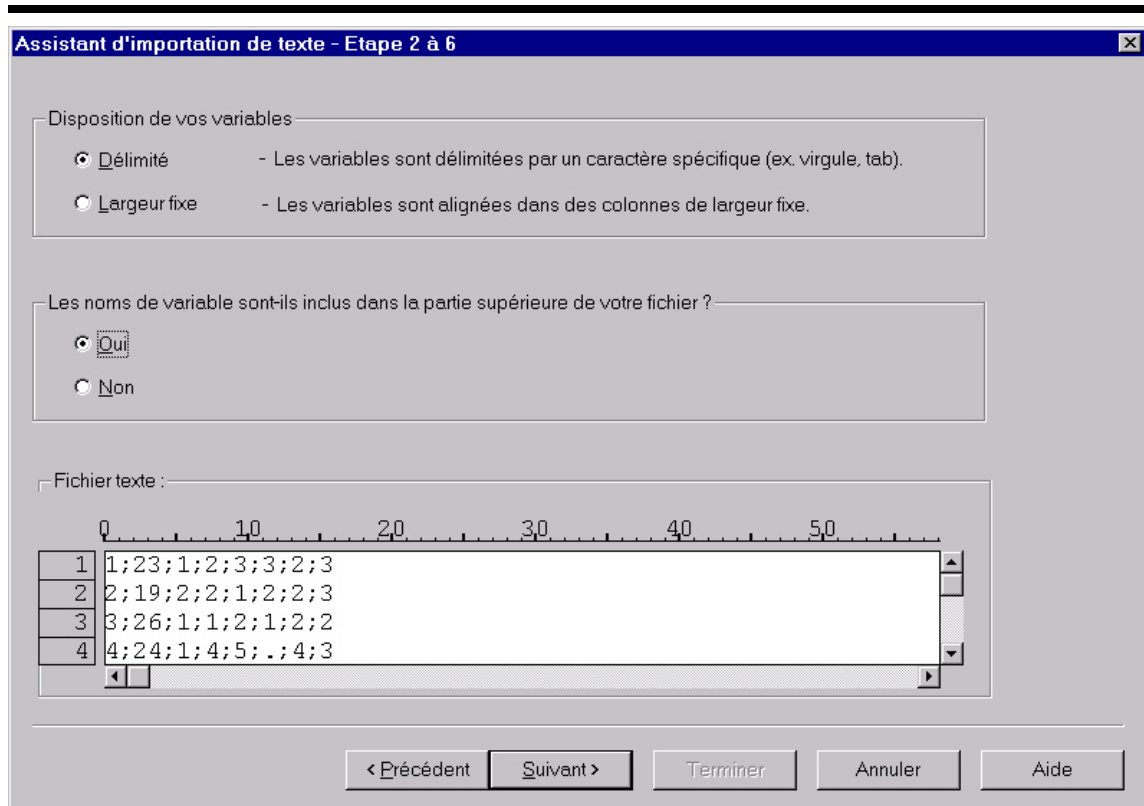


Figure 4.10. Spécification du format des variables

Étape #5. À cette étape, vous devez préciser si vos observations commencent à la première ligne de votre fichier, de même que le nombre de lignes par observation (voir Figure 4.11). Si une ligne vide sépare chaque observation, vous devez l'inclure dans le nombre de lignes par observation. Après avoir indiqué si vous désirez importer toutes vos observations ou seulement une partie, cliquez sur **SUIVANT**.

Étape #6. Cette étape consiste à préciser le type de séparateur dans le fichier de données (voir Figure 4.12). Le séparateur est automatiquement identifié par SPSS pour les données en format délimité mais vous pouvez toujours le modifier, selon vos choix. Dans le cas d'un fichier en format fixe, vous devez indiquer la position de chaque variable dans le fichier et ce, pour **chacune** des lignes d'une même observation. Pour se faire, insérez une ligne verticale entre chacune des variables. Cliquez ensuite sur **SUIVANT**.

Assistant d'importation de texte - Etape délimitée 3 sur 6

La première observation de données commence à la ligne...

Représentation de vos observations

☒ Chaque ligne représente une observation

☐ Un numéro de variables représente une observation :

Combien d'observations souhaitez-vous importer ?

☒ Toutes

☐ La première

☐ Un pourcentage aléatoire (approximatif) : %

Aperçu des données

	0	10	20	30	40	50
1	1;23;1;2;3;3;2;3					
2	2;19;2;2;1;2;2;3					
3	3;26;1;1;2;1;2;2					

< Précédent Suivant > Terminer Annuler Aide

Figure 4.11. Spécification de la taille de chaque observation

Assistant d'importation de texte - Etape délimitée 4 sur 6

Quels séparateurs s'affichent entre les variables ?

☐ Tabulation ☐ Espace

☐ Virgule ☒ Point-virgule

☐ Autre :

Quel est le qualificateur de texte ?

☒ Aucun

☐ Apostrophe

☐ Guillemet

☐ Autre :

Aperçu des données

	0	10	20	30	40	50
1	1;23;1;2;3;3;2;3					
2	2;19;2;2;1;2;2;3					
3	3;26;1;1;2;1;2;2					

< Précédent Suivant > Terminer Annuler Aide

Figure 4.12. Spécification du séparateur

Étape #7. Cette étape consiste à identifier le nom et le format de chaque variable (voir Figure 4.13). Dans cet exemple, vous n'avez pas à spécifier cette information car elle est automatiquement récupérée par SPSS suite à l'examen de la première ligne du fichier texte et des données. Toutefois, pour des données en format *fixe*, vous devez généralement spécifier les noms et formats. Si vous souhaitez ne pas importer certaines variables de vos données, choisissez **NE PAS IMPORTER** dans la zone **FORMAT DE DONNÉES**.

Assistant d'importation de texte - Etape 5 sur 6

Spécifications pour les variables sélectionnées dans l'aperçu des données

Nom de variable : sujet Nom d'origine : sujet

Format de données : Numérique

Aperçu des données

	sujet	age	sexe	q1	q2	q3	q4
1		23	1	2	3	3	2
2		19	2	2	1	2	2
3		26	1	1	2	1	2
4		24	1	4	5	.	4

< Précédent Suivant > Terminer Annuler Aide

Figure 4.13. Spécification du nom et du format des variables

Étape #8. Vous en êtes maintenant à la dernière étape du transfert. Choisissez si vous désirez sauvegarder le format de fichier, afin de pouvoir l'utiliser lors d'un futur transfert de données similaires. Vous pouvez également sauvegarder la syntaxe du transfert. Cliquez ensuite sur **TERMINER** pour finaliser l'opération (voir Figure 4.14). Votre fichier de données SPSS s'affichera enfin à l'écran (voir Figure 4.15). N'oubliez pas de sauvegarder ce fichier, grâce à la commande **FICHIER / ENREGISTRER SOUS**, pour éviter de devoir recommencer le transfert après chaque utilisation des données.

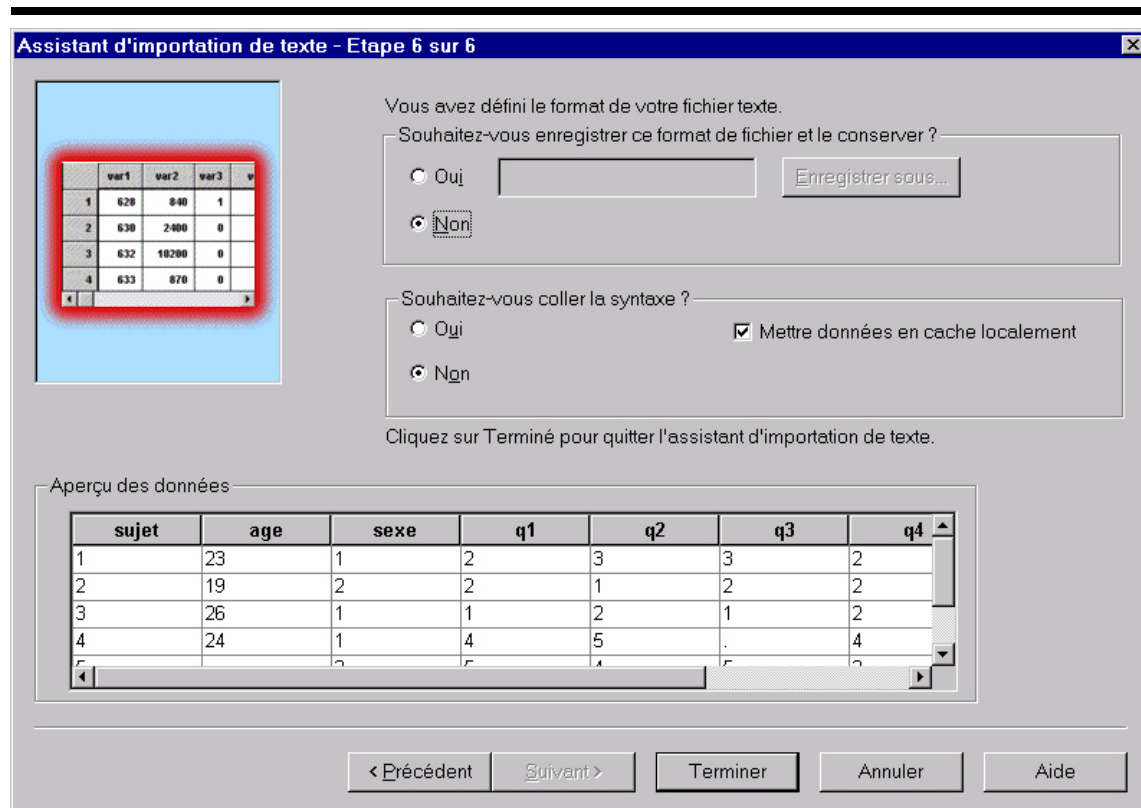


Figure 4.14. Finaliser le transfert des données en format texte

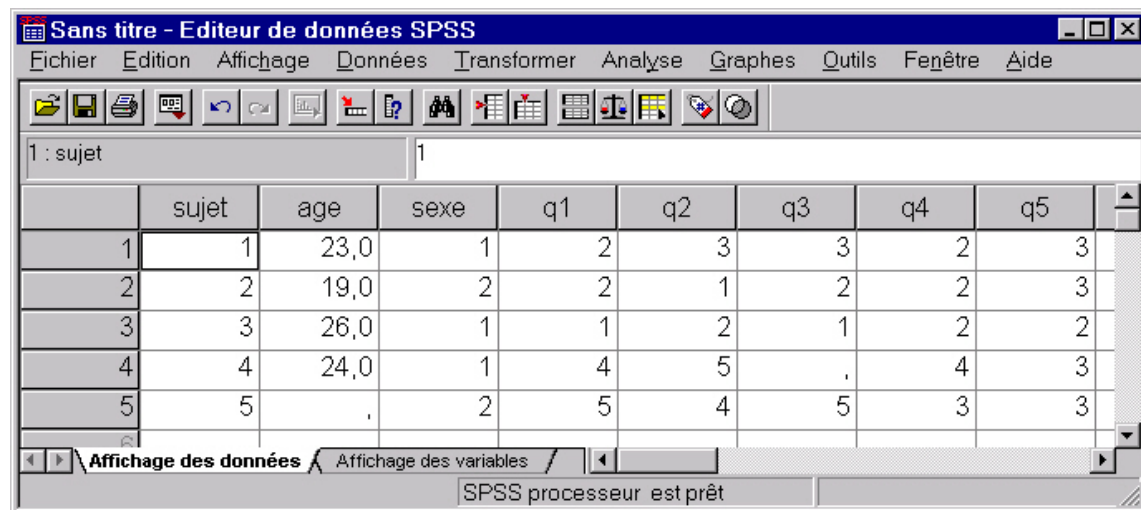


Figure 4.15. Fichier de données SPSS obtenu grâce à l'Assistant d'importation

La manipulation des données

Le menu **DONNÉES** permet la manipulation des données. Cette étape de la gestion de données comprend un vaste ensemble d'opérations, dont l'ajout de variables, l'ajout d'observations, le tri et la sélection d'observations. Le Tableau 4.1 présente en détails les commandes disponibles dans le menu **DONNÉES**.

Tableau 4.1. Les principales commandes du menu DONNÉES

Commande	Fonction(s)
INSÉRER UNE VARIABLE	Insérer une variable dans l' ÉDITEUR DE DONNÉES
INSÉRER UNE OBSERVATION	Insérer une observation dans l' ÉDITEUR DE DONNÉES
TRIER LES OBSERVATIONS	Trier les observations en ordre croissant ou décroissant
RESTRUCTURER...	Transformer des observations en variables et vice-versa
FUSIONNER DES FICHIERS	Ajouter des observations ou des variables provenant d'un autre fichier SPSS
SCINDER UN FICHIER	Séparer les résultats ou l'affichage des analyses selon une variable
SÉLECTIONNER DES OBSERVATIONS	Sélectionner un sous-échantillon des observations selon un critère spécifique
PONDÉRER DES OBSERVATIONS	Pondérer les observations selon un critère

Ajout d'observations et de variables provenant d'autres fichiers SPSS.

Deux situations se produisent régulièrement en recherche expérimentale. La première correspond à une étude où les mêmes variables sont récoltées dans différents sites d'expérimentation, puis saisies dans différents fichiers de données. La seconde situation réfère à une étude où les mêmes participants sont soumis à plusieurs mesures, réparties dans divers fichiers. Généralement, l'analyste est intéressé à regrouper les données pour pouvoir les traiter. Dans le premier cas, l'analyste désire ajouter des observations alors que le second cas illustre l'ajout de variables. La Figure 4.16 schématise l'ajout d'observations.

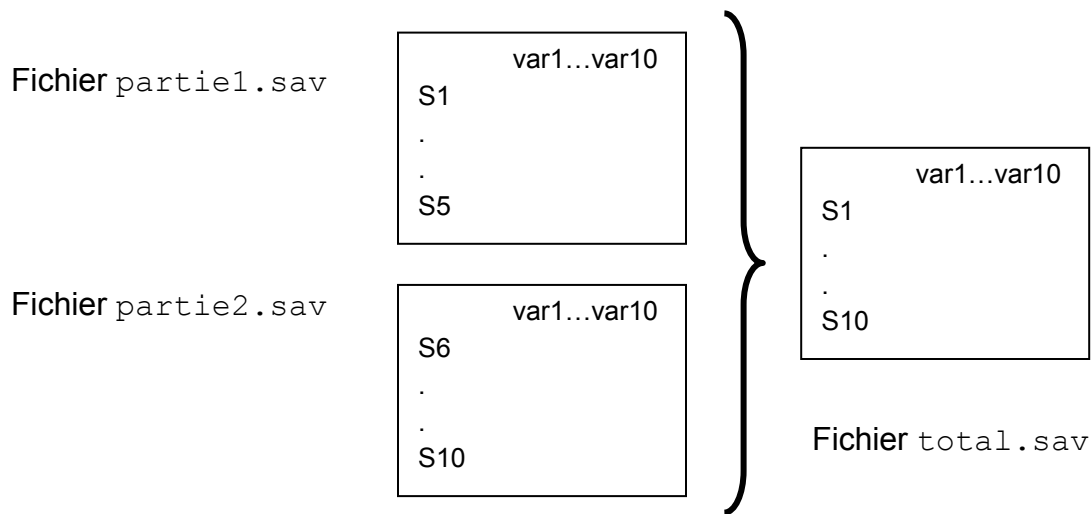


Figure 4.16. Fusion de fichiers par ajout d'observations

Les numéros d'observations étant différents, on peut constater que la fusion des fichiers est réalisée par pairage des variables communes `var1` à `var10`. Si un des fichiers ne possède pas d'information sur une ou plusieurs variables, des données manquantes seront attribuées pour chacune de ces variables.

La Figure 4.17 schématise l'ajout de variables. Ici, comme les noms de variables sont différents, la fusion des fichiers est réalisée par pairage des observations communes `s1` à `s10`. Si un des fichiers ne possède pas d'information sur une ou plusieurs observations, des données manquantes seront attribuées pour chaque variable associée à ces observations.

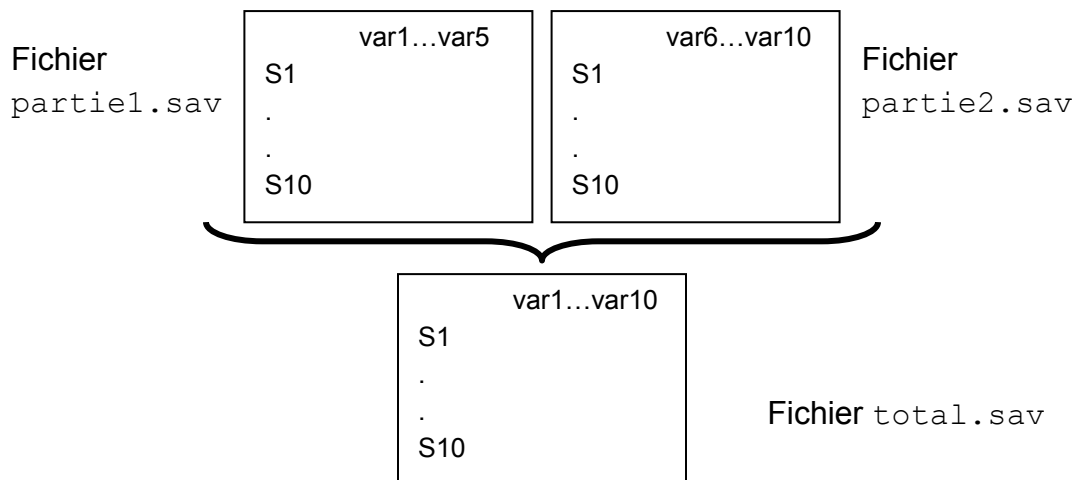


Figure 4.17. Fusion de fichiers par ajout de variables

La commande **FUSIONNER DES FICHIERS**, du menu **DONNÉES**, permet de réaliser ces deux opérations.

Exemple d'ajout d'observations avec SPSS

Etape #1. Ouvrez le fichier correspondant à la première portion de vos données.

Etape #2. Entrez dans le menu **DONNÉES**, choisissez l'option **FUSIONNER DES FICHIERS** suivie de **AJOUTER DES OBSERVATIONS**. SPSS vous demande de sélectionner le fichier contenant les observations que vous désirez ajouter. Cliquez ensuite sur **OUVRIR**. Vous obtiendrez une fenêtre de dialogue similaire à la Figure 4.18.

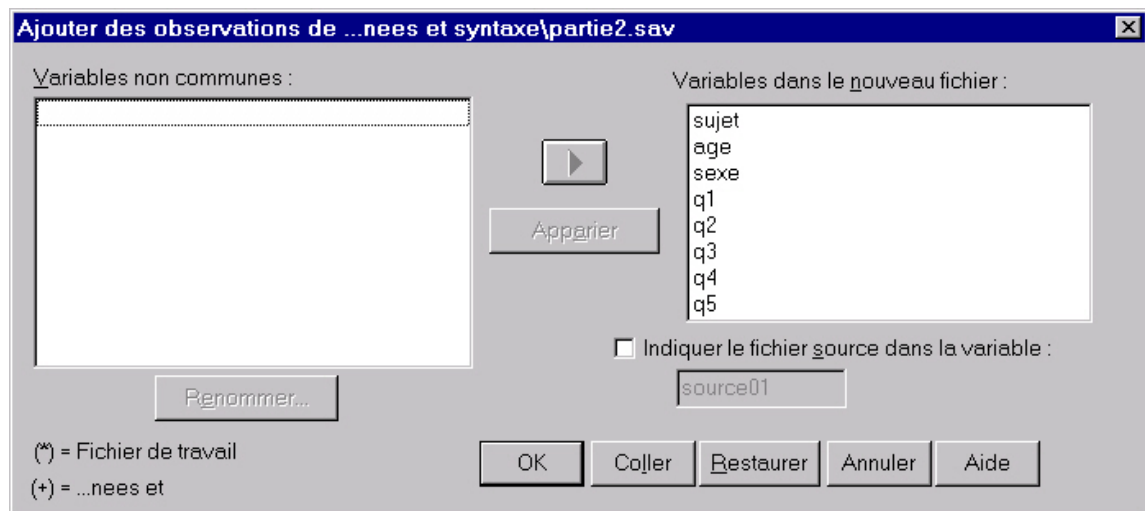


Figure 4.18. Sélection des variables pour ajout d'observations

Dans l'exemple actuel, les deux fichiers contiennent exactement les mêmes variables (sujet, sexe, age, q1 à q5). Ceci explique pourquoi la fenêtre **VARIABLES NON-COMMUNES** est vide. Deux situations peuvent expliquer que certaines variables soient listées dans cette fenêtre : (a) une variable est présente dans un fichier mais absente dans le second, et (b) deux variables identiques sont nommées différemment (e.g., *nosujet* et *sujet*). Dans le premier cas, vous pouvez forcer l'introduction des variables non-communes grâce à la flèche centrale. Le deuxième cas demande de sélectionner les deux variables similaires dans la fenêtre **VARIABLES NON-COMMUNES** (e.g., *nosujet* et *sujet*) puis de cliquer sur le bouton **APPARIER**. Ceci indiquera à SPSS que ces variables portant des noms différents correspondent en fait à la même

information. Notez que c'est le nom de la variable provenant du premier fichier ouvert qui sera imposé dans le fichier fusionné.

Etape #3. Cliquez sur le bouton **OK**. Les observations du second fichier seront ajoutées dans la partie inférieure du premier fichier. Si vous désirez conserver les deux fichiers originaux, il vous suffit d'enregistrer le nouveau fichier fusionné sous un nom différent grâce à la commande **FICHIER / ENREGISTRER SOUS**.

Exemple d'ajout de variables avec SPSS

À la section précédente, nous avons sauvegardé le fichier résultant de la fusion des deux ensembles de cinq observations sous le nom de `data stress.sav`. Nous désirons maintenant y ajouter la variable `note` incluse dans le fichier `notes.sav` qui contient également le numéro de sujet (`sujet`) et le sexe du participant (`sexe`).

Etape #1. Ouvrez le fichier correspondant à la première portion de vos données (e.g., le fichier `data stress.sav`).

Etape #2. Entrez dans le menu **DONNÉES**, choisissez l'option **FUSIONNER DES FICHIERS** suivie de **AJOUTER DES VARIABLES**. Sélectionnez le fichier contenant les variables que vous désirez ajouter puis cliquez sur **OUVRIR**. Vous obtiendrez une fenêtre de dialogue similaire à la Figure 4.19.

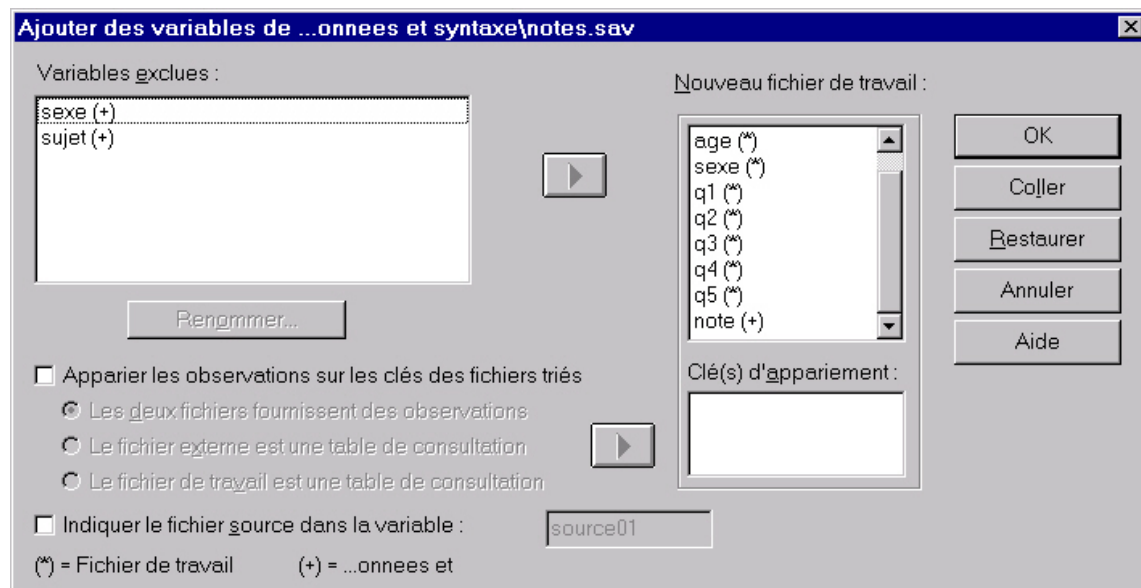


Figure 4.19. Ajout de variables

La fenêtre **NOUVEAU FICHIER DE TRAVAIL** présente la liste des variables qui seront incluses dans le fichier fusionné. Celles-ci peuvent être distinguées selon le fichier d'origine, soit (*) pour les variables provenant du fichier original (i.e., le fichier présent dans l'**ÉDITEUR DE DONNÉES**) et (+) pour les variables provenant du fichier contenant les variables à ajouter (i.e., fichier ouvert avec la commande **FUSIONNER DES FICHIERS**). Vous pouvez constater que la variable `note` provient effectivement du fichier de variables à ajouter. La fenêtre **VARIABLES EXCLUES** présente la liste des variables qui ne seront pas conservées dans le nouveau fichier fusionné. Cette situation provient du fait que des variables possédant le même nom se retrouvent dans les deux fichiers (i.e., `sujet` et `sexe`). Dans ce cas, seules les variables présentes dans le fichier original (i.e., variable avec un (*)) seront conservées.

Etape #3. Il est fortement recommandé de préciser une variable commune aux deux fichiers qui permet à SPSS d'apparier correctement les observations dans deux fichiers (e.g., les données de l'observation #4 doivent être associées avec les données ajoutées de la même observation). Cette variable, appelée « clé », doit être identifiée en cochant l'option **APPARIER LES OBSERVATIONS SUR LES CLÉS DES FICHIERS TRIÉS**. Sélectionnez ensuite la variable « clé » dans la fenêtre **VARIABLES EXCLUES** puis cliquez sur la flèche située à gauche de la fenêtre **CLÉ(S) D'APPARIEMENT**. Dans le présent exemple, la variable clé étant `sujet`, nous avons obtenu la fenêtre de dialogue illustrée à la Figure 4.20.

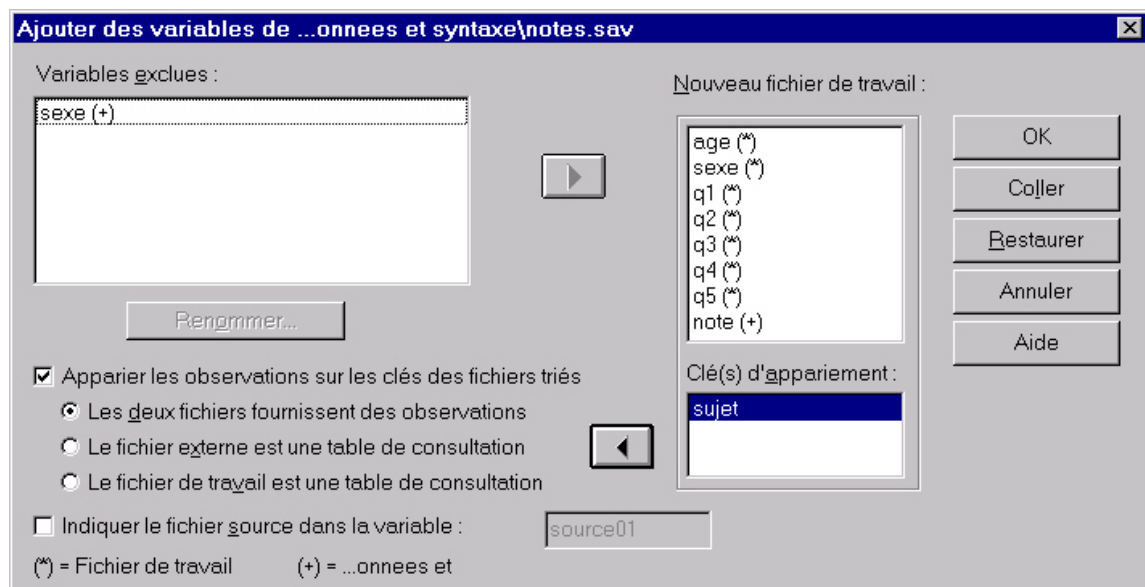


Figure 4.20. Spécification de la clé d'appariement

Étape #4. Cliquez sur le bouton **OK**. Vous obtiendrez un avertissement de SPSS spécifiant que « l'appariement par clé échouera si les données ne sont pas triées dans l'ordre croissant des clés d'appariement ». Notez bien que cette opération doit être réalisée **avant** de débiter la fusion des fichiers. Pour se faire, sélectionnez la commande **TRIER LES OBSERVATIONS** du menu **DONNÉES** et triez les observations de vos deux fichiers selon la variable «clé». Les variables du second fichier seront ajoutées à la droite du premier fichier. Si vous désirez conserver les deux fichiers originaux, il vous suffit d'enregistrer le nouveau fichier fusionné sous un nom différent grâce à la commande **FICHIER / ENREGISTRER SOUS**.

Sélection d'un sous-échantillon d'observations. Un chercheur peut parfois être intéressé à analyser seulement une portion des observations disponibles. Par exemple, il peut vouloir restreindre ses analyses uniquement aux participants masculins, aux participants âgés de plus de 25 ans, aux 50 premiers participants ou à un sous-échantillon aléatoire comprenant 60% des participants originaux. La commande **SÉLECTIONNER LES OBSERVATIONS** du menu **DONNÉES** permet de réaliser cette sélection de participants (voir Figure 4.21).

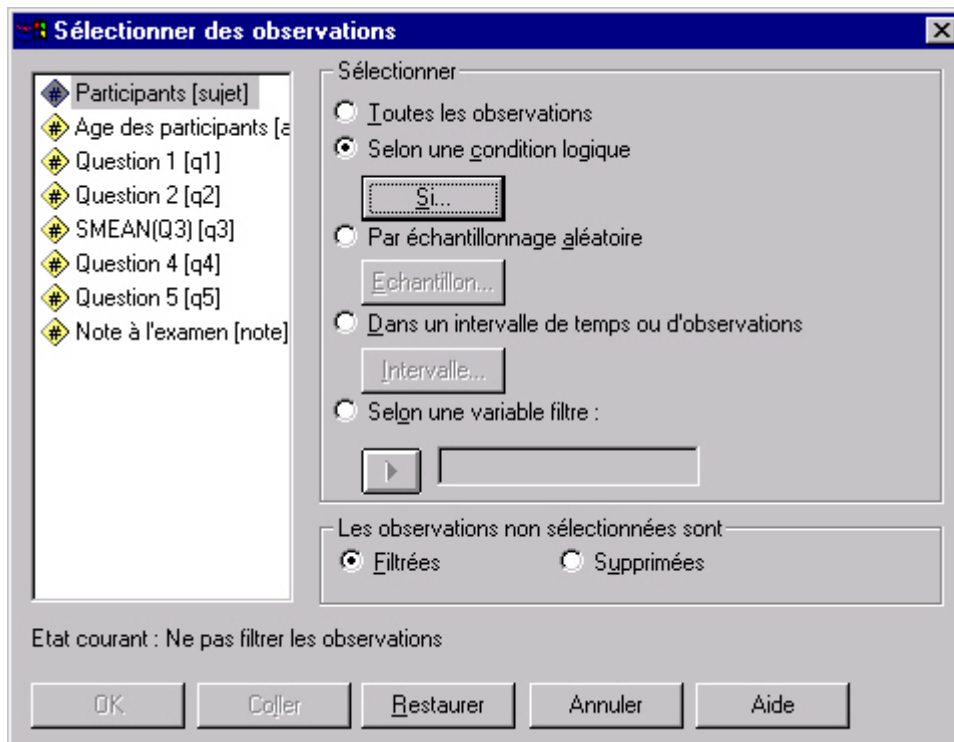


Figure 4.21. Sélectionner des observations

La partie gauche de la fenêtre présente la liste des variables disponibles alors que la partie droite illustre les principaux critères pour sélectionner les observations. Par exemple, l'utilisateur qui désire restreindre les analyses aux participants âgés de 25 ans et plus sélectionne le critère **SELON UNE CONDITION LOGIQUE**. Une fenêtre de dialogue s'ouvre et permet de préciser la nature de condition, soit `age >= 25` (voir Figure 4.22). Après avoir appuyé sur **POURSUIVRE**, l'utilisateur est retourné à la fenêtre principale, où la condition est maintenant précisée. La dernière instruction consiste à préciser, dans la partie inférieure droite, la procédure applicable aux données non-sélectionnées, soit le retrait temporaire (**FILTRÉES**) ou l'effacement (**SUPPRIMÉES**). Il est recommandé de sélectionner le retrait temporaire.

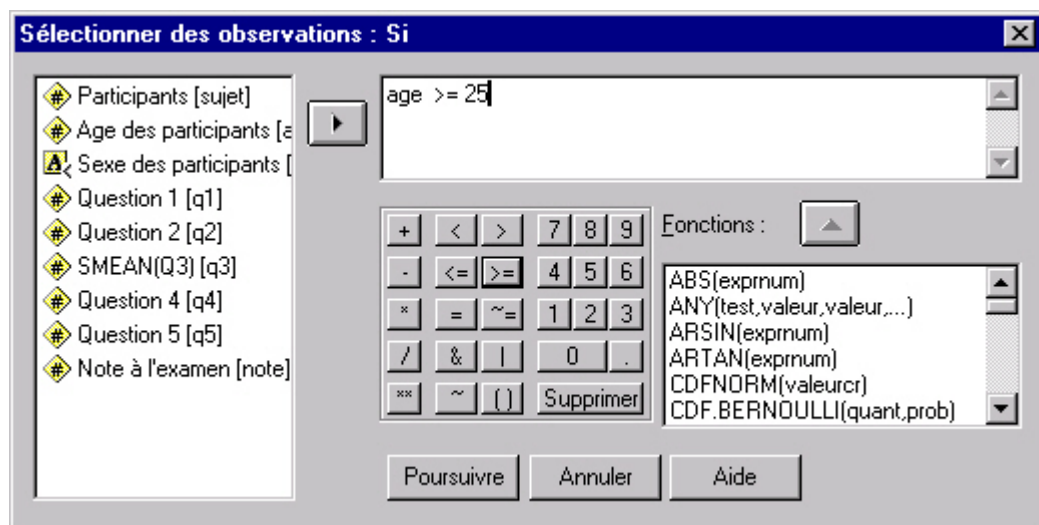


Figure 4.22. Sélectionner des observations – condition logique

Après avoir cliqué sur **OK**, un examen de l'**ÉDITEUR DE DONNÉES** permet de constater que le numéro des observations non-sélectionnées est maintenant barré, indiquant ainsi le statut d'exclusion temporaire des analyses (voir Figure 4.23). De plus, une nouvelle variable, `FILTER_$`, prenant la valeur « 0 » pour les observations retirées et « 1 » pour les observations sélectionnées, a été créée. Pour sélectionner à nouveau l'ensemble des observations, l'utilisateur doit retourner dans la commande **SÉLECTIONNER DES OBSERVATIONS** et cocher l'option **TOUTES LES OBSERVATIONS** dans la fenêtre supérieure droite.

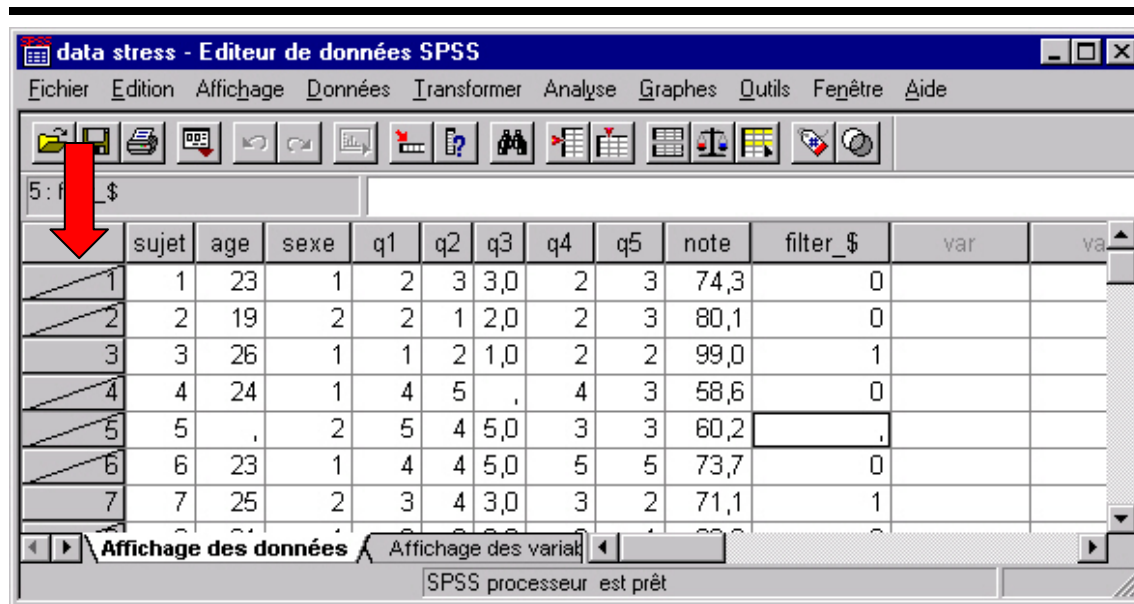


Figure 4.23. Sélectionner des observations

Production d'analyses séparées pour une variable. Une situation rencontrée fréquemment dans l'analyse de données est la production de résultats séparés selon les différentes valeurs d'une variable. Par exemple, un chercheur désire produire une matrice de corrélations entre les cinq questions (e.g., q1-q5) de façon séparée pour les hommes et les femmes. Toutefois, plusieurs commandes d'analyse statistique, dont la commande pour les corrélations, ne permettent pas de spécifier cette séparation. Le logiciel SPSS met à la disposition de l'analyste la commande **SCINDER UN FICHIER** pour permettre de spécifier une *séparation* temporaire des données selon les valeurs d'une variable (voir Figure 4.24).

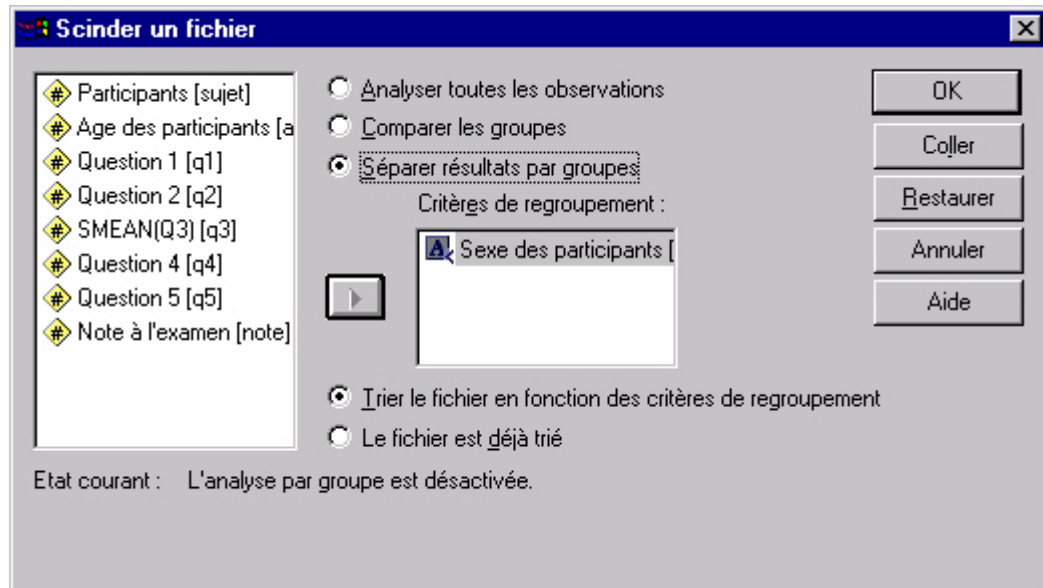


Figure 4.24. Scinder un fichier selon une variable

Cette fenêtre de dialogue comprend, dans sa partie gauche, une liste des variables susceptibles de constituer une variable de séparation, et dans sa partie droite, trois méthodes de séparation des observations. La première méthode **ANALYSER TOUTES LES OBSERVATIONS** est en fait une option pour annuler une précédente séparation. La seconde méthode **COMPARER LES GROUPES** permet de produire une analyse à la fois, selon toutes les valeurs de la variable choisie. Par exemple, pour la variable `sexe`, cette méthode permet de produire chaque analyse pour les deux sexes. La troisième méthode **SÉPARER LES RÉSULTATS PAR GROUPE** permet de produire toutes les analyses, selon une valeur à la fois de la variable choisie. Par exemple, cette méthode produit toutes les analyses pour les hommes, puis toutes les analyses pour les femmes.

La transformation des données

Cette étape de la gestion de données comprend un vaste ensemble d'opérations, dont la création, la définition, la production et la recodification de nouvelles données à partir des données existantes. Le Tableau 4.2 présente en détails les commandes disponibles dans le menu **TRANSFORMER**.

Tableau 4.2. Principales commandes du menu TRANSFORMER

Commande	Fonction(s)
CALCULER	Créer une nouvelle variable ou remplacer une variable existante à partir d'un calcul effectué sur une ou plusieurs variables existantes
GÉNÉRER DES NOMBRES ALÉATOIRES	Produire des nombres aléatoires dans un intervalle spécifique
RANGER LES OBSERVATIONS	Convertir les valeurs des observations en rangs selon leur position
RECODER	Recoder une variable déjà existante (changement d'échelle de mesure)
REPLACER LES VALEURS MANQUANTES	Remplacer les données manquantes par la moyenne ou la projection linéaire

Il est important de noter que toute transformation des données modifie de façon **permanente** les données illustrées dans l'**ÉDITEUR DE DONNÉES**. Il est donc recommandé de réaliser systématiquement une copie de sauvegarde de vos données brutes afin de prévenir une perte irréversible de données.

Gestion des données manquantes. Les données manquantes peuvent être gérées de diverses façons. La première approche consiste à conserver la donnée manquante, sans tenter de l'estimer. Concrètement, vous pouvez donc laisser une case vide dans l'**ÉDITEUR DE DONNÉES**, à l'endroit où se situe la donnée manquante. Les analyses seront produites seulement sur les données disponibles. Toutefois, cette stratégie a comme conséquence de diminuer significativement le nombre d'observations disponibles pour certains types d'analyses, ce qui réduit la puissance statistique des tests d'hypothèses.

La seconde approche consiste à estimer la donnée manquante en la remplaçant par la moyenne ou la médiane de la variable, ou une extrapolation des données situées près de la donnée manquante. Notez qu'il est **essentiel** d'effectuer la gestion des données manquantes avant la transformation des données (e.g., calcul d'une somme ou d'une moyenne). Nous recommandons de réaliser cette opération directement après la fusion de fichiers.

La commande **REPLACER LES VALEURS MANQUANTES**, située dans le menu **TRANSFORMER**, permet d'effectuer l'estimation des données manquantes à l'aide des méthodes décrites précédemment. Vous pouvez constater que la fenêtre de dialogue comprend deux parties (voir Figure 4.25). La partie gauche présente la liste des variables disponibles, alors que la partie droite permet de préciser la méthode d'estimation ainsi que le nom de la nouvelle variable, soit la variable contenant les données brutes disponibles plus les données manquantes estimées.

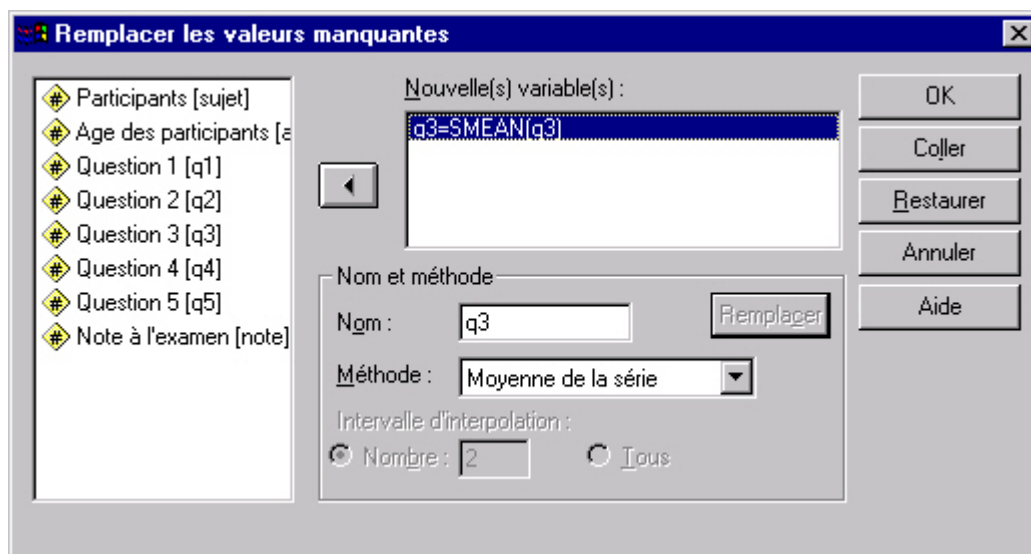


Figure 4.25. La commande Remplacer les valeurs manquantes

Pour expliquer l'estimation de données manquantes, examinons la variable q_3 du fichier illustré à la Figure 4.23. Cette variable comprend une valeur manquante à l'observation 4. La Figure 4.25 présente les instructions nécessaires pour remplacer la donnée manquante par la moyenne de la variable, tout en conservant le même nom de variable. Un examen de la Figure 4.26 nous permet de constater que la valeur manquante a été estimée et remplacée par la valeur 2.8.

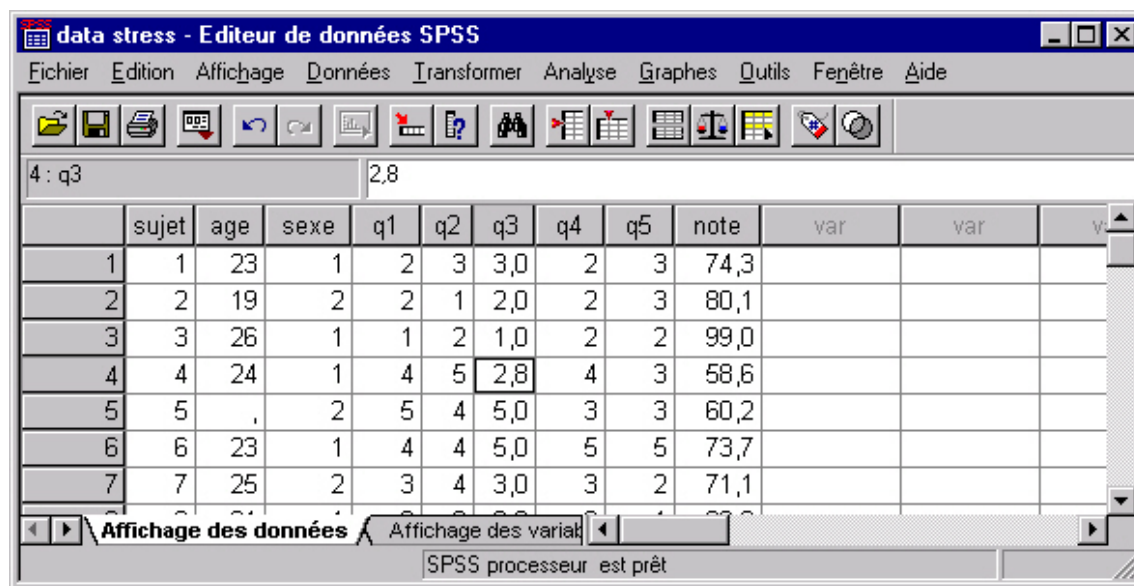


Figure 4.26. Estimation et remplacement d'une valeur manquante

Calcul d'une somme de variables. Dans le menu **TRANSFORMER**, sélectionnez la commande **CALCULER** (voir Figure 4.27). La partie inférieure gauche de la fenêtre présente les variables disponibles (i.e., les variables déjà existantes) pour réaliser le calcul. La partie supérieure gauche permet de spécifier le nom et l'étiquette de la nouvelle variable. La partie droite de la fenêtre comprend une case (partie supérieure) permettant d'indiquer la transformation attendue et une liste des fonctions automatiques disponibles dans SPSS (partie inférieure). Un examen sommaire de la liste des fonctions disponibles (partie inférieure droite) illustre bien l'ensemble des possibilités de transformation de variables de la commande **CALCULER**. Le clavier numérique au centre peut être utilisé avec la souris pour entrer différents opérateurs mathématiques. Toutefois, la saisie directe au clavier de l'ordinateur est généralement suffisante.

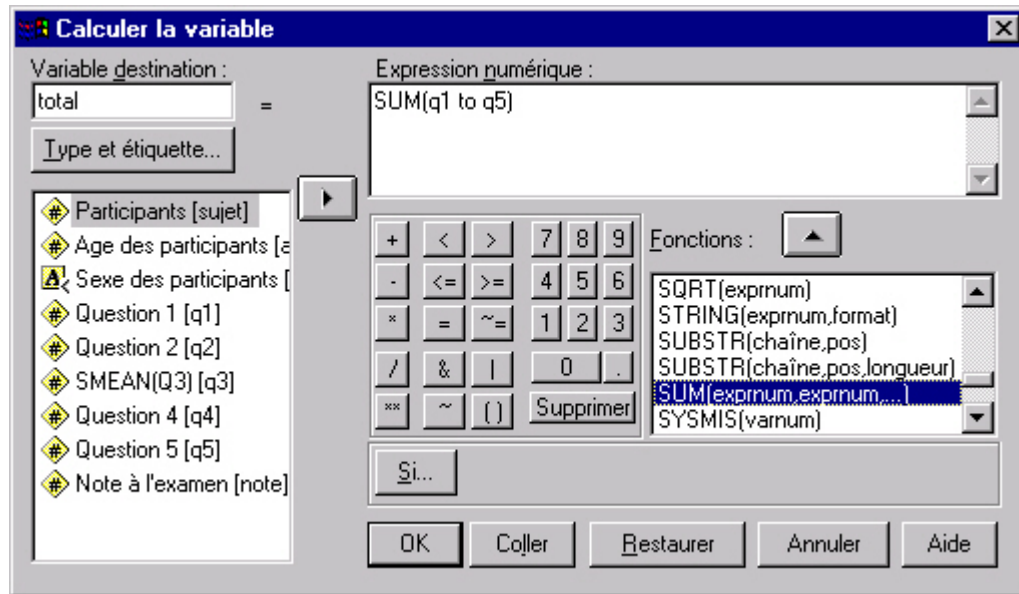


Figure 4.27. Calcul de la somme d'une série de variables

Pour illustrer le fonctionnement de cette fenêtre, nous allons créer une nouvelle variable appelée `total` comprenant la somme des questions 1 à 5 (`q1-q5`). La Figure 4.27 illustre l'apparence de la fenêtre lorsque toutes les instructions nécessaires ont été saisies. Après avoir appuyé sur **OK**, la nouvelle variable `total` sera ajoutée à la suite des variables déjà disponibles dans l'**ÉDITEUR DE DONNÉES**. Cette nouvelle variable peut être soumise à toutes les analyses, comme les variables déjà existantes.

Recodification d'une variable. Dans certaines situations d'analyse de données, l'expérimentateur désire recoder une variable. Généralement, la recodification de variables constitue une transformation d'une variable mesurée sur une échelle d'intervalle en une variable ordinaire ou nominale. Cette transformation peut être effectuée grâce à la commande **RECODER** dans le menu **TRANSFORMER**. Un petit exemple permet d'illustrer cette commande.

Un chercheur désire créer trois groupes de participants : (a) ceux qui ont obtenu au moins 85% dans l'évaluation, tel que mesurée par la variable `note`, (b) ceux qui ont obtenu entre 72% et 85% et (c) ceux qui ont obtenu moins de 72%. D'un point de vue psychométrique, le chercheur désire donc recoder une variable quantitative, allant de 0% à 100%, en une variable discrète, comprenant 3 groupes.

La commande **RECODER** offre deux choix à l'ouverture, soit de recoder les données dans une même variable ou de créer une nouvelle variable. Dans le cas présent, nous désirons conserver la variable *note*. En conséquence, une nouvelle variable doit être créée (voir Figure 4.28). La partie gauche de la fenêtre comprend la liste des variables disponibles pour la recodification. Sélectionnez la variable *note*. La partie droite permet de saisir le nom de la nouvelle variable qui sera créée suite à l'opération. Vous devez saisir le nom de la variable destination, *groupe*, dans cette fenêtre et cliquer sur le bouton **REEMPLACER**. La relation entre les deux variables devrait apparaître dans la fenêtre centrale.

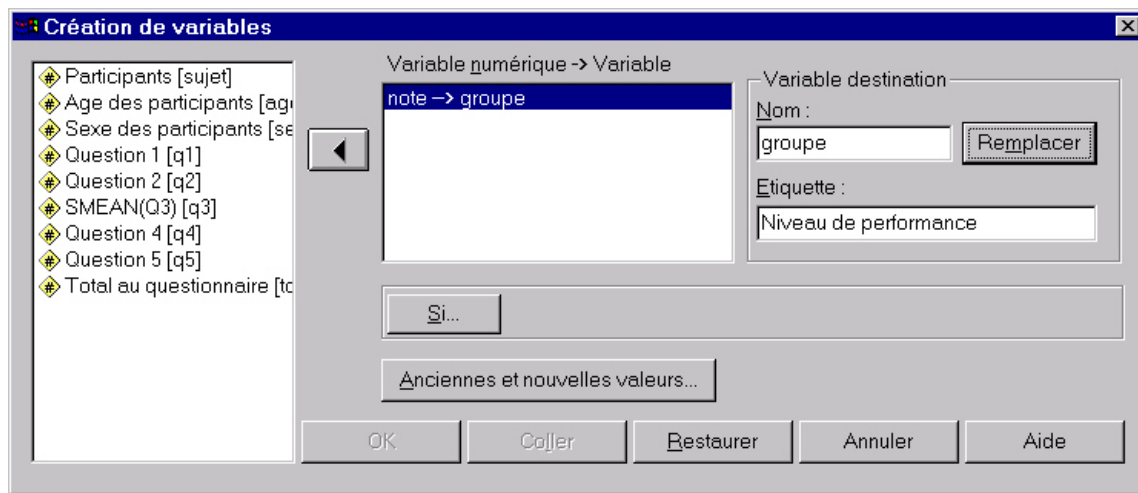


Figure 4.28. Recoder une variable – création de variable

Cliquez sur le bouton **ANCIENNES ET NOUVELLES VALEURS**. Dans la nouvelle fenêtre, la partie gauche permet de spécifier la(les) valeur(s) de *note* et la partie droite, la valeur de la variable *groupe*. Par exemple, pour placer les notes de 85% à 100% dans le groupe #1, les instructions illustrées à la Figure 4.29 doivent être appliquées. Cliquez ensuite sur le bouton **AJOUTER**.

Figure 4.29. Recoder une variable – spécification des nouvelles valeurs

La relation apparaît dans la fenêtre **ANCIENNE-->NOUVELLE**. La Figure 4.30 illustre toutes les opérations pour recoder la variable *note* dans *groupe*. Assurez-vous qu'aucune valeur n'est oubliée durant la transformation afin d'éviter de créer par erreur des valeurs manquantes pour la nouvelle variable. Après cette vérification, cliquez sur **POURSUIVRE** pour revenir dans la fenêtre principale de **RECORDER**, puis cliquez sur **OK** pour créer la nouvelle variable.

Figure 4.30. Recoder une variable – ensemble des instructions

5. L'ANALYSE STATISTIQUE DES DONNÉES

L'analyse statistique des données constitue une des étapes les plus stimulantes d'une recherche scientifique. C'est notamment à cette étape qu'il est possible de vérifier si les données récoltées supportent (confirment) ou ne supportent pas (infirment) les hypothèses de recherche. Le but de la présente section est d'illustrer le fonctionnement général des procédures d'analyse statistique du logiciel SPSS et de présenter les différents sous-menus du menu **ANALYSE**. L'utilisateur pourra donc obtenir un aperçu des principaux tests statistiques disponibles dans le logiciel SPSS. Bien que le but du présent manuel ne soit pas de présenter une à une les procédures d'analyse, quatre exemples seront détaillés : une analyse exploratoire des données, un test t de Student pour échantillons indépendants, une corrélation de Pearson et un test du khi-deux.

Avant d'utiliser le logiciel SPSS pour les analyses statistiques, l'utilisateur doit posséder au préalable une bonne connaissance théorique de la méthode d'analyse statistique prévue. Il est essentiel de connaître notamment les buts d'une méthode d'analyse, les questions de recherche pouvant être répondues à l'aide de cette méthode, ses pré-requis (i.e., conditions d'utilisation) et ses limites. *Ainsi, le logiciel SPSS ne transforme pas un utilisateur novice en un statisticien professionnel: il ne fait qu'accélérer le temps de calcul des statistiques.*

L'utilisation des fenêtres de dialogue

Une des particularités du logiciel SPSS qui en facilite l'utilisation est le recours à des fenêtres de dialogue pour identifier les éléments d'une analyse. Certains logiciels requièrent une programmation informatique ardue des procédures statistiques. Les fenêtres de dialogue du logiciel SPSS simplifient l'analyse de données en réduisant au minimum le recours à la programmation par commande ou syntaxe.

La fenêtre de dialogue d'une analyse apparaît lorsque vous choisissez le type d'analyse statistique que vous désirez effectuer dans le menu **ANALYSE**. Cette fenêtre requiert deux grandes catégories d'informations, soit (a) les variables du fichier de données qui seront soumises à l'analyse choisie, et (b) les options statistiques nécessaires pour préciser l'analyse en question. La Figure 5.1 fournit un exemple d'une fenêtre de dialogue d'une analyse statistique.

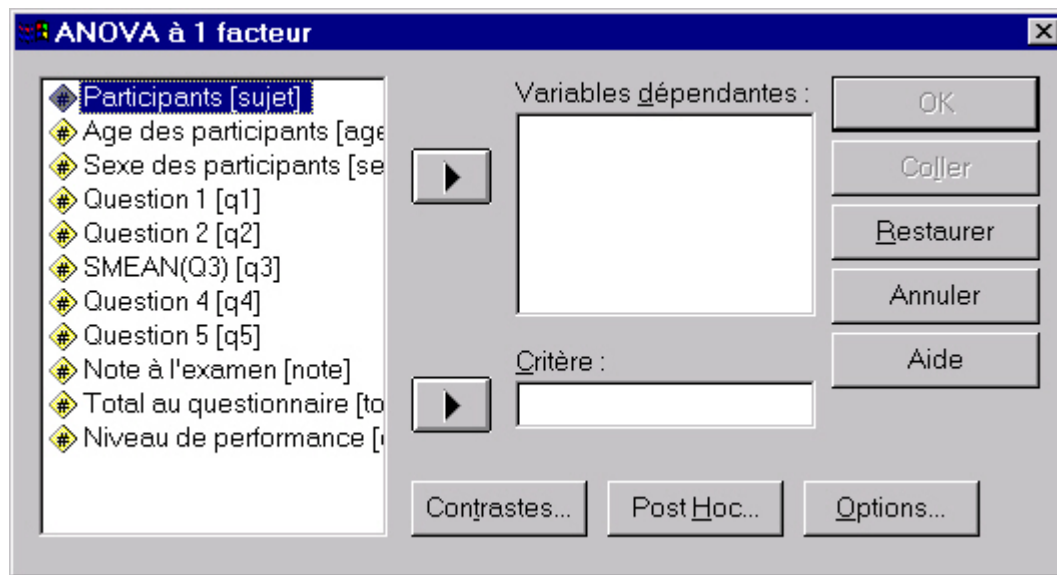


Figure 5.1. Fenêtre de dialogue d'une ANOVA

À l'examen de cette figure, il est possible de constater que la partie gauche de la fenêtre présente toutes les variables disponibles pour l'ANOVA. Pour sa part, la partie droite permet de spécifier une série d'options relatives à l'ANOVA, comme la nature des variables dépendantes et indépendante (**CRITÈRE**), les contrastes *a priori* et les comparaisons multiples *a posteriori* (**POST-HOC**). Arrêtons-nous ici pour l'utilisation des fenêtres de dialogue. Nous illustrerons plus loin l'utilisation de ces fenêtres dans la réalisation des exemples d'analyse statistique.

Les sous-menus du menu Analyse

Les sous-menus du menu **ANALYSE** comprennent la *majorité* des procédures statistiques disponibles dans le logiciel SPSS. Mentionnons que ces sous-menus sont organisés afin de regrouper les procédures d'analyse répondant à une question de recherche similaire. Ainsi, les tests de comparaison de moyennes (i.e., test t de Student pour échantillon unique, pour échantillons indépendants et dépendants, ANOVA à un facteur) sont regroupés dans le sous-menu **COMPARER LES MOYENNES**. La section suivante présente les différents sous-menus du menu **ANALYSE** et énumère les différentes procédures statistiques incluses selon le sous-menu.

RAPPORTS :

Cubes OLAP ; récapitulatif des observations ; tableaux de bord en lignes ; tableaux de bord en colonnes.

STATISTIQUES DESCRIPTIVES :

Fréquences (mesures de position) ; statistiques descriptives (moyenne, écart type, variance, mode, médiane, minimum, maximum, erreur standard, degré d'aplatissement, degré de symétrie) ; explorer (vérification des postulats univariés, graphique en boîtes à moustaches, diagramme en tige et feuilles, histogramme, test de Levene); tableaux croisés (khi-deux, coefficient de contingence, Phi et V de Cramer, Kappa, Statistiques de Cohran et Mantel-Haenszel), statistiques de ratio (coefficient de variation, etc.)

TABULER :

Tableaux statistiques de base, généraux, des réponses multiples et de fréquences (statistiques descriptives par tableaux).

COMPARER LES MOYENNES :

Test t pour un échantillon ; test t pour échantillons indépendants et dépendants ; ANOVA à 1 facteur ; contrastes *a priori* et comparaisons multiples *a posteriori*.

MODÈLE LINÉAIRE GÉNÉRAL (GLM) :

Analyse de variance univariée (ANOVA) et multivariée (MANOVA) ; contrastes *a priori*, comparaisons multiples *a posteriori*, effets simples, interactions traitement-contraste et contraste-contraste; analyse de tendance ; analyse de variance à mesures répétées; analyse de covariance.

MODÈLES MIXTES:

Analyse de variance selon un modèle linéaire mixte (effets aléatoires répétés ou non-répétés). Permet de spécifier la nature de la covariance.

CORRÉLATION :

Corrélations bilatérale et unilatérale (i.e., Pearson, Spearman, Tau-b de Kendall) ; corrélation partielle ; indices de covariation (distance)

RÉGRESSION :

Régression linéaire (simple, multiple, hiérarchique, stepwise, analyse des résidus) ; régression logistique ; régression non linéaire.

ANALYSE LOG-LINÉAIRE :

Analyse de logarithme (généralisée, Logit, analyse des résidus, etc.).

CLASSIFICATION :

Analyse discriminante; analyse de nuées dynamiques (*cluster analysis*); analyse de classification hiérarchique.

FACTORISATION :

Analyse factorielle (analyse factorielle, analyse en composante principale, rotations orthogonales et non orthogonales); analyse des correspondances; codage optimal.

POSITIONNEMENT :

Analyse d'échelles; analyse de fidélité (cohérence interne, qualité des items, etc.); analyse de positionnement multidimensionnel.

TESTS NON PARAMÉTRIQUES :

Test du khi-deux, test binomial ; tests pour 1 échantillon (distribution Gaussienne et Poisson) ; tests pour 2 échantillons indépendants (Mann Whitney, Kolmogorov-Smirnov) ; tests pour plusieurs échantillons indépendants (Kruskal-Wallis) ; tests pour 2 échantillons dépendants (Wilcoxon, McNemar, test du signe) ; tests pour plusieurs échantillons dépendants (Friedman, Kendall, Cochran).

SÉRIES CHRONOLOGIQUES :

Lissage exponentiel ; modèles d'auto-régression (Cochran-Orcutt, Prais-Winsten) ; modèles ARIMA ; désaisonnalisation.

SURVIE :

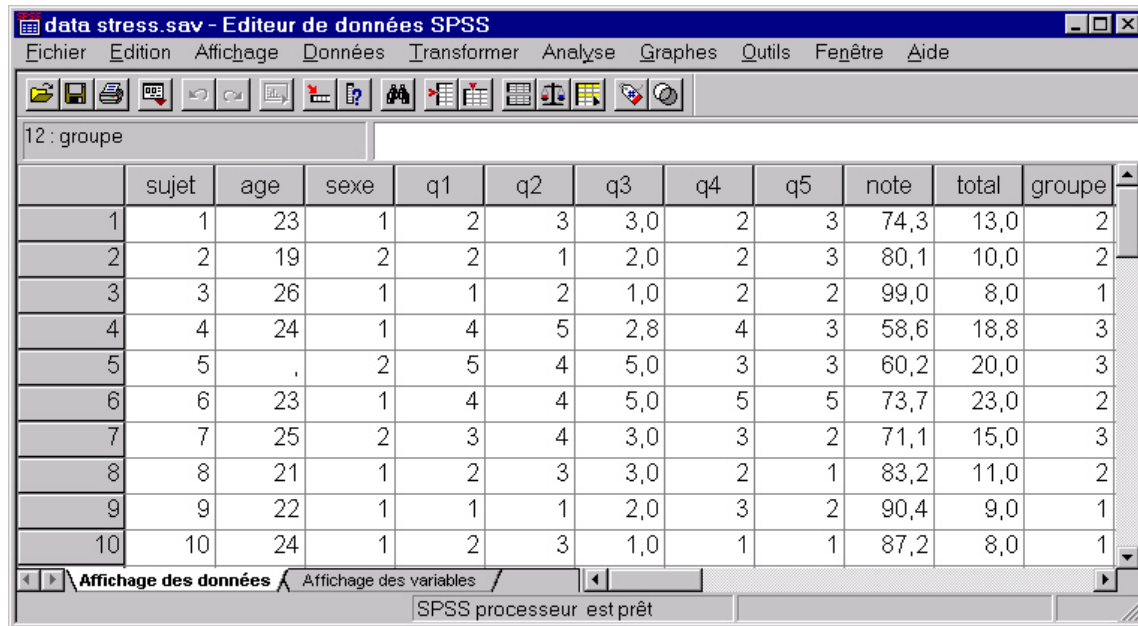
Analyse de durée de vie ; Kaplan Meier ; Modèle de Cox, Modèle de Cox à prédicteur temporel.

RÉPONSES MULTIPLES :

Vecteurs de réponses multiples (questions à plusieurs réponses possibles).

Afin d'illustrer le fonctionnement du logiciel SPSS et l'utilisation des fenêtres de dialogue, des exemples d'analyse sont présentés dans la section suivante. Ces exemples sont réalisés à partir du fichier de données présenté à la Figure 5.2.

Ce fichier a été obtenu suite aux diverses modifications illustrées à la section précédente.



	sujet	age	sexe	q1	q2	q3	q4	q5	note	total	groupe
1	1	23	1	2	3	3,0	2	3	74,3	13,0	2
2	2	19	2	2	1	2,0	2	3	80,1	10,0	2
3	3	26	1	1	2	1,0	2	2	99,0	8,0	1
4	4	24	1	4	5	2,8	4	3	58,6	18,8	3
5	5	.	2	5	4	5,0	3	3	60,2	20,0	3
6	6	23	1	4	4	5,0	5	5	73,7	23,0	2
7	7	25	2	3	4	3,0	3	2	71,1	15,0	3
8	8	21	1	2	3	3,0	2	1	83,2	11,0	2
9	9	22	1	1	1	2,0	3	2	90,4	9,0	1
10	10	24	1	2	3	1,0	1	1	87,2	8,0	1

Figure 5.2. Données utilisées pour les exemples statistiques

Exemple d'une analyse exploratoire des données

Le but de cette analyse est d'explorer les variables `note` et `total`.

Étape #1. Dans le menu **ANALYSE**, choisir le sous-menu **STATISTIQUES DESCRIPTIVES** puis la procédure **EXPLORER**.

Étape #2. À partir de la liste de variables apparaissant dans la partie gauche, sélectionnez les variables `note` et `total` puis cliquez sur la flèche du haut pour les transférer dans la liste des variables dépendantes.

Étape #3. La fenêtre **VARIABLES ACTIVES** vous permet d'effectuer les analyses descriptives en fonction d'un groupe d'observations (i.e., variable indépendante). Par exemple, si vous désirez effectuer les analyses séparément pour chacun des sexes, insérez la variable `sexe` dans cette fenêtre.

Étape #4. Indiquez ensuite si vous désirez obtenir vos résultats sous forme de statistiques, de graphiques ou les deux (voir Figure 5.3).

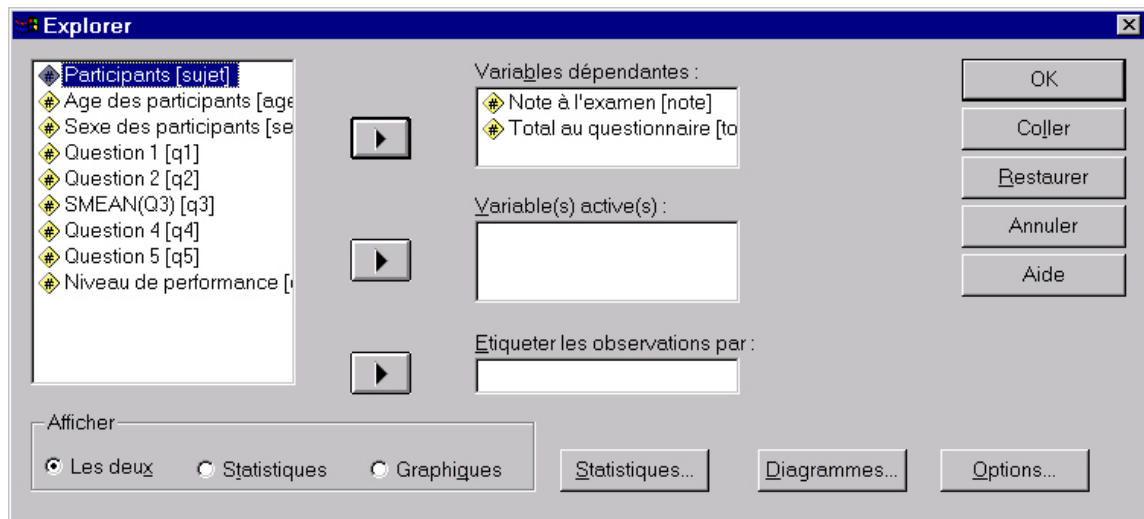


Figure 5.3. Exploration des variables note et total

Étape #5. Après avoir appuyé sur **OK**, une nouvelle fenêtre, soit le **NAVIGATEUR SPSS**, présente le résultat des analyses exploratoires. Un examen détaillé de votre écran vous permettra de constater que ces résultats sont très complets (voir Figure 5.4).

The screenshot shows the SPSS Results Viewer window titled 'Résultats11 - Viewer SPSS'. The menu bar includes 'Fichier', 'Edition', 'Affichage', 'Insérer', 'Format', 'Analyse', 'Graphes', 'Outils', 'Fenêtre', and 'Aide'. The toolbar contains various icons for file operations, editing, and analysis. The main display area shows a table of statistical results for the variable 'Note à l'examen'.

			Statistique	Erreur standard
Note à l'examen	Moyenne		77,780	4,060
	Intervalle de confiance à 95% pour la moyenne	Borne inférieure	68,596	
		Borne supérieure	86,964	
	Moyenne tronquée à 5%		77,667	
	Médiane		77,200	
	Variance		164,817	
	Ecart-type		12,838	
	Minimum		58,6	
	Maximum		99,0	
	Intervalle		40,4	
	Intervalle interquartile		19,625	
	Asymétrie		-,021	,687

At the bottom of the window, a status bar indicates 'SPSS processeur est prêt'.

Figure 5.4. Résultats de l'analyse exploratoire avec la commande Explorer

Exemple d'un test t de Student pour échantillons indépendants

Le but de cette seconde analyse consiste à vérifier si les hommes présentent un niveau de stress significativement différent des femmes de l'échantillon. Deux variables sont nécessaires, soit une variable dépendante (*total*) et une variable indépendante (*sexe*).

Étape #1. Dans le menu **ANALYSE**, choisir le sous-menu **COMPARER LES MOYENNES** puis la procédure **TEST T POUR ÉCHANTILLONS INDÉPENDANTS**.

Étape #2. Sélectionnez la variable *total* dans la partie gauche de la fenêtre de dialogue, puis cliquez sur la flèche vis-à-vis la fenêtre **VARIABLE À TESTER** (i.e., variable dépendante).

Étape #3. Sélectionnez ensuite la variable *sexe* et cliquez sur la flèche vis-à-vis la fenêtre **CRITÈRE DE REGROUPEMENT** (i.e., variable indépendante). La fenêtre de dialogue est illustrée à la Figure 5.5.

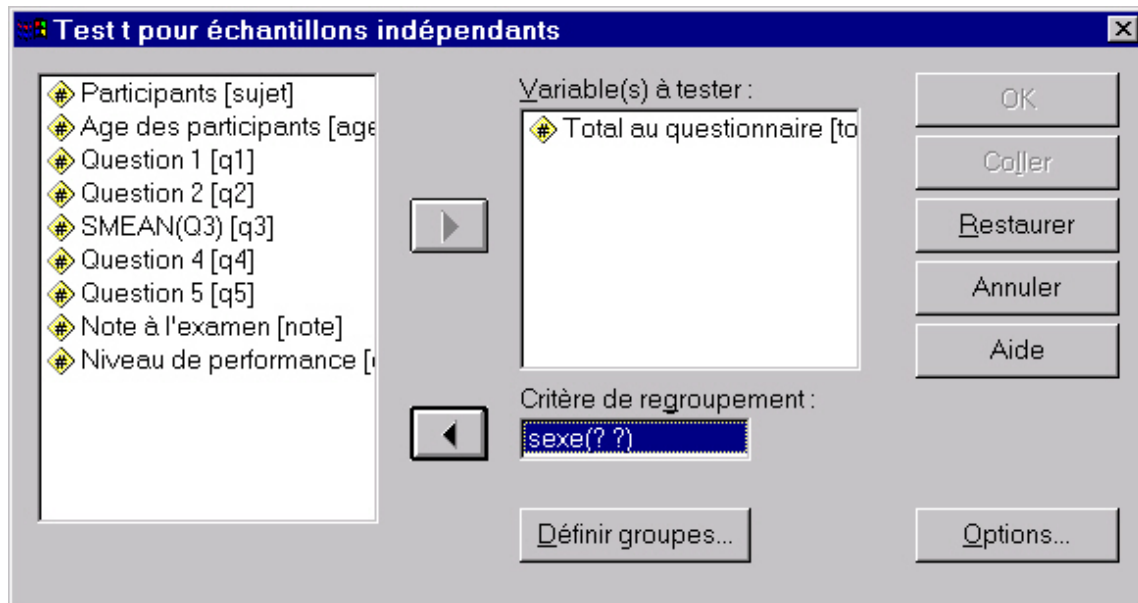


Figure 5.5. Test t pour échantillons indépendants

Étape #4. Vous devez maintenant définir les groupes formés à l'aide de la variable indépendante. Cliquez sur le bouton **DÉFINIR GROUPE**. Inscrire « 1 » dans la fenêtre **GROUPE 1** et « 2 » dans la fenêtre **GROUPE 2** (voir Figure 5.6). Ces valeurs correspondent aux hommes et aux femmes dans l'**ÉDITEUR DE DONNÉES**. Cliquez sur le bouton **POURSUIVRE**.

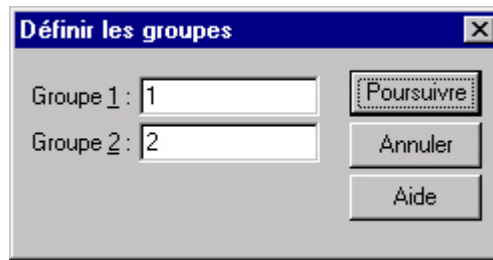


Figure 5.6. Spécification des groupes

Étape #5. Lorsque vos groupes sont définis, cliquez sur le bouton **OK** pour soumettre les instructions au logiciel SPSS. Le **NAVIGATEUR SPSS** s'ouvre automatiquement pour faire apparaître les résultats du test t de Student (voir Figure 5.7).

		Test de Levene sur l'égalité des variances		Test-t pour l'égalité des moyennes		
		F	Sig.	t	ddl	Sig. (bilatérale)
Total au questionnaire	Hypothèse de variances égales	,321	,587	-,522	8	,616
	Hypothèse de variances inégales			-,559	4,497	,603

Figure 5.7. Résultats du test t pour échantillons indépendants

Exemple du calcul d'un coefficient de corrélation de Pearson

Le but de cette troisième analyse est de vérifier la présence d'une association significative entre le niveau de stress (*total*) et la performance scolaire (*note*).

Étape #1. Dans le menu **ANALYSE**, choisir le sous-menu **CORRÉLATION** puis la procédure **BIVARIÉE**.

Étape #2. Dans la liste des variables disponibles (i.e., partie gauche de la fenêtre de dialogue), sélectionnez les variables `note` et `total`. Cliquez sur la flèche pour les transférer dans la fenêtre **VARIABLES** (voir Figure 5.8).

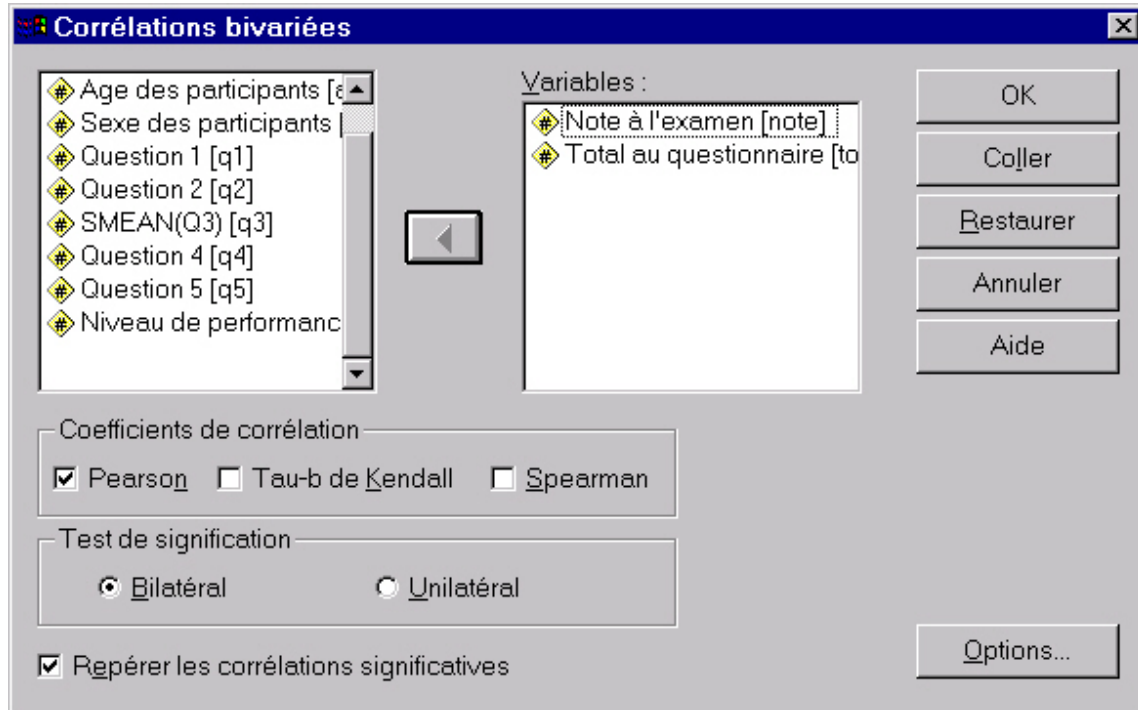


Figure 5.8. Spécification des variables pour une analyse de corrélation

Étape #3. Spécifiez le type de corrélation que vous désirez obtenir (e.g., **PEARSON**) et la direction du test de signification (**UNILATÉRAL** ou **BILATÉRAL**). Vous pouvez cliquer sur le bouton **OPTION** afin de modifier les options de l'analyse (e.g., pour obtenir les moyennes et écart types des variables).

Étape #4. Après avoir complété ces étapes, cliquez sur le bouton **OK** pour soumettre les instructions au logiciel SPSS. Le **NAVIGATEUR SPSS** s'ouvre automatiquement pour présenter les résultats de l'analyse de corrélation (voir Figure 5.9). À l'examen de cette figure, vous pourrez noter que le logiciel produit trois indices numériques par cas, soit le coefficient de corrélation, le degré de signification et le nombre d'observations. Pour obtenir le nombre de degrés de liberté pour ce coefficient, vous devez soustraire un au nombre d'observations ($n - 1$).

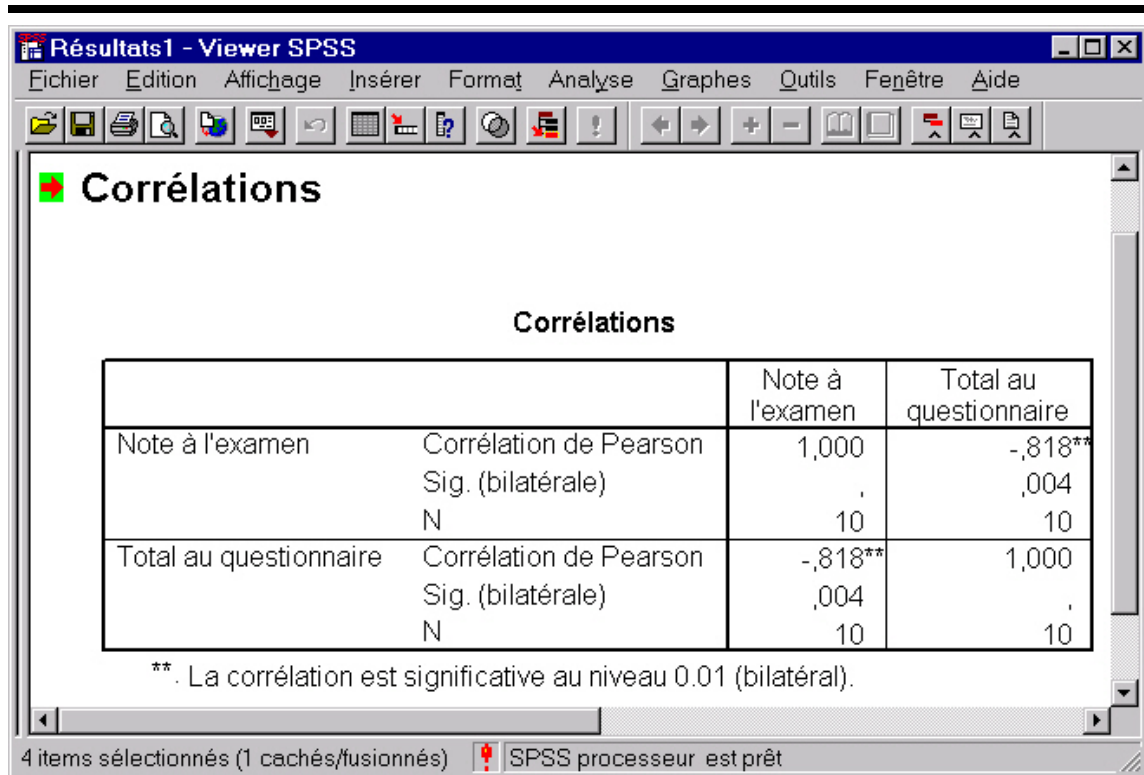


Figure 5.9. Corrélation de Pearson entre la note et le niveau de stress

Exemple du calcul d'un coefficient du khi-deux

Le but de cette analyse est de vérifier la présence d'une différence significative sur une variable nominale (*groupe*) en fonction d'une autre (*sexe*).

Étape #1. Dans le menu **ANALYSE**, choisir le sous-menu **STATISTIQUES DESCRIPTIVES** puis la procédure **TABLEAUX CROISÉS**. Si l'analyste désire réaliser un test d'ajustement (*goodness-of-fit*) sur une seule variable nominale, il doit plutôt sélectionner le sous-menu **TESTS NON-PARAMÉTRIQUES** puis la commande **KHI-DEUX**.

Étape #2. Sélectionnez la variable *groupe* dans la partie gauche de la fenêtre de dialogue, puis cliquez sur la flèche vis-à-vis la fenêtre **LIGNE(S)** (i.e., variable dépendante).

Étape #3. Sélectionnez ensuite la variable *sexe* et cliquez sur la flèche vis-à-vis la fenêtre **COLONNE(S)** (i.e., variable indépendante). La fenêtre de dialogue est illustrée à la Figure 5.10.

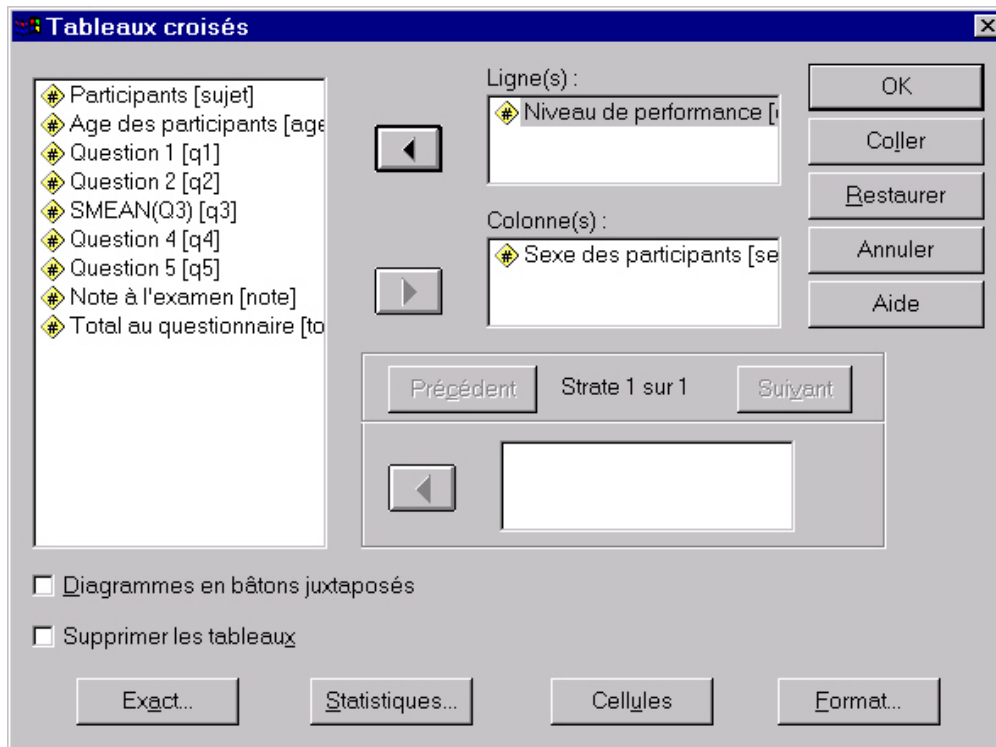


Figure 5.10. Spécification des variables pour une analyse de khi-deux

Étape #4. Cliquez sur le bouton **STATISTIQUES** et cochez la case **KHI-DEUX**, puis cliquez sur le bouton **POURSUIVRE**. Cliquez ensuite sur le bouton **CELLULES**, cochez toutes les cases dans la section **POURCENTAGES** puis cliquez sur le bouton **POURSUIVRE**.

Étape #5. Après avoir précisé ces options, vous pouvez cliquer sur le bouton **OK** de la fenêtre principale pour soumettre les instructions au logiciel SPSS. Le **NAVIGATEUR SPSS** s'ouvre automatiquement pour présenter les résultats de l'analyse du khi-deux (voir Figure 5.11). À l'examen de cette figure, vous pourrez noter que le logiciel produit trois indices numériques, soit le coefficient du khi-deux, le coefficient de vraisemblance et l'association linéaire par linéaire. De plus, il informe l'analyste du pourcentage des cellules ayant une fréquence théorique inférieure à 5 et pouvant invalider la statistique du khi-deux. Enfin, bien que ces données ne soient pas illustrées dans la Figure 5.11, cette commande SPSS produit un tableau croisé avec la répartition des données en fréquence brute et en pourcentage selon les deux variables nominales.

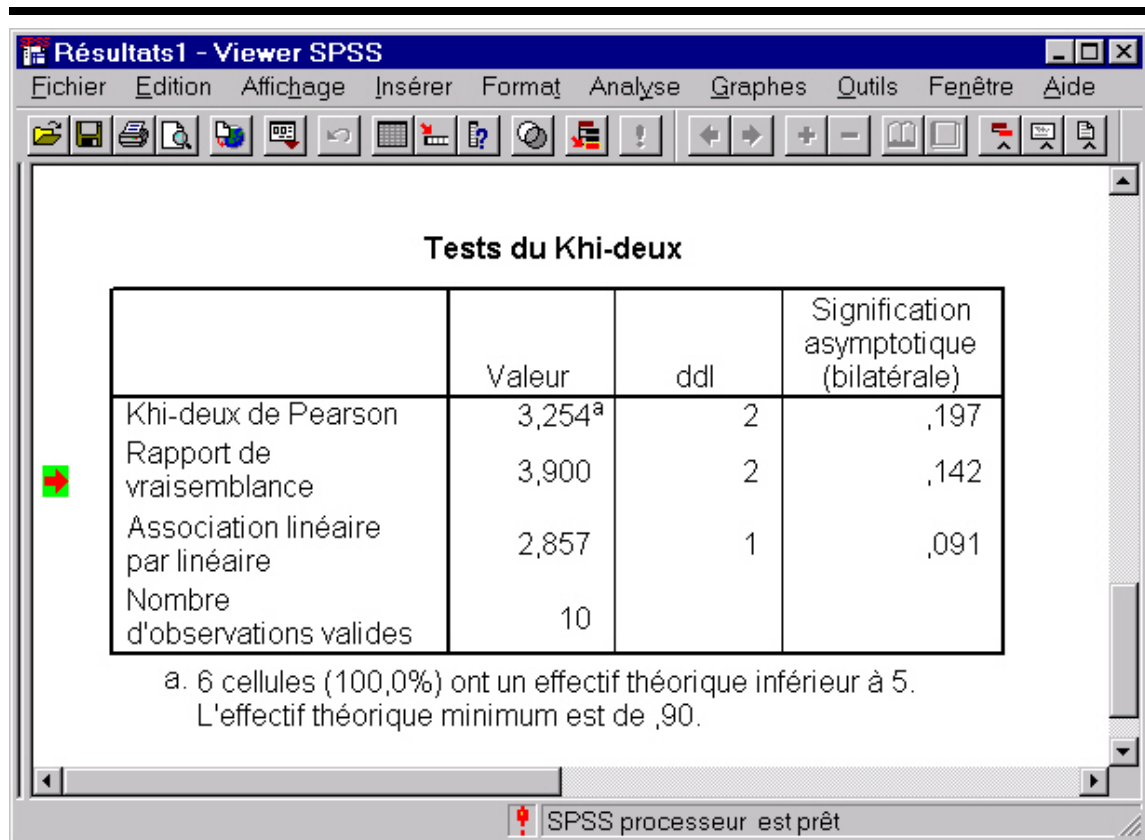


Figure 5.11. Tests du khi-deux pour le groupe en fonction du sexe

6. LA REPRÉSENTATION GRAPHIQUE DES DONNÉES

Le logiciel SPSS place à la disposition de l'utilisateur le menu **GRAPHES** pour représenter graphiquement les données. Il produit également certaines analyses complémentaires, telles les diagrammes en tige et feuilles ainsi que boîte et moustache, les diagrammes de dispersion, les courbes ROC en psychométrie et le calcul de l'autocorrélation totale et partielle pour l'analyse de séries chronologiques.

L'objectif de cette section est d'illustrer les fonctionnalités de ce menu à l'aide d'un court exemple. Évidemment, cette présentation n'est pas exhaustive et le lecteur est encouragé à explorer le menu **GRAPHES** par lui-même.

Illustration de la dispersion des données

Dans un exemple de la section précédente, le chercheur était intéressé à étudier le degré de covariance (la corrélation) entre le niveau de stress mesuré avant un examen et la performance à cet examen. Dans ce contexte, l'exploration visuelle de la covariance entre ces deux variables constitue une excellente analyse complémentaire.

L'objectif de cette analyse est donc d'illustrer la covariance entre les variables *note* et *total*, en réalisant un graphique du score d'une variable en fonction de l'autre. La commande **DIAGRAMMES DE DISPERSION** permet de réaliser facilement ce graphique.

Étape #1. Ouvrez le menu **GRAPHES** et sélectionnez la commande **DIAGRAMME DE DISPERSION**. Une fenêtre identique à la Figure 6.1 devrait s'ouvrir.

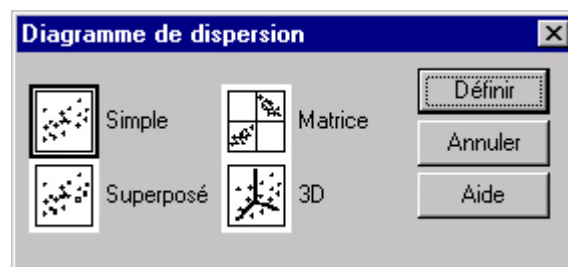


Figure 6.1. Menu de la commande Diagramme de dispersion

Étape #2. Sélectionnez l'option **SIMPLE** et cliquez sur **DÉFINIR**.

Étape #3. Une fenêtre de dialogue s'ouvre (voir Figure 6.2). La partie gauche présente l'ensemble des variables disponibles pour le graphique alors que la partie droite permet de spécifier les variables utilisées pour l'axe des X (en abscisse) et l'axe des Y (en ordonnée). Sélectionnez les variables *total* et *note*. Au besoin, vous pouvez également spécifier un titre, un sous-titre ainsi qu'une étiquette pour chaque axe du graphique.

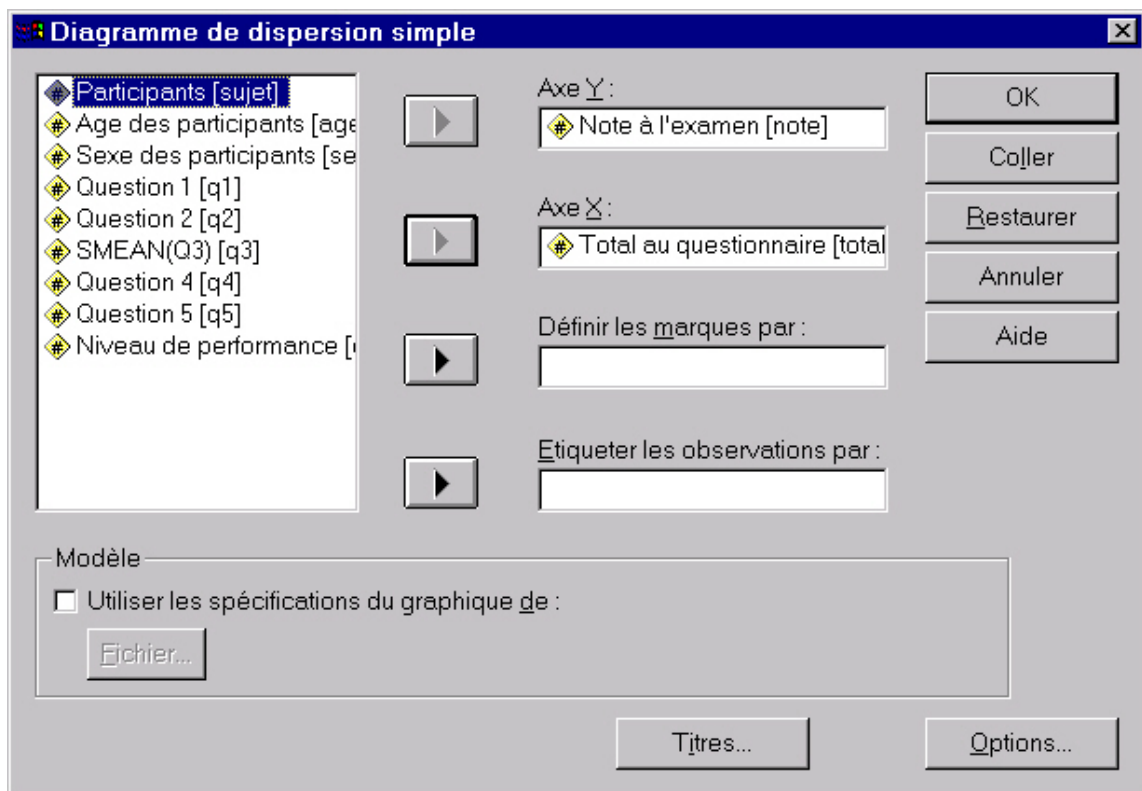


Figure 6.2. Spécification des variables dans un diagramme de dispersion

Étape #4. Cliquez sur **OK**. Le graphique sera produit dans le **NAVIGATEUR SPSS** (voir Figure 6.3). Vous pourrez imprimer directement ce graphique, l'exporter dans un autre logiciel ou le sauvegarder pour une utilisation future (pour plus de détails, consultez la section intitulée « Le Navigateur SPSS »).

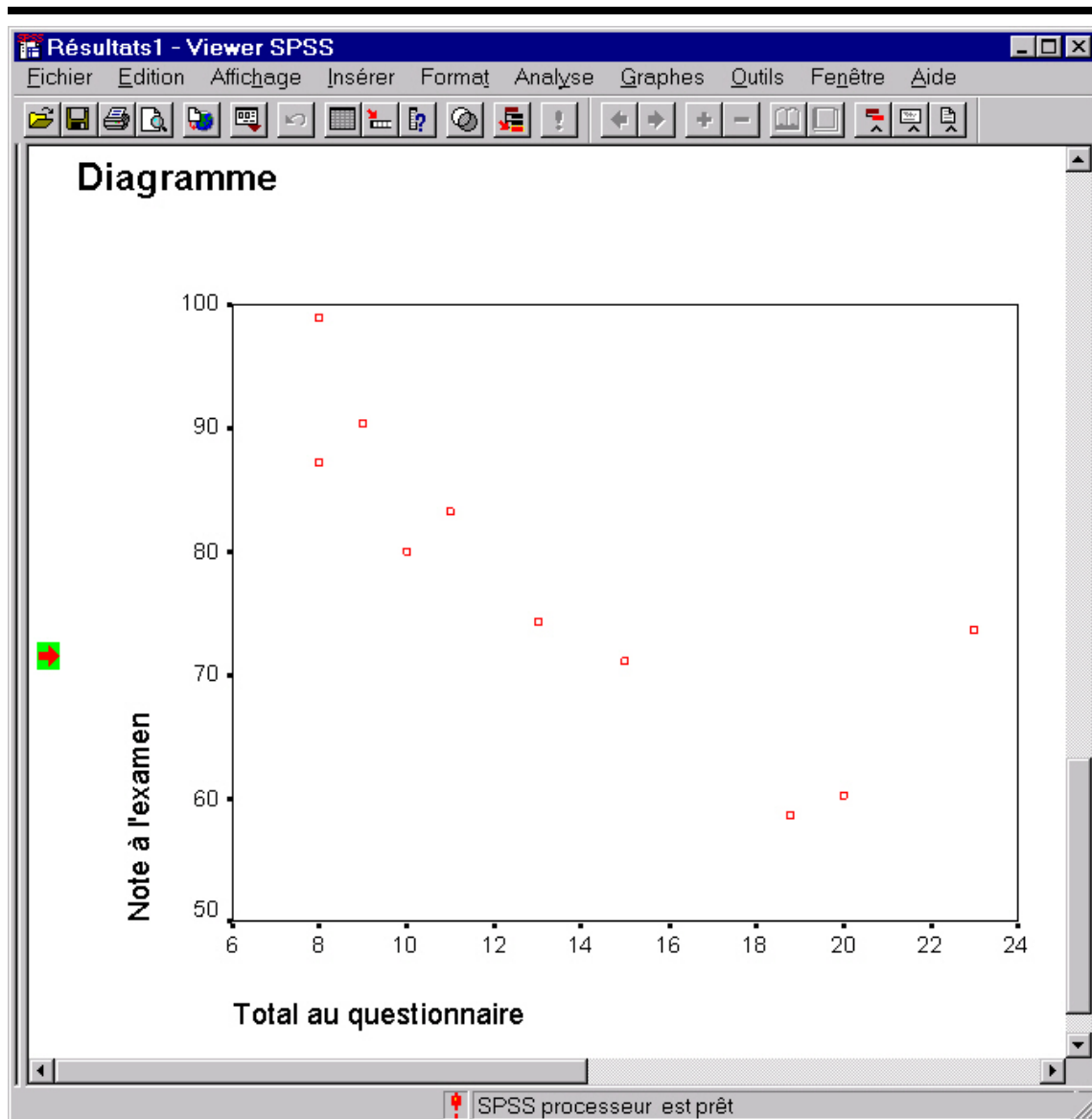


Figure 6.3. Affichage du diagramme de dispersion dans le Navigateur SPSS

7. LE NAVIGATEUR SPSS

Description générale

Le **NAVIGATEUR SPSS** est une fenêtre où les résultats générés par une procédure SPSS vont s'afficher automatiquement. Elle se divise en deux parties (voir Figure 7.1). La partie de gauche affiche les résultats SPSS en mode « Légende ». Ainsi, si vous effectuez plusieurs procédures statistiques, cette fenêtre vous offre la possibilité de repérer rapidement vos résultats. La partie droite contient tous les « Résultats détaillés » soit les tableaux statistiques, les graphiques et le texte. Pour accéder directement aux résultats que vous désirez visualiser, vous n'avez qu'à cliquer sur le titre apparaissant dans la partie « Légende ». Le **NAVIGATEUR** affiche directement les résultats demandés dans la partie droite.

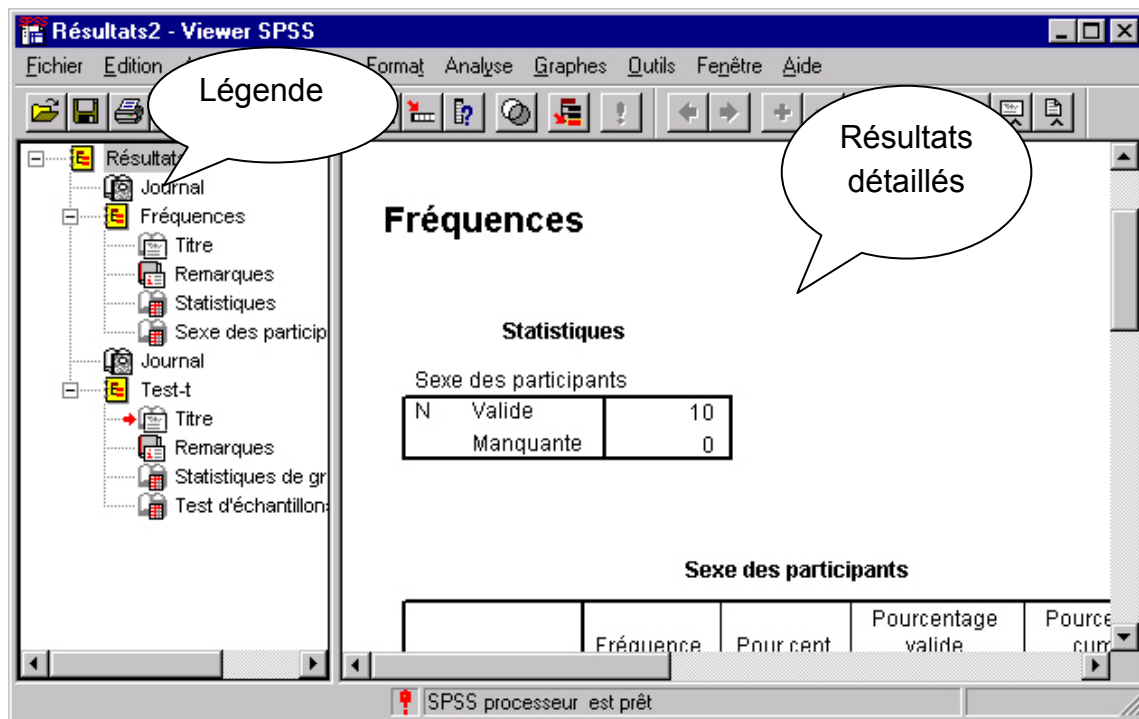


Figure 7.1. Vue complète du Navigateur SPSS

Il est également possible de déplacer ou d'éliminer des éléments de la partie « Légende » à l'aide de votre souris. Par exemple, cliquez sur l'élément à déplacer en prenant soin de tenir enfoncé le bouton gauche de votre souris, faites glisser l'élément à l'endroit désiré, puis relâchez le bouton de votre souris.

Vous pouvez également masquer certains résultats apparaissant dans la partie « Résultats détaillés » sans les supprimer. En fait, une icône en forme de livre situé à la droite d'un item de la partie « Légende » indique si l'élément est visible ou non dans la partie « Résultats détaillés ». Si le livre est ouvert, les résultats sont visibles sinon, ils sont masqués. Pour rendre visible ou masquer des résultats, cliquez deux fois sur l'icône en forme de livre pour l'ouvrir ou le fermer.

L'identification et la transformation du contenu

Une fois les tableaux statistiques et les graphiques affichés dans le **NAVIGATEUR SPSS**, il est possible de les nommer et de les modifier. Par défaut, SPSS attribue le nom de la commande utilisée comme titre au tableau et au graphique. Pour modifier ce titre, cliquez deux fois sur le nom par défaut et entrez le titre désiré. Le même principe s'applique à la transformation des tableaux et des graphiques. Le double clic active le tableau ou le graphique sélectionné et fait apparaître un menu permettant de modifier ses propriétés.

La sauvegarde des résultats SPSS

La sauvegarde des résultats s'effectue de la même façon que la sauvegarde des données présentée précédemment (voir la section sur la gestion des données). Choisissez le menu **FICHIER** dans le **NAVIGATEUR SPSS**, puis **ENREGISTRER SOUS**. Entrez un nom pour votre fichier puis cliquez sur le bouton **SAUVEGARDER**. Vous noterez cependant dans la fenêtre **ENREGISTRER SOUS** que le fichier est maintenant sauvegardé en format *.spo, soit un fichier de résultats SPSS.

L'impression des résultats SPSS

Avant de procéder à l'impression de vos résultats, vous devez choisir les éléments que vous désirez imprimer. Trois choix sont offerts : (a) imprimer seulement l'élément ou les éléments sélectionnés, (b) imprimer seulement les résultats affichés dans la partie « Résultats détaillés » ou encore (c) imprimer l'ensemble des résultats.

Pour imprimer un seul élément ou une partie des éléments, sélectionnez-les dans la partie « Légende » du **NAVIGATEUR SPSS**. Ensuite, choisissez le menu **FICHIER**, puis **IMPRIMER**. Vous noterez que l'option **SÉLECTION** est activée.

Cliquez sur **OK** pour confirmer l'impression. Pour imprimer seulement les résultats affichés dans la partie « Résultats détaillés », choisissez le menu **FICHIER**, puis **IMPRIMER**. Cochez ensuite l'option **TOUS LES RÉSULTATS AFFICHÉS** et cliquez sur **OK** pour confirmer. Finalement, pour imprimer l'ensemble des résultats, sélectionnez tous les résultats dans la partie « Légende » et effectuez les étapes permettant d'imprimer une sélection.

L'exportation des résultats SPSS.

La dernière fonction du **NAVIGATEUR SPSS** illustrée dans ce guide consiste à exporter un tableau ou un graphique de la partie « Résultats détaillés » dans une autre application (e.g., Microsoft Word). Dans la fenêtre **NAVIGATEUR**, cliquez sur le tableau, le graphique ou le texte que vous désirez exporter. Notez que vous pouvez réaliser cette sélection directement dans la partie « Résultats détaillés » ou encore dans la partie « Légende » du **NAVIGATEUR**. Choisissez ensuite le menu **ÉDITION** de la fenêtre puis **COPIER**. Dirigez-vous maintenant dans l'autre application (e.g., Microsoft Word) et sélectionnez **COLLAGE SPÉCIAL** du menu **ÉDITION** de cette même application.

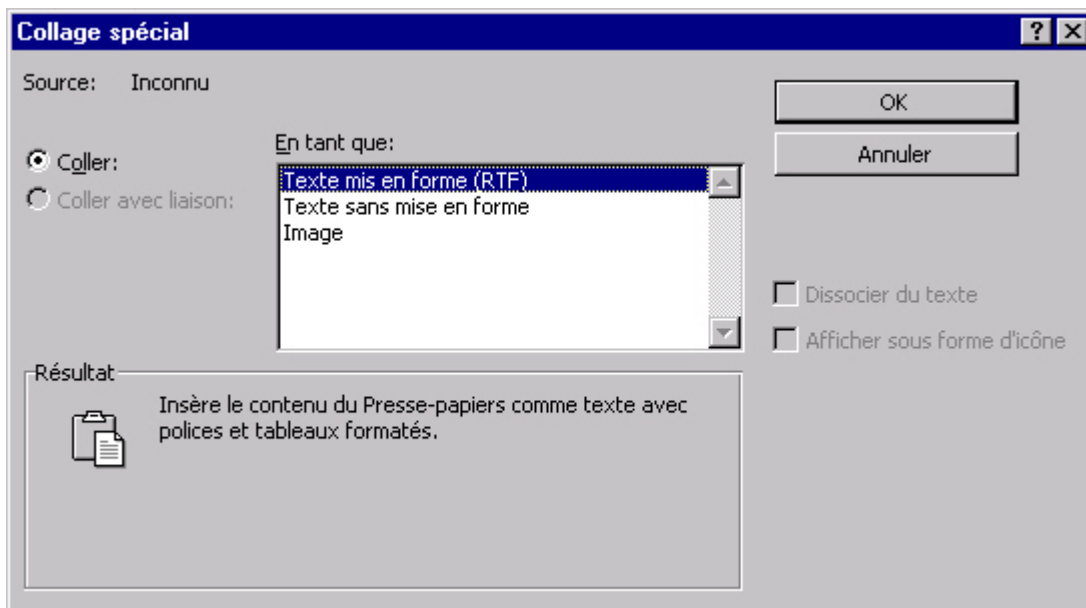


Figure 7.2. Collage spécial dans le logiciel Word

La fenêtre illustrée à la Figure 7.2 vous offre trois possibilités de transfert. Premièrement, vous pouvez **COLLER** votre sélection (résultats) dans votre nouvelle application sous la forme d'une **IMAGE**. Vous obtiendrez alors vos résultats sous une forme identique à celle apparaissant dans le logiciel SPSS.

Ceci est notamment utile lorsque vous désirez conserver vos formats d'origine. Vous pouvez également choisir de **COLLER** votre sélection en tant que **TEXTE MIS EN FORME (RTF)** ou **TEXTE NON MIS EN FORME**. Le premier format convertit vos résultats en texte avec des polices et tableaux formatés, tandis que le second convertit vos résultats en texte non formaté.

8. LA PROGRAMMATION EN MODE SYNTAXE

Toutes les commandes présentées dans les sections précédentes ont été illustrées à l'aide du « mode graphique », soit les fenêtres, les fenêtres de dialogue, les boutons de choix, et l'utilisation de la souris. Toutefois, le logiciel SPSS permet à l'utilisateur de parler directement avec le *processeur SPSS*, soit la partie du logiciel responsable de traduire les demandes de l'utilisateur en analyses statistiques. Cette discussion est réalisée grâce à une série de commandes ou *instructions* respectant une orthographe et une grammaire spécifique, d'où le nom de « mode syntaxe ».

Le mode syntaxe présente de nombreux avantages. Premièrement, l'inventaire des instructions disponibles en mode syntaxe est beaucoup plus étendu que celui du mode graphique. Deuxièmement, l'utilisation du mode syntaxe permet au chercheur de conserver une *copie* de toutes les transformations et analyses effectuées sur les données, ce que le mode graphique ne permet pas. Enfin, dans le cas d'analyses répétitives où seules quelques variables sont modifiées d'une analyse à l'autre, le mode syntaxe est parfois plus rapide que le mode graphique. Il est donc fortement recommandé de se familiariser avec ce mode de programmation du logiciel SPSS.

L'Éditeur de syntaxe

Pour utiliser le mode syntaxe, vous devez être dans l'**ÉDITEUR DE SYNTAXE**. Vous pouvez ouvrir cet éditeur en sélectionnant la commande **NOUVEAU / SYNTAXE** dans le menu **FICHIER** de l'**ÉDITEUR DE DONNÉES**. L'**ÉDITEUR DE SYNTAXE** se présente sous la forme d'une grande feuille blanche où les instructions peuvent être directement saisies à l'écran (voir Figure 8.1). Le menu de cet éditeur est similaire au menu de l'**ÉDITEUR DE DONNÉES** mais il se restreint aux analyses statistiques et à la représentation graphique des données. Un nouveau menu, intitulé **EXÉCUTER**, vous permet de contrôler la soumission de vos instructions au processeur SPSS. Évidemment, comme pour les fichiers de données et les fichiers de résultats, un fichier de syntaxe peut être sauvegardé et imprimé pour un usage futur. Ces commandes sont disponibles dans le menu **FICHIER**. Deux exemples spécifiques permettent d'illustrer le fonctionnement de l'**ÉDITEUR DE SYNTAXE**. Le premier porte sur une transformation complexe de données, alors que le second illustre une analyse statistique.

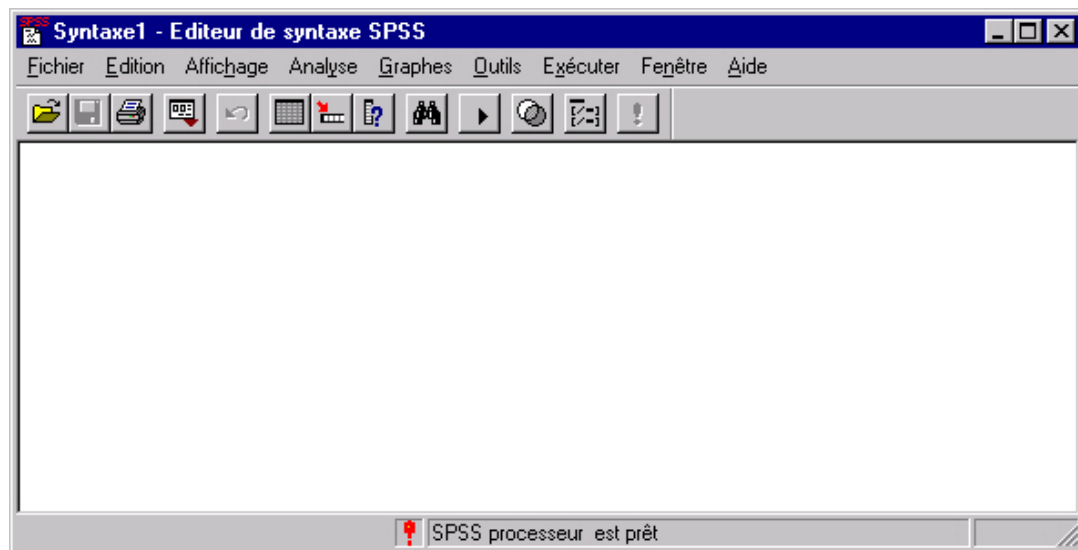


Figure 8.1. Éditeur de syntaxe SPSS

Exemple d'une transformation de données en mode syntaxe

L'utilisation de la syntaxe SPSS permet à l'analyste d'exécuter des commandes très complexes, qui n'ont pas d'équivalent en mode graphique. Pour illustrer ce propos, nous allons présenter un tel exemple. Un chercheur, sachant que l'âge et le sexe influencent le niveau de stress, décide de recalculer le total d'une mesure brute de stress (*total*) en pondérant pour l'âge (*age*) et le sexe (*sexe*) du répondant. Il décide d'ajouter 4 points aux hommes, qui ont tendance à sous-évaluer leur niveau de stress, et de retirer 2 points aux femmes âgées de 25 ans ou plus, car certains événements associés à la vie active d'adulte produisent une surévaluation du niveau de stress.

Les transformations peuvent être résumées par l'algorithme suivant, qui permet de produire un total pondéré (*totalp*) du niveau de stress :

```
par défaut, totalp = total.  
si sexe = homme,  
    totalp = total + 4.  
si sexe = femme et age >= 25,  
    totalp = total - 2.
```

Cet algorithme a été traduit en syntaxe SPSS à la Figure 8.2.

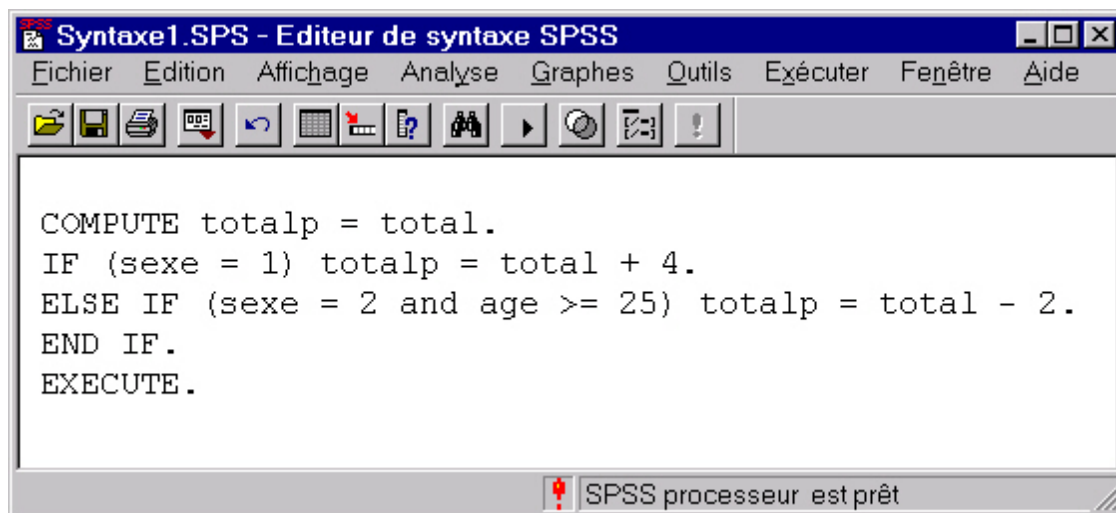


Figure 8.2. Instructions de syntaxe pour transformer une variable

L'utilisateur doit exécuter les instructions en sélectionnant le menu **EXÉCUTER** puis la commande **TOUT**. Les instructions sont alors acheminées au processeur SPSS. Le **NAVIGATEUR SPSS** s'ouvre et, en cas d'erreurs de syntaxe, il indique où se situent les erreurs dans les instructions de l'**ÉDITEUR DE SYNTAXE**. Comme nous avons demandé la création d'une nouvelle variable (`totalp`) à partir d'une ancienne variable (`total`), les résultats ne sont pas affichés dans le **NAVIGATEUR SPSS** mais bien dans l'**ÉDITEUR DE DONNÉES** (voir Figure 8.3).

The screenshot shows the 'data stress.sav - Editeur de données SPSS' window. The menu bar includes 'Fichier', 'Edition', 'Affichage', 'Données', 'Transformer', 'Analyse', 'Graphes', 'Outils', 'Fenêtre', and 'Aide'. The toolbar contains icons for opening, saving, printing, running, and other functions. The main data grid shows the following data:

	age	sexe	q1	q2	q3	q4	q5	note	total	groupe	totalp
1	23	hom	2	3	3,0	2	3	74,3	13,0	2	17,0
2	19	femm	2	1	2,0	2	3	80,1	10,0	2	10,0
3	26	hom	1	2	1,0	2	2	99,0	8,0	1	12,0
4	24	hom	4	5	2,8	4	3	58,6	18,8	3	22,8
5		femm	5	4	5,0	3	3	60,2	20,0	3	20,0
6	23	hom	4	4	5,0	5	5	73,7	23,0	2	27,0
7	25	femm	3	4	3,0	3	2	71,1	15,0	3	15,0
8	21	hom	2	3	3,0	2	1	83,2	11,0	2	15,0
9	22	hom	1	1	2,0	3	2	90,4	9,0	1	13,0
10	24	hom	2	3	1,0	1	1	87,2	8,0	1	12,0

The status bar at the bottom indicates 'SPSS processeur est prêt'.

Figure 8.3. Résultats de la transformation en mode syntaxe

Exemple d'une analyse statistique en mode syntaxe

L'exemple de la section précédente a pu démontrer la puissance du mode syntaxe mais également sa complexité. Généralement, il n'est pas nécessaire de connaître l'ensemble des instructions SPSS « par cœur ». L'aide en ligne du logiciel, ainsi que les manuels fournissent tout le support nécessaire pour réaliser une excellente programmation (voir la section intitulée « Ressources d'aide du logiciel SPSS »). Toutefois, il existe une méthode pour produire plus rapidement la syntaxe des analyses statistiques.

Pour illustrer l'utilisation de cette méthode, nous allons reproduire le calcul du coefficient de corrélation de Pearson, tel que présenté aux Figures 5.8 et 5.9. Rappelons que le but de cette analyse est de vérifier la présence d'une association significative entre le niveau de stress (`total`) et la performance scolaire (`note`).

Étape #1. Déplacez-vous dans l'**ÉDITEUR DE DONNÉES** puis, dans le menu **ANALYSE**, choisissez le sous-menu **CORRÉLATION** puis la procédure **BIVARIÉE**.

Étape #2. Dans la liste des variables disponibles dans la partie gauche de la fenêtre de dialogue, sélectionnez les variables `note` et `total`. Cliquez sur la flèche pour les transférer dans la fenêtre **VARIABLES**. Spécifiez le type de corrélation que vous désirez obtenir (e.g., **PEARSON**) et la direction désirée du test de signification (e.g., **BILATÉRAL**). Vous pouvez cliquer sur le bouton **OPTION** afin de modifier les options de l'analyse (e.g., pour obtenir les moyennes et écart types des variables). Vous devriez obtenir une fenêtre de dialogue identique à la Figure 5.8.

Étape #3. Après avoir complété ces étapes, cliquez sur le bouton **COLLER** (et non sur le bouton **OK**). Cette opération transfère les instructions de l'analyse corrélationnelle dans l'**ÉDITEUR DE SYNTAXE**. Ainsi, vous n'avez pas à saisir manuellement chaque instruction, au risque de faire des erreurs. L'**ÉDITEUR DE SYNTAXE** s'ouvre automatiquement pour présenter ces instructions (voir Figure 8.4). Vous n'avez qu'à soumettre les instructions au processeur SPSS, en sélectionnant le menu **EXÉCUTER** puis la commande **TOUT**. Le **NAVIGATEUR SPSS** présente alors les résultats de l'analyse corrélationnelle. Le contenu devrait être identique à la Figure 5.9.

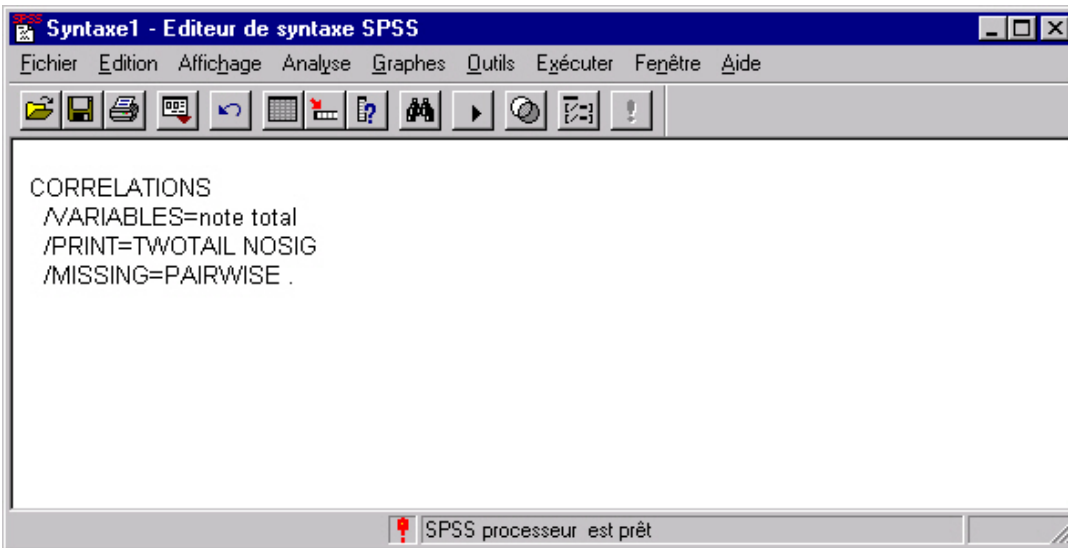


Figure 8.4. Transfert des instructions en mode syntaxe grâce au bouton Coller

Si le mode syntaxe vous rebute, vous pouvez produire automatiquement les instructions relatives à chaque opération réalisée sur vos données. Pour se faire, vous devez activer l'écriture automatique de toutes les instructions dans le **NAVIGATEUR SPSS**. Sélectionnez la commande **OPTIONS** dans le menu **ÉDITION** de l'**ÉDITEUR DE DONNÉES**. Dans la partie supérieure de la fenêtre de dialogue, sélectionnez le deuxième onglet intitulé **VIEWER**. Une fenêtre semblable à la Figure 8.5 devrait apparaître. Dans la partie inférieure gauche, vous devez cocher la case intitulée **AFFICHER COMMANDES**. Cliquez sur **OK** pour quitter.

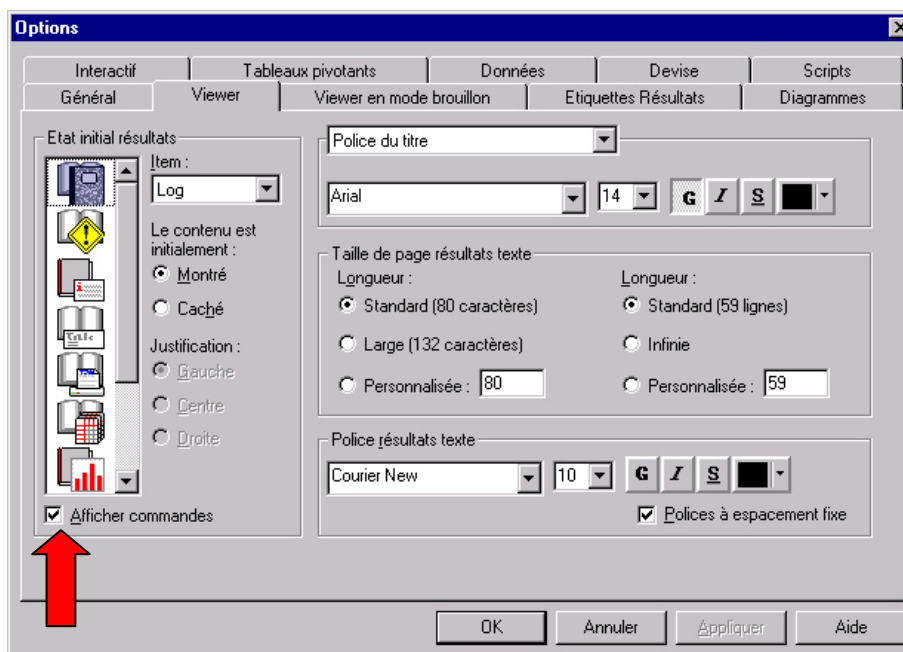


Figure 8.5. Option du logiciel SPSS pour activer l'affichage des instructions

Liste des commandes de syntaxe utilisées dans ce manuel

Afin d'illustrer quelques fonctions du mode syntaxe, cette section présente un sommaire des commandes associées aux procédures utilisées dans ce manuel. Vous trouverez une information complète à propos de chaque commande dans le guide de syntaxe de base de SPSS, présenté dans la section « Les ressources d'aide du logiciel SPSS ».

Tableau 8.1. Sommaire des commandes SPSS du manuel

Section	Opération	Commande SPSS
Gestion des données	Ouverture d'un fichier	GET
	Sauvegarde d'un fichier	SAVE
	Format d'une variable	FORMATS
	Etiquette de variable	VARIABLE LABELS
	Etiquette de valeur	VALUE LABELS
	Echelle de mesure	VARIABLE LEVEL
	Largeur des colonnes	VARIABLE WIDTH
	Alignement de variable	VARIABLE ALIGNMENT
	Gestion des données manquantes	MISSING VALUES / RECODE
	Importer des données	GET DATA
	Fusion des fichiers	ADD FILES / MATCH FILES
	Création d'une variable	COMPUTE
	Recodification d'une variable	RECODE
	Sélection d'observations	SELECT IF
		USE ALL / FILTER BY
Analyse des données	Scission d'un fichier	SPLIT FILE
	Exploration des données	EXAMINE
	Test-t de Student	T-TEST
	Corrélation de Pearson	CORRELATIONS
	Test du khi-deux	CROSSTABS
Graphiques	Graphique de dispersion	GRAPH

9. LES RESSOURCES D'AIDE DU LOGICIEL SPSS

Nous terminons ce guide en vous présentant les ressources à votre disposition pour vous aider dans l'utilisation du logiciel SPSS, en distinguant les ressources d'aide internes et externes.

Les ressources d'aide internes au logiciel SPSS

Les rubriques SPSS. Les rubriques d'aide SPSS sont accessibles à partir du menu **AIDE** de l'**ÉDITEUR DE DONNÉES** ou du **NAVIGATEUR SPSS**. Elles offrent trois options : (a) **SOMMAIRE DE L'AIDE**, (b) **INDEX** et (c) **RECHERCHER** (voir Figure 9.1). La première option permet de consulter les rubriques par thème (e.g., la gestion des données). La deuxième permet de consulter les rubriques à partir d'un mot clé, comme « Bonferroni ». Les mots clé de l'**INDEX** sont présentés en ordre alphabétique. La troisième option permet de consulter l'aide suite à une recherche d'un mot dans le texte. Par exemple, si vous recherchez le mot « moyenne », toutes les rubriques d'aide contenant ce mot seront affichées.

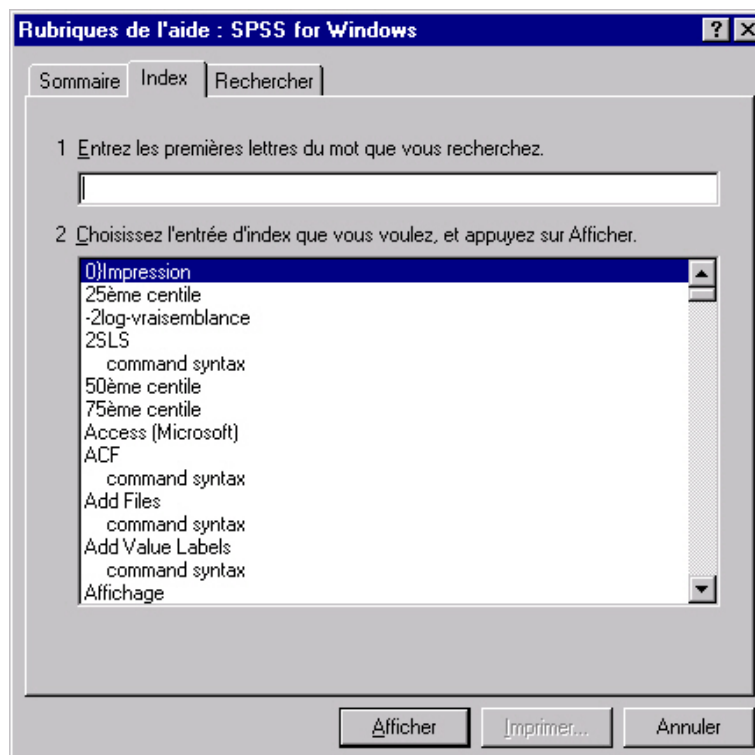


Figure 9.1. Rubriques d'aide du menu Aide

LE DIDACTICIEL SPSS. Il vous est probablement déjà arrivé, lors de la lecture d'un mode d'emploi ou durant l'utilisation d'un logiciel, de vous dire « Si quelqu'un était là pour me montrer comment faire, cela serait beaucoup plus simple ! ». Le didacticiel SPSS a été créé pour répondre à ce besoin puisqu'il présente, de façon visuelle et séquentielle, les étapes à suivre pour réaliser différentes opérations. Par exemple, le didacticiel peut illustrer, de façon très détaillée, la démarche à suivre pour créer un diagramme représentant vos données. Ce type de support ressemble beaucoup à certaines sections du présent guide, où une série de figures présente les étapes à suivre pour réaliser une procédure.

L'aide dans les fenêtres de dialogue. Deux types d'aide sont disponibles lorsque vous travaillez à l'intérieur d'une fenêtre de dialogue SPSS relative à une opération sur les données. La première consiste à cliquer sur l'onglet **AIDE** disponible dans la plupart des fenêtres. Ceci vous conduira directement à la rubrique d'aide relative à l'opération que vous réalisez (voir Figure 9.2). À l'ouverture de cette rubrique, vous apercevrez deux onglets, **COMMENT** et **SYNTAXE**, dans la partie supérieure droite. Le premier onglet illustre la démarche par étapes pour utiliser la fenêtre de dialogue, alors que le second présente l'ensemble des instructions de syntaxe disponibles pour cette commande.

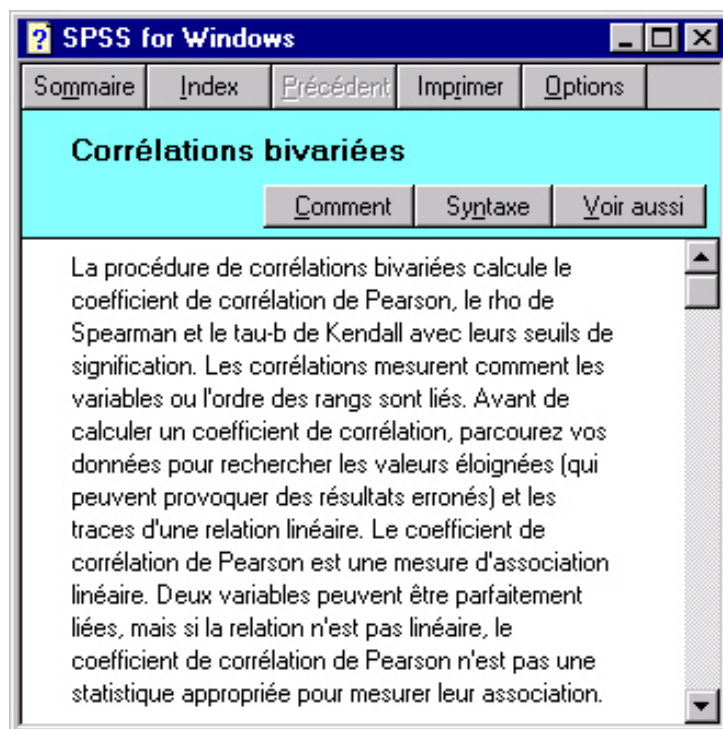


Figure 9.2. Aide pour la commande *Corrélations bivariées*

Le deuxième type d'aide s'applique lorsque vous désirez obtenir de l'information sur une partie de la fenêtre de dialogue. Pour l'obtenir, cliquez avec le bouton droit de votre souris directement sur la caractéristique de la fenêtre de dialogue pour laquelle vous avez besoin d'informations supplémentaires. Si disponibles, des informations seront alors affichées sur la commande.

L'assistant de résultat. L'assistant de résultat est une ressource d'aide mise à votre disposition pour vous diriger dans l'interprétation de vos résultats. Pour y accéder, vous n'avez qu'à sélectionner un tableau de résultats puis cliquez sur le bouton droit de votre souris. Une fois le menu affiché, sélectionnez **ASSISTANT DE RÉSULTAT** (voir Figure 9.3).

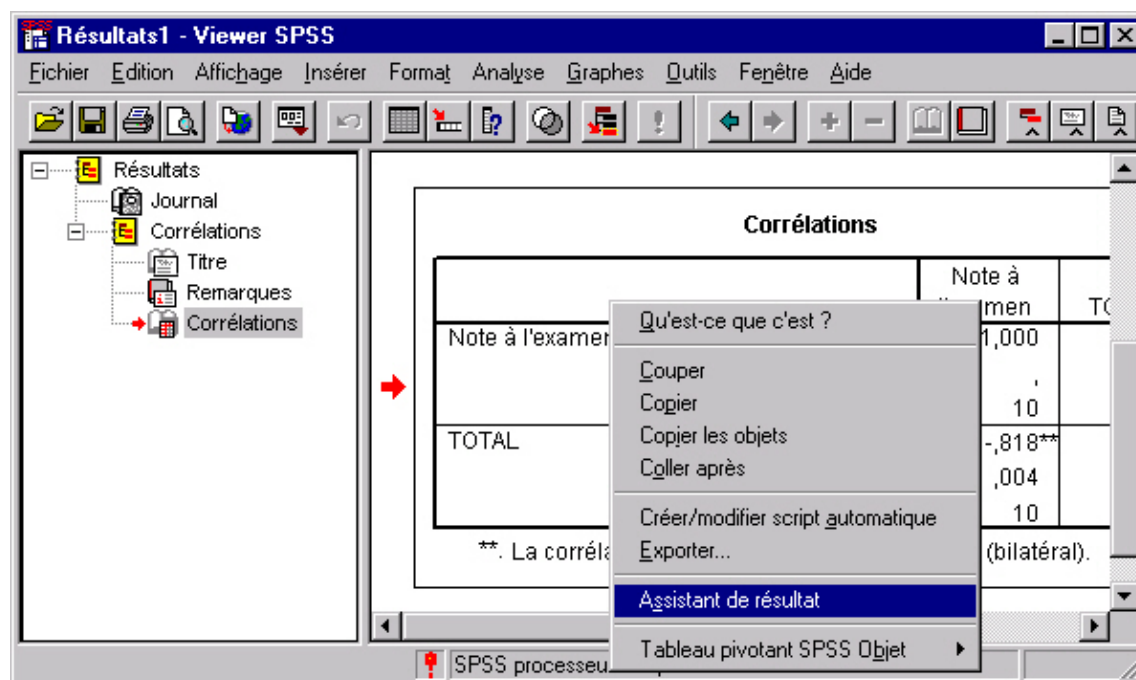


Figure 9.3. Affichage de l'Assistant de résultat pour un tableau de corrélation

L'assistant statistique. L'assistant statistique fournit aux utilisateurs de SPSS une aide permettant de choisir adéquatement le type d'analyse statistique à effectuer selon le but recherché. Pour utiliser l'assistant statistique, choisissez **ASSISTANT STATISTIQUE** dans le menu **AIDE** du logiciel. Des questions simples sont alors posées pour définir vos besoins d'analyse. Des exemples visuels sont également présentés par l'assistant pour vous aider.

Les ressources d'aide externes au logiciel SPSS

Les manuels de syntaxe SPSS en format PDF. Tous les manuels SPSS relatifs aux modules installés sont disponibles en format PDF et sont lisibles grâce au logiciel Acrobat Reader. La commande **SYNTAX GUIDE** du menu **AIDE** de l'**ÉDITEUR DE DONNÉES** permet l'accès aux différents manuels (voir Figure 9.4).

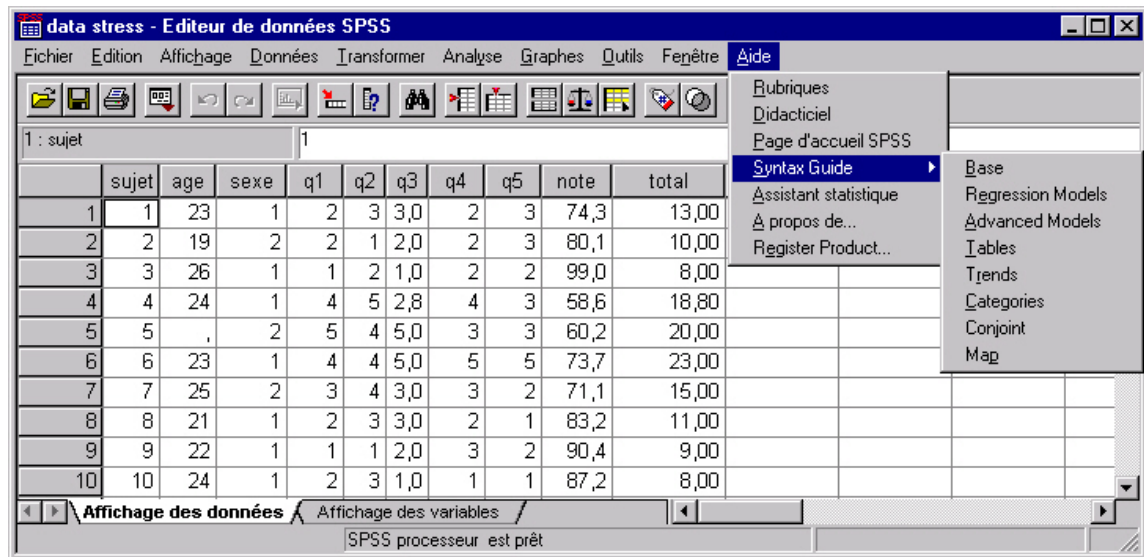


Figure 9.4. Manuels de syntaxe disponibles dans le logiciel SPSS

Le support technique SPSS. Les utilisateurs enregistrés de SPSS peuvent soumettre leurs questions techniques à un professionnel de la compagnie via le site Internet <http://www.spss.com>. Complétez le formulaire Internet en consultant la section **SUPPORT / SPSS BUSINESS INTELLIGENCE PRODUCT STATISTICAL SUPPORT / TECHNICAL SUPPORT CONTACT INFORMATION**. La réponse est envoyée dans les deux jours ouvrables. Vous pouvez également télécharger gratuitement les correctifs logiciels pour obtenir une mise à jour de votre version de SPSS. Les correctifs ne permettent pas de passer d'une version majeure à une autre (e.g., de la version 10 à la version 11). Ceux-ci sont disponibles sur le site Internet de SPSS <http://www.spss.com/tech/downloads/> dans la section **PATCHES, UPDATES AND UTILITIES / SPSS CLIENT & SERVER**.

Le Service d'Aide à la Recherche (SAR). À chaque trimestre, le SAR assiste les étudiants et les professeurs de l'École de psychologie de l'Université Laval dans leurs activités de recherche. Les services offerts sont principalement regroupés en deux catégories, soit (a) les dépannages hebdomadaires et (b) les ateliers de formation au logiciel SPSS. Les personnes intéressées à prendre

rendez-vous pour un dépannage ou qui veulent assister à une formation SPSS doivent s'inscrire sur le babillard du SAR au pavillon Félix-Antoine-Savard, local 042B. Pour plus de détails concernant les modalités reliées aux dépannages et connaître les dates des formations SPSS, consultez le site Internet du SAR.

Le Service d'Aide à la Recherche (SAR)

Pavillon Félix-Antoine-Savard, local 040

École de psychologie, Université Laval

Téléphone : (418) 656-2131 #8867

Télécopieur : (418) 656-3646

Courrier électronique : sar@psy.ulaval.ca

Site Internet : <http://www.psy.ulaval.ca/~sar/>

Le Service de Consultation Statistique (SCS) de l'Université Laval. Le département de mathématiques et statistique offre un service gratuit de consultation hebdomadaire en statistiques pour les étudiants chercheurs de la communauté universitaire. Pour plus d'informations, consultez le site Internet:

<http://www.mat.ulaval.ca/pages/scs/>

Les manuels SPSS. Plusieurs manuels sont disponibles mais seulement quelques-uns sont mis à jour selon les changements de versions du logiciel. Soyez attentif à ce point lorsque vous achetez un manuel pour vos besoins personnels.

Vous pouvez également vous procurer les livres développés par la compagnie SPSS en consultant, sur le site de la compagnie (<http://www.spss.com>), la section **STORE / BOOKS**.

10. RÉFÉRENCES

SPSS Corporation (site américain)

<http://www.spss.com>

SPSS Corporation (site français)

<http://www.spss.com/france/>

Information sur le *Graduate Pack*

<http://www.spss.com/spssbi/gradpack/>