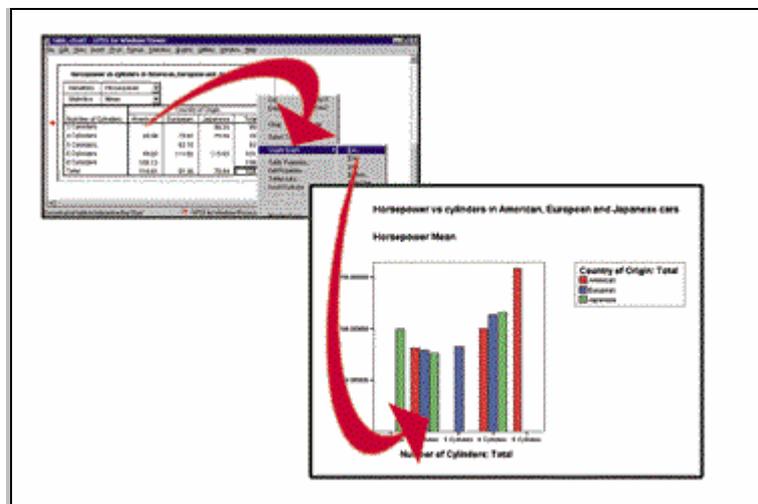


Analyse de données sur les ménages à l'aide de SPSS pour Windows



Manuel accompagnant le cours "Statistique pour la Politique Sociale, 1999"

assemblé par Klaus Röder - octobre 1999

Table de Matière

1. INTRODUCTION.....	1
1.1. LES CONVENTIONS DES STYLE DE TEXTE	2
1.2. COMMENT UTILISER LES MENUS	2
2. LA VUE D'ENSEMBLE.....	3
2.1. DEMARRER SPSS ET QUITTER DU LOGICIEL	3
2.2. LES TYPES DES FICHIERS EN SPSS	3
2.3. EXEMPLE: LE QUESTIONNAIRE ET L'ENQUETE EXEMPLAIRE:.....	4
2.4. LE FICHIER DE DONNÉES	5
2.5. SAUVEZ VOTRE FICHIER DE DONNEES	7
2.6. LISTER LES OBSERVATIONS	7
2.7. L'ANALYSE DE FREQUENCE DES DONNEES	10
3. LA TRANSFORMATION DES VARIABLES.....	14
3.1. DEFINIR LES VARIABLES	14
3.2. LES NOMS DES VARIABLES	14
3.3. LES TYPES DES VARIABLES.....	14
3.4. LES VALEURS MANQUANTS	15
3.5. LES ETIQUETTES DES VARIABLES.....	15
3.6. CALCULE DE VARIABLES	18
3.7. RECODAGE DE VARIABLES	18
3.8. TRANSPOSER LES VARIABLES	19
3.9. TRIER LES OBSERVATIONS.....	19
3.10. SELECTIONNER LES OBSERVATIONS	19
4. LA TRANSFORMATION DES FICHIERS DE DONNEES	24
4.1. FUSIONNER LES FICHIERS: CONSIDERATIONS GENERALES.....	24
4.2. FUSIONNER LES FICHIERS: AJOUTER DES OBSERVATIONS	25
4.3. FUSIONNER LES FICHIERS: SELON LES VARIABLES	26
4.4. AGREGER LES DONNEES	27
5. LES FONCTIONS POUR L'ANALYSE INITIALE DES DONNEES.....	31
5.1. LE TABLEAUX DE FREQUENCE.....	31
5.2. L'ANALYSE EXPLORATIVE	36
5.3. LES TABLEAUX CROISEE.....	41

6. LES FICHIERS DE SYNTAXE.....	45
6.1. CONSIDERATIONS GENERALES.....	45
6.2. UTILISER LA FENETRE DE SYNTAXE.....	45
6.3. EDITION D'UN FICHIER DE SYNTAXE.....	46
6.4. EXECUTER LES COMMANDES.....	48
6.5. LA SYNTAXE DES COMMANDES.....	48
7. LA PRESENTATION DES DONNEES.....	51
7.1. LES TABLEAUX RECAPITULATIVES.....	51
7.1.1. Tableaux de bord en lignes.....	51
7.1.2. Tableaux de bord en colonnes.....	54
7.2. LE MODULE TABLEAUX.....	59
7.2.1. Tableaux de base.....	59
7.2.2. Tableaux généraux.....	60
7.2.3. Les tableaux de réponses multiples.....	63
7.2.4. Tableaux de fréquences.....	64
7.3. LES GRAPHIQUES.....	70
7.3.1. Création d'un diagramme.....	71
7.3.2. Les principaux diagrammes.....	71
7.3.3. Edition d'un diagramme.....	71
7.3.4. Création de diagrammes interactifs depuis le menu Graphes.....	72
8. COMPARAISON ENTRE LES POSSIBILITÉS DE SPSS ET EXCEL.....	76
8.1. PREMIER EXEMPLE : LES STATISTIQUES DESCRIPTIVES.....	77
8.2. DEUXIÈME EXEMPLE: LES TABLEAUX CROISES.....	79

1. Introduction

L'objectif de ce cours est de familiariser les participants avec l'outil statistique SPSS . Ce logiciel est parmi les plus connus dans le monde informatique et par son ancienneté certainement un des meilleurs établis dans la symbiose ordinateur - analyse statistique. Avec l'évolution des micro-ordinateurs l'évolution des outils pour l'analyse statistique n'était pas moins dramatique. Actuellement SPSS existe dans la version plus récente no. 9. Même si les améliorations de cette version face aux prédécesseurs sur le plan théorique, des méthodes etc. n'était pas si impressionnant, les améliorations pour l'utilisateur étaient notables. L'ambiance de WINDOWS facilite l'utilisation de ce logiciel notamment. Surtout la présentation des graphiques est résolue d'une forme très avertie. L'accès commun d'utilisateur (ACU) de WINDOWS et le système d'aide intégré facilite l'utilisation et l'apprentissage. L'utilisateur avertie ne profitera pas tant que le débutant mais devrait réjouir surtout de l'intégration plus performante des autres logiciels, mais aussi de l'aide plus facilement accessible.

L'objectif donc de cette introduction ne sera pas d'approfondir les méthodes statistiques, ni de répéter plus que le essentiel des théories statistiques, pour cela existent l'étude de cas dans le contexte de ce cours et la littérature spécialisée. Cette introduction s'adresse au débutant ainsi comme au expérimenté en SPSS pour connaître le logiciels et faire expérience avec sa manipulation. L'objectif sera que le participant sera capable de maîtriser SPSS avec une certaine confiance et sera en état de résoudre les premières tâches d'analyse statistiques avec l'aide de SPSS.

Cette introduction ne remplace pas un manuel de référence de SPSS. L'utilisation de ces pages avec le système d'aide SPSS doit être suffisant quand même pour les besoins du cours actuel

1.1. Les conventions des style de texte

Pour reconnaître le texte dans son contexte, le style d'impression signifie le suivant:

[Alt]	Vous devez taper la touche du clavier ou sélectionnez un bouton ou une option
[ALT] + [Shift]	Vous devez taper les deux touche du clavier ensemble
[End],[]	Vous devez taper les deux touche un après l'autre
"nrques"	Vous devez taper ce texte entre les guillemets
[Fichier/ Ouvrir/Donnés]	Les commandes du menu (avec un sous- menu suivant le choix /Ouvrir). Vous pouvez utiliser les menus avec l'aide du souris ou du clavier
quest.sav	Un mot clé du système DOS/WINDOWS comme un nom d'un fichier ou le choix du menu
Nouvelle données	Pour le texte du système SPSS comme un nom du fenêtre
GET	Un mot clé de la syntaxe de SPSS comme une commande qui apparaît dans le texte.
GET	Le contenu d'un fichier de texte (soit données, résultat ou syntaxe).
4	Au début du chapitre Vous trouverez une description brève du contenu et du but du chapitre
*	Un tuyaux spécial; quelque chose considéré un peu difficile de comprendre
-	Quelque chose à faire, comme p.e. de remplir un questionnaire ou compléter une exercice guidé par Votre professeur.

1.2. Comment utiliser les menus

Pour récapituler; il existe deux méthodes pour choisir des éléments dans un menu:

Avec une souris: Cliquez sur le nom du menu que vous voulez. Le nom du menu apparaît en surbrillance et la liste des options de ce menu se déroule vers le bas. Vous pouvez alors afficher les options et cliquer sur l'option que vous souhaitez. (Ou cliquer sur le nom du menu en maintenant le bouton de la souris enfoncé. Sans relâcher le bouton, faites glisser le pointeur sur la liste des options de menu. Elles apparaissent chacune tour à tour en surbrillance au fur et à mesure que vous faites glisser la souris. Lorsque l'option désirée est en surbrillance, relâchez le bouton de la souris.)

Pour annuler: cliquez en dehors du menu et de la barre de menus.

Au clavier: Appuyez sur la touche **[ALT]** et la lettre soulignée (souvent la première) dans le nom du menu. Lorsque la liste des options du menu se déroule, appuyez sur la lettre soulignée dans le nom de l'option voulue.

Pour annuler: appuyez sur **[ECHAPPEMENT]**.

2. La vue d'ensemble

- 4 L'objectif de ce chapitre sera de Vous donner une vue globale de SPSS sans entrer dans les détails. Vous devez connaître les types des fichiers et les fenêtres associés.

2.1. Démarrer SPSS et quitter du logiciel

- Cliquez deux fois sur le icône de SPSS pour Démarrer
- Choisissez [Fichier/Quitter] pour sortir du logiciel

2.2. Les types des fichiers en SPSS

SPSS utilise les types de fichiers et des fenêtres suivants:

Editeur de données : Cette fenêtre affiche le contenu du fichier de données. Vous pouvez créer de nouveaux fichiers de données ou modifier des fichiers de données existants avec l'Editeur de données. La fenêtre Editeur de données s'ouvre automatiquement lorsque vous lancez une session SPSS. Vous ne pouvez avoir qu'un seul fichier de données ouvert à la fois.

Viewer/le Navigateur : Tous les résultats, tableaux et graphiques différents s'affichent dans le Viewer. Vous pouvez modifier les résultats et les enregistrer pour une utilisation ultérieure. Une fenêtre Viewer s'ouvre automatiquement la première fois que vous exécutez une procédure qui génère des résultats.

Les résultats sont affichés dans le Navigateur. Le panneau gauche du Navigateur contient la légende. Le panneau droit contient les tableaux statistiques, les graphiques et les textes. Vous pouvez utiliser les barres de défilement pour faire défiler les résultats. Vous pouvez également cliquer sur un item affiché dans la légende pour accéder directement au tableau ou au diagramme correspondant. Vous pouvez cliquer et faire glisser le bord droit de la fenêtre de légende pour modifier la largeur de la fenêtre.

Par l'utilisation du navigateur Vous pouvez lancer des Editeurs suivants:

- **Viewer en mode brouillon.** Vous pouvez afficher un résultat comme texte simple (plutôt que sous forme de tableaux pivotants interactifs) dans le Viewer en mode brouillon.
- **Editeur de tableaux pivotants :** Les résultats affichés dans les tableaux pivotants peuvent être modifiés de diverses manières grâce à l'Editeur de tableaux pivotants. Vous pouvez modifier le texte, permuter les données dans les lignes et colonnes, ajouter de la couleur, créer des tableaux multidimensionnels, et masquer et afficher les résultats de façon sélective.
- **Editeur de graphiques :** Vous pouvez modifier les graphiques et diagrammes à haute résolution dans les fenêtres de graphiques. Vous pouvez changer les couleurs, sélectionner des polices de taille ou de type différent, permuter les axes horizontal et vertical, faire pivoter les diagrammes de dispersion 3-D, et même changer le type de graphique.

- **Editeur de résultats** : Les résultats qui ne sont pas affichés dans les tableaux pivotants peuvent être modifiés grâce à l'Editeur de résultats. Vous pouvez modifier les résultats et changer les caractéristiques des polices (type, style, couleur, taille).

Editeur de syntaxe : Vous pouvez coller les choix des boîtes de dialogue dans une fenêtre de syntaxe, lorsque vos choix apparaissent sous forme de syntaxe de commande. Vous pouvez alors modifier la syntaxe de commande pour utiliser les fonctions spéciales de SPSS qui ne sont pas disponibles dans les boîtes de dialogue. Vous pouvez enregistrer ces commandes dans un fichier pour les utiliser dans des sessions SPSS ultérieures.

* Si vous avez plus d'une fenêtre Viewer ouverte, le résultat est dirigé vers la fenêtre Viewer désignée. Si vous avez plusieurs fenêtres Editeur de syntaxe ouvertes, la syntaxe de commande est collée dans la fenêtre Editeur de syntaxe **désignée**. Les fenêtres désignées sont indiquées par un point d'exclamation (!) dans la barre d'état. Vous pouvez changer les fenêtres désignées à tout moment.

Il ne faut pas confondre fenêtre désignée et fenêtre **active**, qui est la fenêtre sélectionnée. Si plusieurs fenêtres se chevauchent, la fenêtre active apparaît en premier plan. Si vous ouvrez une nouvelle fenêtre Editeur de syntaxe ou Viewer, cette fenêtre devient automatiquement la fenêtre active et la fenêtre désignée. Exemple: The questionnaire of the Munich Centre's sample survey

2.3. Exemple: Le questionnaire et l'enquête exemplaire:

Pour illustrer les contenus théoriques du texte nous appliquons ces connaissances sur un exemple très simple: L'enquête exemplaire parmi les participants du cours du Centre de Recyclage de Munich:

Le questionnaire

Pour connaître mieux l'information sur l'Allemagne dans l'extérieur répondez les questions suivants s.v.p.		
1:Sexe	masculin	<input type="checkbox"/>
	féminin	<input type="checkbox"/>
2.Age		<input type="checkbox"/>
3.Quel est le personnage allemand vivant plus connu dans votre pays (ne indiquez qu'une personne)	Gerhard Schröder	<input type="checkbox"/>
	Boris Becker	<input type="checkbox"/>
	Claudia Schiffer	<input type="checkbox"/>
	Franz Beckenbauer	<input type="checkbox"/>
	Helmut Koh	<input type="checkbox"/>

Pour connaître mieux l'information sur l'Allemagne dans l'extérieur répondez les questions suivants s.v.p.		
	Juliane Behrisch	<input type="checkbox"/>
	Aucun/Autre	<input type="checkbox"/>
	Ne sais pas	<input type="checkbox"/>
4. Quelle est la société allemande jouissant du plus grand prestige dans votre pays (ne indiquez qu'une société)	Mercedes-Benz	<input type="checkbox"/>
	Carl-Duisberg Gesellschaft	<input type="checkbox"/>
	Bayer	<input type="checkbox"/>
	MAN	<input type="checkbox"/>
	BMW	<input type="checkbox"/>
	Siemens	<input type="checkbox"/>
	Aucun/Autre	<input type="checkbox"/>
	Ne sais pas	<input type="checkbox"/>

Cet exemple n'est pas très réaliste mais sert de comprendre bien du fonctionnement de SPSS

Remplissez Votre copie personnelle du questionnaire, faites passer Votre copie au voisin à cote droite. Terminez après d'avoir rempli tous les questionnaires et puis utilisez SPSS pour analyser les données complets de Votre cours .

2.4. Le fichier de données

Ouvrez un fichier nouveau de données:

[Fichier/ Nouvelles /Donnés] puis un fenêtre de données apparait avec le nom:
Nouvelle données

Pour les cinq variables définissons les noms suivants:

Question	Nom de variable
No du questionnaire	nrques
1.	sexe
2.	age
3.	pealp_co
4.	soalp_co

Cliquez sur la première colonne du tableau:

[Donnés/ Définir la Variable]

puis un fenêtre de données apparait et Vous indiquez le nom de la variable, tapez:"nrques" dans le champ indiqué.

Ceci est Votre premier exercice interactive en SPSS. Ceci doit être fait ensemble avec et guidé par votre professeur. Essayez de suivre les pas indiqués exactement comme ils sont prescrits. Vous pouvez compléter ces pas quand même a Votre vitesse appropriée.

1. Remplissez ensuite les noms des autres variables. Ignorez les autres options de la boite de commande.



2. Entrez les valeurs des observations (chaque questionnaire représente une observation ou une ligne de donnés. Utilisez les touches de commande ou la souri pour mettre une cellule en surbrillance, puis tapez une valeur. Codez les réponses de la forme suivante:

- question 1: masculin =0, feminin =1,
- question 2 : Codez l'âge en années directement,
- question 3 et 4: première réponse=1, deuxième réponse =2 etc.

[ENTREE] accepte la valeur et fait descendre le curseur d'une ligne. [TABULATION] accepte la valeur et déplace le curseur d'une colonne vers la droite.

Figure 1: Définir les Variables

2.5. Sauvez Votre fichier de données

Choisissez

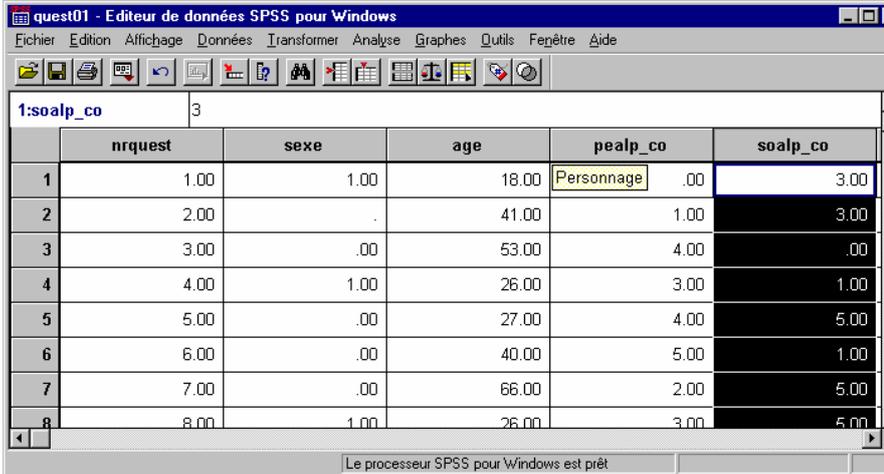
[Fichier/ Enregistrer sous...]

puis tapez le nom d fichier, par exemple tapez: *quest01*.

Le nom fichier sera enregistré comme *quest01.sav* .

Le tableau, la fenêtre de données doit apparaître ainsi.

Faites attention, ou et comment Vous sauvez Votre fichier. C'est une bonne idée de créer un enregistrement spécifique pour Vos exercices. Vous devez faire ceci dans le WINDOWS-Explorer.



	nrquest	sexe	age	pealp_co	soalp_co
1	1.00	1.00	18.00	Personnage .00	3.00
2	2.00	.	41.00	1.00	3.00
3	3.00	.00	53.00	4.00	.00
4	4.00	1.00	26.00	3.00	1.00
5	5.00	.00	27.00	4.00	5.00
6	6.00	.00	40.00	5.00	1.00
7	7.00	.00	66.00	2.00	5.00
8	8.00	1.00	26.00	3.00	5.00

Figure 2: Sauver le fichier de données

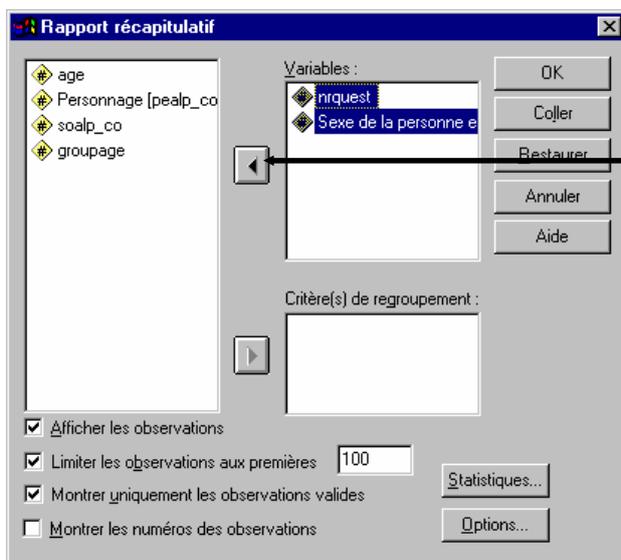
Généralement SPSS sauve les fichier dans l'enregistrement actuel, par exemple dans l'enregistrement ou SPSS maintien les fichiers du système SPSS. Ceci n'est pas un bon endroit pour les fichiers de travail.

2.6. Lister les observations

Vous voulez voir les données enregistrées, *listez* les observations enregistrés:

[Statistiques/ Récapituler /Lister les observations...]

puis choisissez les variables par l'utilisation de la fenêtre suivant:



Utilisez ce touche pour ajouter ou éliminer des variables de la liste, puis tapez [ENTRÉE]

Les résultats sont affichés dans le Viewer.
Vous pouvez utiliser le Viewer pour :

- parcourir les résultats ;
- afficher ou masquer les tableaux et diagrammes sélectionnés ;
- modifier l'ordre d'affichage des résultats en déplaçant les éléments sélectionnés ;
- déplacer des éléments entre le Viewer et d'autres applications.

Figure 3: Sélectionner les variables pour récapituler



Le Viewer est divisé en deux panneaux :

Le panneau de gauche du Viewer contient une légende du contenu du résultat.

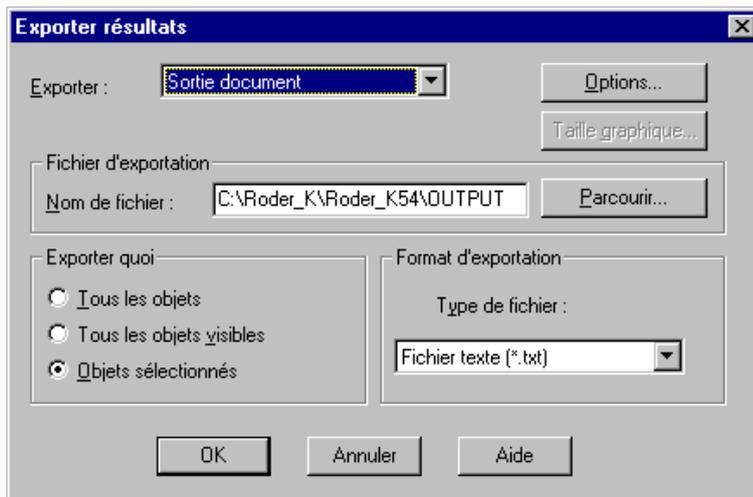
Le panneau droit contient les tableaux statistiques, les graphiques et les textes.

Vous pouvez sauver directement les résultats comme un fichier du Viewer (.spo) qui sera accessible par le Viewer uniquement: Si Vous voulez utiliser les résultats dans un éditeur du texte, exportez les résultats par [Fichier/Exporter...] et complétez la boîte de commande

Figure 4: Le SPSS Viewer

suivante:

Ou utiliser le Clipboard de WINDOWS ([Strg+C] et [Strg+V]) pour copier les résultats en texte et/ou graphique.



Le contenu du fichier résultat sera:

Figure 5: Exporter les résultats

Récapitulatif des observations(a)

	NRQUEST	SEXE	AGE	PEALP_CO	SOALP_CO
1	1.00	1.00	18.00	1.00	4.00
2	2.00	.00	41.00	2.00	4.00
3	3.00	.00	53.00	5.00	1.00
.....					
29	29.00	1.00	33.00	1.00	2.00
30	30.00	.00	47.00	2.00	1.00
N	30	30	30	30	30

φ
Limité aux 100 premières observations

2.7. L'analyse de fréquence des données

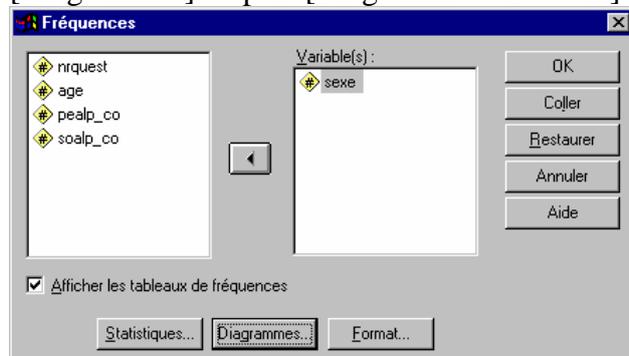
Depuis commencez avec la première analyse statistique: Combien des femmes et des homes trouvez Vous dans la groupe des personnes enquêtes ?

Depuis analysez la fréquence du sexe des répondeur:

[Analyse/ Statistiques Descriptives/ Fréquences ...]

depuis choisissez les variables par l'utilisation de la fenêtre suivant et choisissez l'option

[Diagrammes] et puis [Diagrammes en bâtons]



Avec le choix de **[Poursuivre]** et le choix de **[Coller]** dans la fenêtre précédent Vous obtenez le suivant: un fichier de syntaxe (syn01.sps) La commande sera exécuté par **[Exécuter/Tous]** dans la fenêtre de syntaxe après avoir marqué le texte de l commande ou pale bouton

Figure 6: Choisir les variables pour calculer les fréquences

Le Viewer enregistrera les résultats: Marquez et Exportez un fichier de résultat (res02.txt) et un fichier graphique (res01.htm), le graphique sera enregistré

comme res010.jpg et peut être utilisé directement dans un rapport

Le contenu du fichier de syntaxe:

```
FREQUENCIES
  VARIABLES=sexe
  /BARCHART PERCENT
  /ORDER ANALYSIS .
```

Le contenu du fichier du résultat (partiel):

Fréquences

Statistiques
SEXE

N	Valide	30
	Manquante	0

ç

SEXE

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	.00	18	60.0	60.0	60.0
	1.00	12	40.0	40.0	100.0
	Total	30	100.0	100.0	

¢

The content of the output file:

The two tables:

Statistics

SEX

N	Valid	30
	Missing	0

SEX

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	16	53.3	53.3	53.3
	2.00	14	46.7	46.7	100.0
	Total	30	100.0	100.0	

Le contenu du fichier graphique

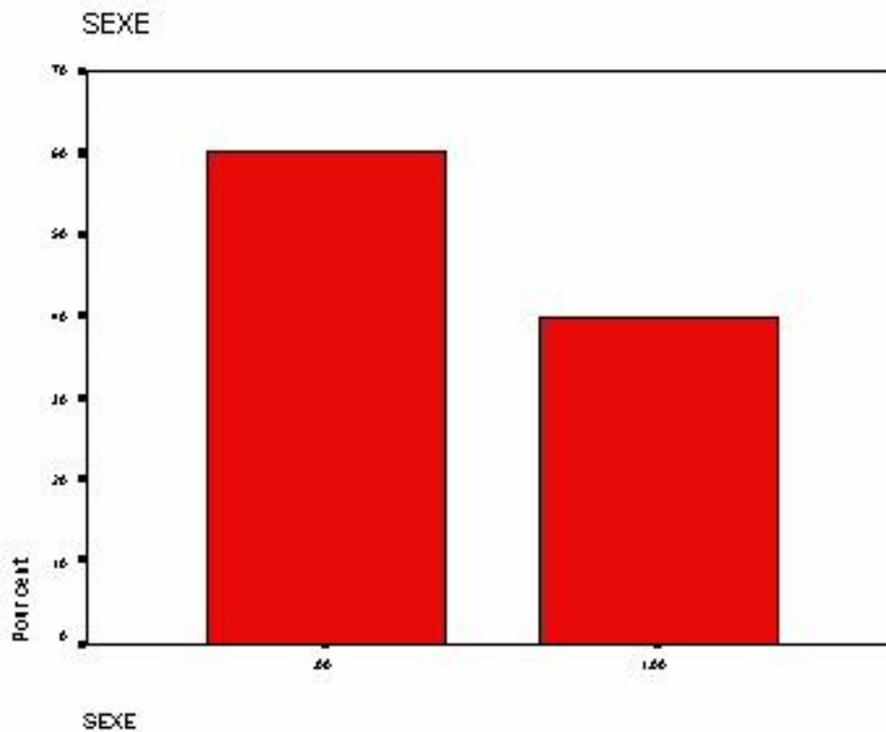


Figure 7: Résultat graphique de la commande FREQUENCIES

Exercice: Enquête Camerounaise Auprès des Ménages (ECAM):

Dans l'annexe Vous trouverez la description de l'Enquête Camerounaise Auprès des Ménages. Les données principales Vous trouverez sous le nom PSS01_01.SAV

Utilisez SPSS pour manipuler et analyser les données:

1. Commencez de lire les données. Comptez le nombre des variables et des observations
2. Listez le contenu des 50 observations
3. Imprimez la fréquences des hommes et des femmes dans l'univers.
4. Imprimez un Diagramme en bâtons de la variable *sexe*.
5. Imprimez un histogramme de la variable correspondant a l'âge et commentez le différences entre la distribution de cette variable contre la distribution normale (Utilisez l'Aide de SPSS

pour plus des informations) .

3. La transformation des variables

- 4 Avant d'être capable d'analyser les données, les variables doivent être traités d'une façon préparatoire. Ce chapitre s'occupe de ces transformations en détail. Vous devez connaître les possibilités et les raisons pour lesquelles ces transformations sont exécutés.

3.1. Définir les variables

Les variables, les noms des colonnes, sont définies de manière implicite ou explicite.

Pour définir de nouvelles variables ou modifier le format des variables existantes, cliquez deux fois sur un en-tête de colonne dans l'éditeur de données, ou sélectionnez une cellule dans la colonne et choisissez Définir la variable dans le menu Données. De manière implicite Vous tapez directement les données d'une observation dans une colonne vide. Le nom de la variable sera attribué par SPSS: la première variable nommée sera *VAR00001*, la deuxième *VAR00002* etc.

3.2. Les noms des variables

Dans SPSS, les noms de variables peuvent comporter jusqu'à 8 caractères. Le premier caractère du nom d'une variable défini par vous doit obligatoirement être une lettre ou @. Les caractères qui suivent peuvent être des lettres, des chiffres, les caractères point ou soulignement ou un des signes \$#@, à la seule réserve près que le dernier caractère ne peut pas être un point. Les lettres majuscules et minuscules sont équivalentes dans les noms de variables. Vous ne pouvez pas utiliser *ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, WITH* comme noms de variables.

3.3. Les types des variables

SPSS utilise deux types principales des variables:

Variable numérique

Une variable numérique ne contient que des valeurs numériques. Les variables numériques ont de nombreux formats d'affichage. Vous pouvez spécifier le nombre de chiffres et de décimales qui s'afficheront ainsi que le format des symboles monétaires, des dates et des heures. Aucun de ces formats n'affecte la façon dont SPSS représente les données au niveau interne. Les variables numériques peuvent être converties en variables alphanumériques, ou vice-versa, à l'aide des fonctions *STRING* et *NUMBER* fournies par les commandes de transformation. Dans la syntaxe des commandes SPSS, vous pouvez spécifier des valeurs numériques sous forme de nombres entiers, de fractions décimales, ou en notation scientifique (7.29735E-3).

et **Variable chaîne**

Les valeurs d'une variable chaîne ne sont pas numériques, et ne sont donc pas utilisées dans les calculs. Dans les listes de variables des boîtes de dialogue, les variables alphanumériques de 8 caractères ou moins sont identifiées par le symbole <, et les variables alphanumériques longues, de plus de 8 caractères, par le symbole >. La longueur maximale d'une variable alphanumérique est déterminée lors de la définition de la variable. Les variables alphanumériques peuvent être converties en variables numériques, et vice-versa, à l'aide des fonctions *NUMBER* et *STRING* du langage SPSS. Dans la syntaxe des commandes SPSS, la valeur des variables chaîne est spécifiée entre apostrophes ou guillemets.

3.4. Les valeurs manquants

Il existe deux types de valeurs manquantes dans SPSS, les valeurs manquantes spécifiées et les valeurs manquantes par défaut. La valeur manquante par défaut, qui s'affiche comme un point, est affectée à toutes les cellules numériques vides qui se trouvent dans un rectangle de données. Vous pouvez aussi affecter des valeurs qui identifient une information manquante pour des raisons spécifiques, puis ordonner à SPSS d'identifier ces valeurs comme des valeurs manquantes. Les valeurs manquantes par défaut et les valeurs manquantes spécifiées sont spécialement prises en charge par des procédures statistiques de SPSS et par des commandes de transformation de données.

3.5. Les étiquettes des variables

Une étiquette de variable contient un texte descriptif pouvant comporter jusqu'à 120 caractères et permettant d'identifier plus facilement les résultats SPSS. L'étiquette de variable apparaît généralement à côté du nom de la variable, et sert à décrire la signification ou l'origine de la variable. (Dans de nombreuses procédures les résultats n'indiquent qu'une quarantaine de caractères de l'étiquette de la variable.)

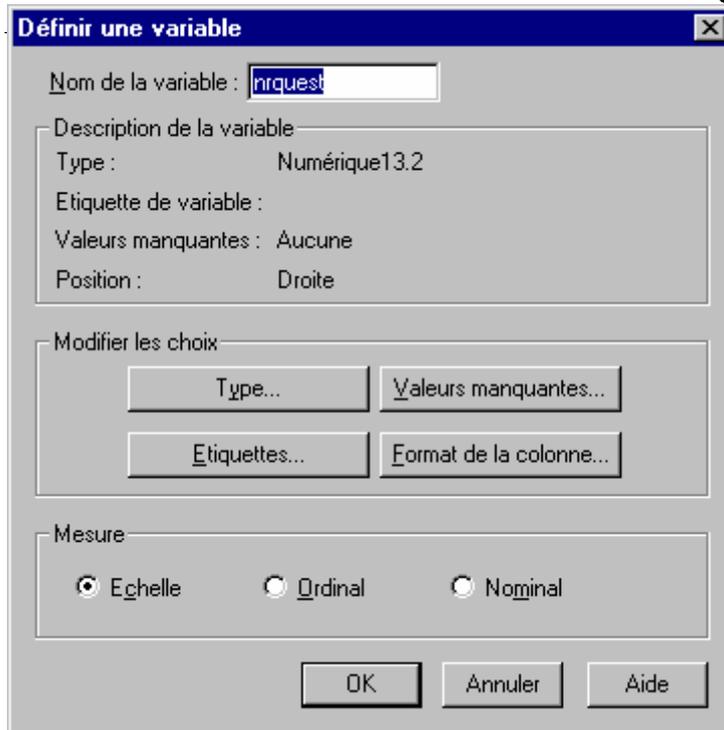


Figure 8: Définir une variable

Tous ces informations ainsi comme le format de la colonne Vous définissez par la boîte de dialogue apparaissant après avoir choisi le menus [Données/ Définir une variable..]

Cette commande est disponible si l'éditeur des données est active. Vous pouvez définir une nouvelle variable ou changer le format d'une variable active. Tapez le nom de la variable ou remplace le nom existant.

Pour définir une variable

Activez la fenêtre de l'Editeur de données.

Double-cliquez sur le nom de la variable en haut de la colonne ou

cliquez n'importe où dans la colonne de la variable, et choisissez dans les menus :

[Données/Définir une variable...]

Entrez le nom d'une variable.

Choisissez le niveau de mesure.

Cliquez sur [Type...] pour changer le format de la donnée.

Cliquez sur [Etiquettes...] pour affecter des étiquettes descriptives de la variable et des valeurs.

Cliquez sur [Valeurs manquantes...] pour spécifier des codes pour les valeurs manquantes.

Cliquez sur [Format de colonne...] pour changer la largeur de la colonne ou son alignement.

Maintenant Vous devrez éditer les étiquettes des variables.

Dans ce document toutes les commandes de SPSS sont affichés faire semblant que un fichier de syntaxe avait été édité et exécuté. -

Comme débutant et dans les premières chapitres Vous sélectionnez et exécutez certainement plutôt les commandes à l'aide des menus. Graduellement Vous devez Vous familiariser avec les fichiers de syntaxe de SPSS et 'écrire' un programme de SPSS doit être la tâche la plus normale pur Vous. Au moins c'est l'espoir du auteur de ce document. *

S'il Vous plaît utilisez les étiquetés dans le fichier de syntaxe suivant:

Exemple: l'enquête exemplaire du CdeM

Nom du fichier: ex03_1.sps

```
*****
* fichier syntaxe: ex03_1.sps
* fichier données début: QUE01.sav
* fichier données fin : QUE03_1.sav
* fichier résultats : -----
* fichier graphique : -----
* fichier résultats : -----
* commentaire      :
* commentaire      : K54
* auteur           : CDG-FZ München/klaus röder
*****
GET FILE='C:\Roder_K\Roder_K54\QUE01.SAV'.
VARIABLE LABELS
nrquest 'numéro du questionnaire'
/sexe 'Sexe'
/age 'Age'
/pealp_co 'Personnage allemand vivant plus connu'
/soalp_co 'Société allemande jouissant du plus grand prestige'
.
VALUE LABELS
/sexe 0 'masculin' 1 'féminin'
/pealp_co 1 'Gerhard Schröder'
          2 'Boris Becker'
          3 'Claudia Schiffer'
          4 'Franz Beckenbauer'
          5 'Helmut Kohl'
          6 'Juliane Behrisch'
          7 'Aucun/Autre'
          8 'Ne sais pas'
/soalp_co 1 'Mercedes-Benz'
          2 'Carl-Duisberg Gesellschaft'
          3 'Bayer'
          4 'MAN'
          5 'BMW'
          6 'Siemens'
          7 'Aucun/Autre'
          8 'Ne sais pas'
.
SAVE OUTFILE=' C:\Roder_K\Roder_K54\que03_1.sav'.
```

3.6. Calcule de variables

Menu: [Transformer/Calculer]

Langage de commandes: *COMPUTE, IF*

Pour calculer les valeurs de toutes les observations, tapez le nom d'une variable destination unique. Il peut s'agir d'une variable existante, ou d'une nouvelle variable devant être ajoutée à votre fichier de données courant. Créez une Expression pour les valeurs devant être affectées à cette variable. Si vous préférez calculer les valeurs de façon sélective, pour des observations répondant à une certaine condition logique, sélectionnez également le bouton de commande **[Si]**. Si la variable destination est ou sera une variable chaîne vous devez sélectionner **[Type et Etiquette]** pour indiquer sa longueur. Si vous calculez une nouvelle variable, vous pouvez sélectionner **[Type et Etiquette]** pour spécifier une étiquette pour cette variable. Pour créer une expression, vous pouvez soit coller ses divers éléments dans la zone de texte *Expression*, soit placer le curseur dans cette zone et taper ces éléments. Collez les noms de variable en les copiant à partir de la liste source située à gauche. Collez les nombres et les opérateurs de la calculatrice. Collez les fonctions en les copiant à partir de la liste de fonctions située à droite. Les constantes alphanumériques doivent être comprises entre guillemets ou apostrophes. Les constantes numériques doivent être tapées au format américain, en utilisant le point comme séparateur décimal.

3.7. Recodage de variables

Menu: [Transformer/ Recoder/ Recodage des variables] et

Menu: [Transformer/ Recoder/ Création des variables]

Langage de commandes: *RECODE*

Faites passer les variables que vous souhaitez recoder selon les mêmes spécifications dans la zone de liste Variables. Sélectionnez Anciennes et nouvelles valeurs et spécifiez la manière dont la/les variables doivent être recodées.

Si vous ne souhaitez recoder que les observations satisfaisant à une condition logique, sélectionnez **[Si]** et spécifiez la condition. Si l'option de transformation choisie (sous Edition/Préférences) est *Calculer les valeurs immédiatement* (choix initial), SPSS lira le fichier de données et réalisera la transformation de données dès que vous aurez sélectionné **[OK]**. Si l'option *Calculer les valeurs avant utilisation* est choisie, SPSS effectuera la transformation lorsque la procédure suivante lira les données, et l'éditeur de données ne montrera pas la transformation avant ce moment.

3.8. Transposer les variables

Menu: [Transformer /Transposer]

Langage de commandes: *FLIP*

Sélectionnez au moins une variable. Ces variables sont des colonnes dans le fichier de travail, et elles seront transposées dans des observations (lignes) dans le nouveau fichier de travail. Les variables non sélectionnées seront éliminées. SPSS indique le nom des variables dans le fichier transposé, et crée une variable *CASE_LBL* contenant les anciens noms de variables. Si le fichier de travail contient une variable d'identification ou de nom ayant des valeurs uniques, vous pouvez faire passer celle-ci dans la zone Variable de nom, et ses valeurs seront converties en noms de variables pour le fichier de données transposé.

3.9. Trier les observations

Menu: [Transformer /Trier les observations]

Langage de commandes: *SORT CASES*

Sélectionnez une variable et faites-la passer dans la zone Trier par. Pour que le tri soit effectué dans l'ordre décroissant, sélectionnez la variable figurant dans la zone Trier par et appuyez sur Décroissant.

Vous pouvez sélectionner des variables de tri supplémentaires. Chaque variable de tri supplémentaire spécifie l'ordre des observations du groupe ayant des valeurs identiques pour toutes les variables de tri précédentes.

Croissant: L'ordre croissant est représenté par le symbole (A) dans la liste des variables de tri. Les variables alphanumériques sont triées dans l'ordre alphabétique.

Décroissant: L'ordre décroissant est représenté par le symbole (D) dans la liste des variables de tri. Les variables alphanumériques sont triées dans l'ordre alphabétique inverse.

3.10. Sélectionner les observations

Menu: [Données/Sélectionner les observations]

Langage de commandes: *SELECT IF; SAMPLE; FILTER; USE*

Choisissez l'une des options de la liste Sélectionner:

Toutes les observations indique que vous souhaitez traiter toutes les observations du fichier de travail. Si la condition est remplie indique que vous souhaitez traiter uniquement les observations satisfaisant à une condition logique. Si vous avez spécifié une condition, elle apparaîtra à côté du bouton de commande **[Si]**. Dans le cas contraire, ou si vous souhaitez changer la condition spécifiée, cliquez sur **[Si]**. Cette option permet de définir la variable filter_\$.

Choisissez l'une des options de la liste Les observations non sélectionnées sont:
[Filtrées] exclut du traitement les observations non sélectionnées, mais les conserve dans le fichier de données. Vous pouvez entrer à nouveau dans cette boîte de dialogue et modifier l'option de Sélectionner pour restaurer les observations.**[Supprimées]** élimine les observations du fichier de données. Pour les récupérer, vous devez Ouvrir à nouveau le fichier de données initial (en supposant que vous ne l'avez pas remplacé par la version réduite).

Exemple: l'enquête exemplaire du CdeM

Appliquons les transformations initiales aux données du questionnaire exemplaire. Les résultats seront affichés dans un fichier de syntaxe.

1. Créez les groupe d'âge en pas des dix années.
2. Appliquez les étiquettes pour chaque valeur, ' ≤ 10 ' pour valeur 1, ' $10 < 20$ ' pour valeur 2 etc.
3. Introduisez des valeurs manquantes avec des valeurs et des étiquettes
4. Triez les valeurs selon les groupes d'âge
5. Appliquez un filtre, sélectionnez les observations plus jeunes que 30

Nom du fichier: ex03_2.sps

```
* fichier syntaxe: ex03_2.sps
* fichier données début: QUE03_1.sav
* fichier données fin : QUE03_2.sav
* fichier résultats : -----
* fichier graphique : -----
* fichier résultats : -----
* commentaire      :
* commentaire      :
* auteur           : CDG-FZ München/klaus röder
*****
```

```
GET FILE='C:\Roder_K\Roder_K54\QUE03_1.SAV'.
COMPUTE gr_age = age .
VARIABLE LABELS gr_age "groupe d'âge" .
IF (age <= 10) gr_age = 1 .
```

```
IF (age > 10 & age <= 20) gr_age = 2 .
IF (age > 20 & age <= 30) gr_age = 3 .
IF (age > 30 & age <= 40) gr_age = 4 .
IF (age > 40 & age <= 50) gr_age = 5 .
IF (age > 50 & age <= 60) gr_age = 6 .
IF (age > 60 & age <= 70) gr_age = 7 .
IF (age > 70      ) gr_age = 8 .
MISSING VALUES
/gr_age (-1 -2).
VALUE LABELS
/gr_age
-1 'aucune réponse'
-2 "valeur d'age inconnu"
1 ' <= 10'
2 '10< <=20'
3 '20< <=30'
4 '30< <=40'
5 '40< <=50'
6 '50< <=60'
7 '60< <=70'
8 '70< '
.
EXECUTE .
SAVE OUTFILE=' C:\Roder_K\Roder_K54\QUE03_2.sav'.
SORT CASES BY gr_age .
SPLIT FILE
  BY gr_age .

SORT CASES BY
  age (A) .

USE ALL.
COMPUTE filter_$=(gr_age <= 3).
VARIABLE LABEL filter_$ 'gr_age <= 3 (FILTER)'.

VALUE LABELS filter_$ 0 'Non sélectionné' 1 'Sélectionné'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .
FILTER OFF.
USE ALL.
SELECT IF(gr_age <= 3).
EXECUTE .
```



Carl Duisberg
Gesellschaft e.V

SAVE OUTFILE=' C:\Roder_K\Roder_K54\QUE03_3.sav'.

Exercice: Enquête Camerounaise Auprès des Ménages (ECAM):

Dans l'annexe Vous trouverez la description de l'Enquête Camerounaise Auprès des Ménages. Les données principales Vous trouverez sous le nom CHS01_01.SAV

Utilisez SPSS pour manipuler et analyser les données:

Transformez les données:

1. Attribuez les étiquettes des variables et des valeurs selon le questionnaire
2. Définissez les valeurs manquants pour au moins une variable.
3. Choisissez le format des colonnes approprié.
4. Définissez une variable nouvelle *gr_âge*
5. Récodez les valeurs de *gr_age* pour chaque groupe comprenant 10 années (*gr_age* = 1 pour age entre 0 et 10 années, *gr_age* = 2 pour age entre 10 et 20 années etc.)
6. Eliminez les observations pour les personnes plus âgées que 20 ans et calculez la fréquence de la variable *Fréquentation scolaire*.
7. Recodez la variable *sex* (1 -> 2 et 0 -> 1).
8. Triez les observations suivant la variable *groupe d'âge*.

Sauvez votre travail. Sauvez le fichier des données modifié, sauvez le fichier de syntaxe, sauvez la partie désirée du fichier de résultat. Choisissez des noms des fichiers appropriés

c

4. La transformation des fichiers de données

4 Très souvent les fichiers de données doivent être manipulés pour répondre aux besoins de l'analyste. Soit que quelques observations doivent être exclues de l'analyse, soit que plusieurs fichiers doivent être fusionnés. Ce chapitre s'occupe des fusions des fichiers de données, des éliminations des variables et des observations qui ont été traitées auparavant. Vous devez connaître les possibilités et les raisons pour lesquelles ces transformations sont exécutées.

4.1. Fusionner les fichiers: Considérations générales

Menu: [Données/Fusionner les fichiers]

Langage de commandes: *MATCH FILES*

La boîte de dialogue avec la fonction Fusionner les fichiers met en œuvre deux fichiers de données: le fichier de travail, et un fichier de données enregistré sous le format SPSS (le fichier externe). Deux types de fusion sont possibles par les menus suivants:

Ajouter des observations

ajoute toutes les observations du fichier externe à la fin du fichier de travail. Par défaut, le nouveau fichier comprendra toutes les variables communes aux deux fichiers. Vous pouvez choisir d'inclure les variables présentes dans un seul fichier; les observations provenant de l'autre fichier afficheront des valeurs manquantes par défaut ou vides pour les variables de ce type. Vous pouvez également exclure des variables.

Ajouter des variables

ajoute toutes les variables du fichier externe, ou certaines d'entre elles, aux observations correspondantes du fichier de travail. Vous pouvez spécifier les observations correspondantes de l'une des trois manières suivantes:

Ordre des observations.

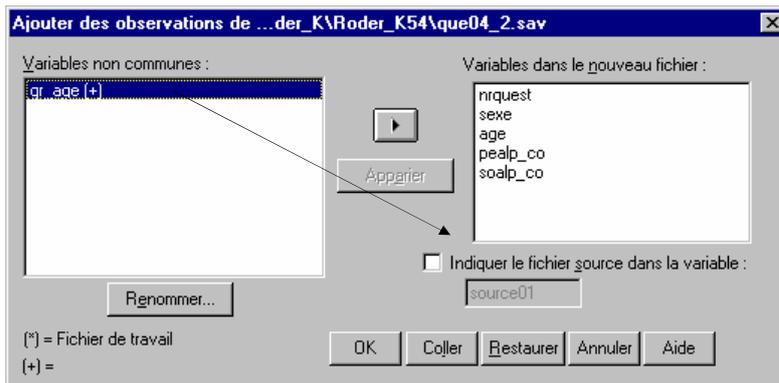
Si vous ne spécifiez pas de clé d'appariement, les observations seront simplement disposées dans le même ordre que dans les fichiers. Cela est risqué, à moins que vous ne soyez certain que les ordres se correspondent d'un fichier à l'autre.

Clé(s) d'appariement.

Si vous spécifiez une clé d'appariement, les observations seront positionnées selon les valeurs de cette clé, dont les valeurs d'identification sont généralement uniques. Au besoin, vous pouvez utiliser plusieurs clés d'appariement pour spécifier la correspondance entre les observations. Les deux fichiers doivent être triés par la/les clés d'appariement. Les observations non appariées dans l'un quelconque des fichiers entreront dans le nouveau fichier de travail, avec des valeurs manquantes par défaut ou des valeurs vides pour les variables issues de l'autre fichier.

Consultation de table.

Si vous spécifiez une clé d'appariement et indiquez que l'un des fichiers est une table, les observations seront positionnées en fonction de la clé, mais les observations apparaissant uniquement dans la table ne contribueront pas au nouveau fichier de travail. La table fournit



seulement des données pour les nouvelles variables ajoutées à l'autre fichier. Les deux fichiers doivent être triés par la/les clés d'appariement.

Figure 10: Ajouter des observations

4.2. Fusionner les fichiers: Ajouter

des observations

Menu: [Données/Fusionner les fichiers /Ajouter des observations]

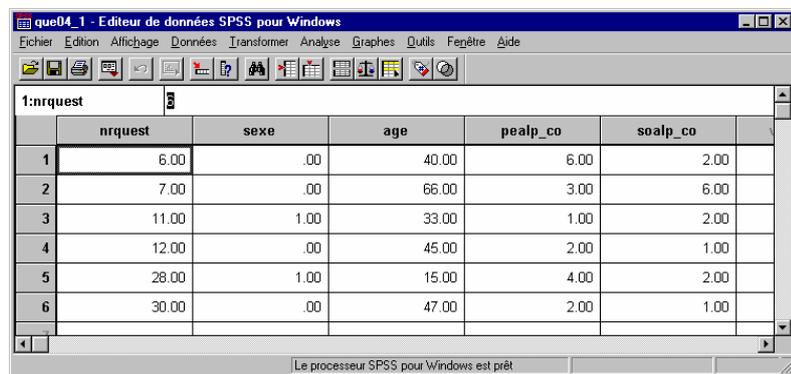
Langage de commandes:
ADD FILES

Une boîte de dialogue initiale vous demande d'identifier le fichier dans lequel vous souhaitez obtenir des observations.

La zone de liste située à droite: [Variables dans le nouveau fichier], comprend toutes les variables communes aux deux fichiers. Si cette liste vous satisfait, appuyez sur **[OK]**.

La zone de liste située à gauche, Variables non communes, contient les variables n'ayant pas pu être appariées d'un fichier à l'autre en raison de différences de noms ou de types. Les variables repérées par le symbole (*) se trouvent dans le fichier de travail. Celles repérées par un (+) sont dans le fichier externe.

Si vous sélectionnez [Identifier le fichier source dans la variable], une variable ayant la valeur 0 pour les observations provenant du fichier de travail initial et la valeur 1 pour celles issues du fichier externe sera ajoutée au nouveau fichier de travail. Pour modifier le nom d'une variable dans la liste des Variables non communes, sélectionnez la variable puis appuyez sur **[Renommer]**.



	nrquest	sexe	age	pealp_co	soalp_co
1	6.00	.00	40.00	6.00	2.00
2	7.00	.00	66.00	3.00	6.00
3	11.00	1.00	33.00	1.00	2.00
4	12.00	.00	45.00	2.00	1.00
5	28.00	1.00	15.00	4.00	2.00
6	30.00	.00	47.00	2.00	1.00

Figure 9: Le fichier QUE04_1.SAV

4.3. Fusionner les fichiers : Selon les variables

Menu:
[Données/Fusionner les fichiers /Ajouter des variables]
Langage de commandes:
MATCH FILES,
MATCH TABLES

	nrquest	sexe	age	pealp_co	soalp_co	gr_age	source01
1	1.00	2.00	1
2	2.00	5.00	1
3	3.00	6.00	1
4	4.00	3.00	1
5	6.00	.00	40.00	6.00	2.00	.	0
6	7.00	.00	66.00	3.00	6.00	.	0
7	11.00	1.00	33.00	1.00	2.00	.	0
8	12.00	.00	45.00	2.00	1.00	.	0
9	20.00	3.00	1
10	22.00	7.00	1
11	28.00	1.00	15.00	4.00	2.00	.	0
12	29.00	4.00	1
13	30.00	.00	47.00	2.00	1.00	.	0

Figure 13: Résultat de **MATCH FILES**
Figure 11: Résultat de **ADD FILES..** Fichier
Fichier QUE04_4.SAV QUE04_03.SAV

Spécifiez le type d'appariement:

Si vous souhaitez appairer les observations des deux fichiers dans l'ordre séquentiel seulement, ne cochez pas la case [Appairer les observations] sur les clés des fichiers triés. Si les deux fichiers sont triés selon une ou plusieurs clés d'appariement et si vous voulez appairer les observations en fonction de leur valeur: sélectionnez [Appairer les observations sur les clés des fichiers triés], choisissez l'une des trois options de type d'appariement, puis transférez la/les clés de la liste des Variables communes à la liste Clés d'appariement.

Spécifiez les variables pour le nouveau fichier de travail:

La zone de liste située à droite, Variables dans le nouveau fichier, contient initialement toutes les variables du fichier de travail, repérées par le symbole (*), suivies de toutes les variables nommées différemment dans le fichier externe, repérées par le symbole (+). Vous pouvez ajouter des variables à cette liste, ou en éliminer.

La zone de liste située à gauche, Variables communes, présente les variables qui normalement ne seront pas comprises dans le nouveau fichier de travail car elles portent le même nom que certaines variables du fichier de travail. A l'origine, toutes ces variables proviennent du fichier externe, et sont repérées par le symbole (+). Les variables repérées par le symbole (*) proviennent du fichier de travail (ancien).

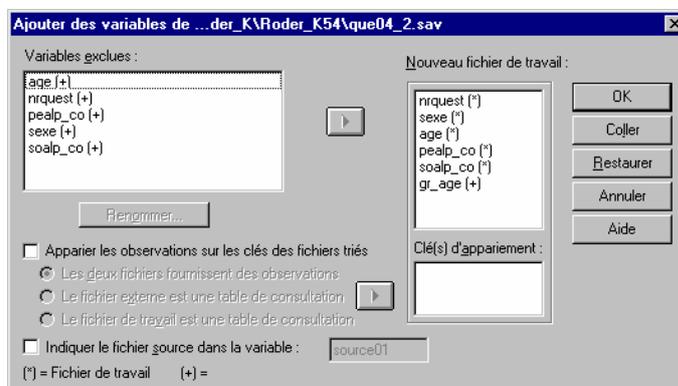


Figure 12: Boite de dialogue **MATCH FILES**

Si vous sélectionnez [Identifier le fichier source dans la variable], une variable ayant la valeur 0 pour les observations provenant du fichier de travail initial et la valeur 1 pour les observations issues du fichier externe, et pouvant donc être contestable, est ajoutée au nouveau fichier de travail. (Cette option est désactivée lorsque vous sélectionnez l'un des choix de table de consultation, puisque les observations ne sont jamais obtenues d'une table de consultation.)

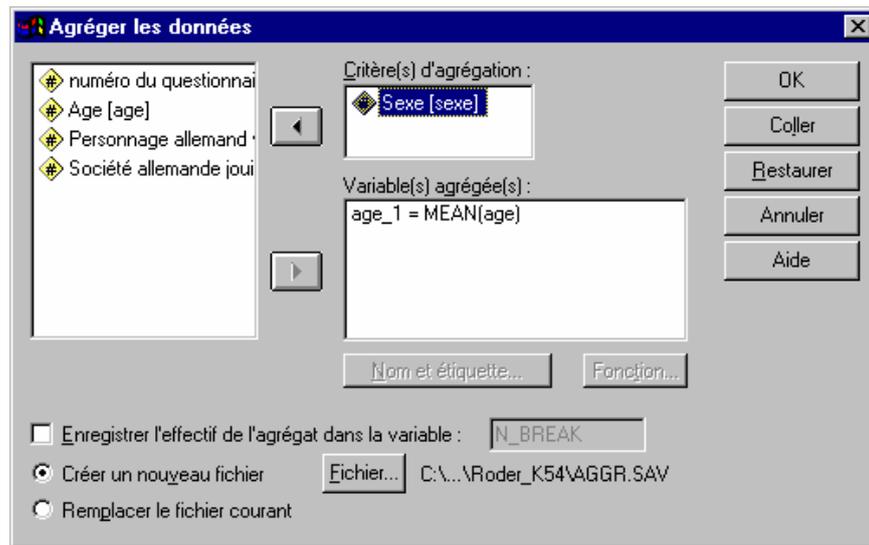
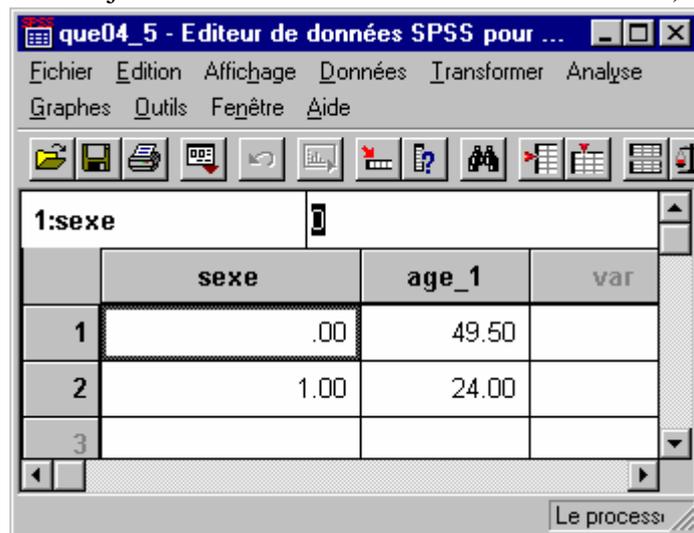


Figure 15: Agréger les données

Le fichier externe ou fichier actif est une table codée : Une table codée ou un fichier de consultation de table, est un fichier dans lequel les données de " chaque observation " peuvent s'appliquer à plusieurs observations de l'autre fichier. Par exemple, si un des fichiers contient des informations sur les différents membres d'une famille (sexe, âge, niveau scolaire) et l'autre des informations générales sur la famille (revenu global, taille, habitat), vous pouvez utiliser le fichier sur la famille comme fichier de consultation et appliquer les informations générales à chaque membre de la famille dans le fichier fusionné.



	sexe	age_1	var
1	.00	49.50	
2	1.00	24.00	
3			

Figure 14: Résultat de AGGREGATE

Fichier QUE04_5.SAV

4.4. Agréger les données

Menu: [Données/Agréger]

Langage de commandes: *Aggregate*

Sélectionnez un ou plusieurs Critères d'agrégation. Sélectionnez la/les variables agrégées voulues dans le nouveau fichier agrégé. Si vous voulez utiliser une autre fonction Aggregate

que la Moyenne, appuyez sur Fonctions. Enfin, cliquez sur Spécification de fichier pour indiquer la destination du fichier de données agrégées.

SPSS affecte un nom et une étiquette de variable à chaque nouvelle variable du fichier agrégé. Appuyez sur [Nom & Etiquette] pour modifier l'un des deux.

Sélectionnez [Enregistrer l'effectif de l'agrégat] dans la variable pour créer une nouvelle variable comprenant le nombre d'observations de l'agrégat. Vous pouvez modifier le nom de cette variable.

Appliquons les transformations aux données du questionnaire exemplaire. Les résultats seront affichés dans un fichier de syntaxe.

Exemple: l'enquête exemplaire du CdeM

Nom du fichier: ex04_1.sps

```
*****
* fichier syntaxe: ex04_1.sps
* fichier données début: QUE04_1.sav,QUE04_2.sav
* fichier données fin : QUE04_3.sav,QUE04_4.sav,QUE04_5.sav
* fichier résultats : -----
* fichier graphique : -----
* commentaire :
* commentaire : K54
* auteur : CDG-FZ München/klaus röder
*****
GET
 FILE='C:\Roder_K\Roder_K54\que04_1.sav'.
EXECUTE .
ADD FILES /FILE=*
 /FILE='C:\Roder_K\Roder_K54\que04_2.sav'.
EXECUTE.
SAVE OUTFILE='C:\Roder_K\Roder_K54\que04_3.sav'
 /COMPRESSED.
MATCH FILES /FILE=*
 /FILE='C:\Roder_K\Roder_K54\que04_2.sav'
 /RENAME (age pealp_co sexe soalp_co = d0 d1 d2 d3)
 /BY nrquest
 /DROP= d0 d1 d2 d3.
EXECUTE.
SAVE OUTFILE='C:\Roder_K\Roder_K54\que04_4.sav'
 /COMPRESSED.
AGGREGATE
 /OUTFILE='C:\Roder_K\Roder_K54\que04_5.sav'
 /BREAK=sexe
 /age_1 = MEAN(age).
```

Exercice: Enquête Camerounaise Auprès des Ménages (ECAM):

Dans l'annexe Vous trouverez la description de l'Enquête Camerounaise Auprès des Ménages. Les données principales Vous trouverez sous le nom CHS01_01.SAV et CHS01_01.SAV les fichiers individuelles et CHS02_01.SAV le fichier du ménage.

Utilisez SPSS pour manipuler et analyser les données:

1. Ajoutez les observations des deux fichiers individuelles. Sauvez le fichier
2. Créez une variable BRID: l'identifiant du ménage ($\text{No du ménage} + 1000 * \text{No du Ilot} + 1.000.000 * \text{No de la strate}$) pour le fichier du ménage et pour le fichier individuel
3. Créez une variable BRIDP: l'identifiant individuel ($\text{No d'ordre} + 100 * \text{No du ménage} + 100.000 * \text{No du Ilot} + 100.000.000 * \text{No de la strate}$) pour le fichier individuel
4. Triez les fichiers sur les Identifiant
5. Ajoutez les variables des deux fichiers individuelles au fichier du ménage. Traitez le fichier du ménage comme table codée

Page laissée vide intentionnellement

5. Les fonctions pour l'analyse initiale des données

- 4 Dans ce chapitre Vous apprenez d'utiliser des fonctions statistiques de SPSS. Cette introduction se limite d'expliquer la théorie statistique très brièvement. La manipulation de SPSS et l'interprétation des résultats sera l'objectif principale. Vous devez connaître les possibilités et l'utilisation des fonctions statistiques descriptives et être capable d'interpréter et manipuler les résultats de ces méthodes en SPSS.

5.1. Le tableaux de fréquence

Menu: [Analyse /Statistiques descriptives/ Fréquences]

Langage de commandes: *Frequencies*

Pourquoi: Vous voulez analyser la fréquence des réponses aux questions 3. et 4. du questionnaire exemplaire.

Faites passer une ou plusieurs variables dans la liste Variable(s). Un tableau de fréquences sera produit pour chaque variable de cette liste et indiquera l'effectif pour chaque valeur.

Sélectionnez [Statistiques] pour demander des statistiques univariées.

Sélectionnez [Diagrammes] pour demander des diagrammes en barres ou des histogrammes.

Sélectionnez [Format] pour gérer l'ordre des valeurs dans le tableau, la disposition de celui-ci sur la page, l'affichage d'un index, et l'affichage des étiquettes de valeur.

Si vous désélectionnez la zone Afficher les tableaux de fréquences, aucun tableau de fréquences ne s'affichera, mais toute statistique ou tout diagramme que vous demandez apparaîtra.

[Statistiques..]

Quartiles Valeurs qui divisent les observations en quatre groupes de taille égale.

Moyenne Moyenne arithmétique; somme divisée par le nombre d'observations.

Partition en classes égales

Médiane Valeurs qui divisent les observations en plusieurs groupes détaille égale.

Valeur qui est supérieure à la moitié des données réelles et inférieure à l'autre moitié. 50ème centile.

Centiles Valeurs au-dessus et au-dessous desquelles se situent certains pourcentages des observations.

Mode	Valeur le plus souvent rencontrée. Si plusieurs valeurs à la fois présentent la plus grande fréquence d'occurrences, chacune d'entre elles est un mode. La procédure Fréquences signale le plus petit de ces modes multiples.
Somme	Somme ou total des valeurs, pour toutes les observations comportant des valeurs non manquantes.
Ecart-type	Mesure de dispersion autour de la moyenne, égale à la racine carrée de la variance. L'écart-type est mesuré avec les mêmes unités que la variable.
Minimum	Plus petite valeur d'une variable numérique.
Asymétrie	Mesure d'asymétrie d'une distribution. Une asymétrie positive signifie que les valeurs les plus extrêmes sont supérieures à la moyenne; une asymétrie négative indique que les valeurs les plus extrêmes sont inférieures à la moyenne.
Variance	Mesure de dispersion autour de la moyenne, égale à la somme du carré des écarts par rapport à la moyenne divisée par l'effectif moins un. La variance se mesure en unités qui correspondent au carré de celles de la variable.
Maximum	Plus grande valeur d'une variable numérique.
Aplatissement	Mesure de la lourdeur de la queue d'une distribution par rapport à une distribution normale. Un aplatissement positif indique que les queues extrêmes comptent plus d'observations que dans une distribution normale ayant la même variance.
Intervalle	Différence entre les valeurs maximale et minimale d'une variable numérique; le maximum moins le minimum.
E-T de la moyenne	Erreur-type de la moyenne de la caractéristique de la distribution pour toutes les observations de la cellule. Disponible pour les caractéristiques et leurs totaux.

[Diagrammes..]

Un [diagramme en bâtons] représente la fréquence de chaque valeur sous forme de barre distincte. Les diagrammes en barres sont appropriés pour les variables mesurées au niveau nominal, ou pour les variables présentant peu de valeurs distinctes. Les valeurs dont l'occurrence est nulle n'apparaissent pas. Un histogramme affiche la fréquence des intervalles de valeurs sous forme de barre distincte.

Un [diagramme en secteurs] montre la participation de chaque partie à l'ensemble. Chaque secteur du diagramme correspond à un groupe défini par une simple variable de regroupement. Un diagramme en bâtons montre l'effectif de chaque valeur ou de chaque modalité sous la forme d'un bâton distinct, ce qui vous permet de comparer les modalités visuellement.

Les [histogrammes] sont appropriés pour les variables dont les valeurs adjacentes peuvent être regroupées de façon significative. Un espace est laissé pour les intervalles dans lesquels aucune observation n'est comprise. Pour vérifier plus aisément si la variable est distribuée normalement, une courbe normale peut être superposée à l'histogramme.

[Format..]

Le tableau de fréquences peut être disposé en fonction des valeurs réelles des données, ou en fonction du comptage (fréquence d'occurrence) de ces valeurs, dans l'ordre croissant ou décroissant.

Exemple: l'enquête exemplaire du CdeM

Appliquons la fonction aux données du questionnaire exemplaire. Les résultats des opérations en boîtes de dialogue seront affichés dans un fichier de syntaxe, un fichier de résultat et un fichier graphique.

Nous voulons connaître la distribution des réponses concernant le personnage le plus connu et la société la plus connue

```
*****
* Syntax file      : ex05_1.sps
* Data File charged : que03_2.sav
* Output File     : res05_1.txt
* Graphics File   : gra05_1.jpg,gra05_2.jpg, gra05_3.jpg
* Comments       :
* Comments       :
* Comments       :
* Comments       : K54
* Author         : CDG-FZ München/klaus röder
*****
* Syntax File
GET
  FILE='C:\Roder_K\Roder_K54\Que03_2.sav'.
EXECUTE .
FREQUENCIES
  VARIABLES=sexe age pealp_co soalp_co gr_age
  /NTILES= 4
  /STATISTICS=MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT
  /BARCHART  FREQ
  /ORDER  ANALYSIS .
```

Résultats sélectionnés:

Statistiques

		Sexe	Age	Personnage allemand vivant plus connu	Société allemande jouissant du plus grand prestige	groupe d'âge
N	Valide	30	30	30	30	30
	Manquante	0	0	0	0	0
Moyenne		.4000	38.0333	3.5667	3.4667	4.2333
Médiane		.0000	35.0000	4.0000	3.0000	4.0000
Mode		.00	26.00	1.00	2.00	3.00
Asymétrie		.430	.811	-.174	.158	.518
Erreur std. d'asymétrie		.427	.427	.427	.427	.427
Aplatissement		-1.950	.374	-1.350	-1.723	-.476
Erreur std. d'aplatissement		.833	.833	.833	.833	.833
Centiles	25	.0000	25.2500	2.0000	2.0000	3.0000
	50	.0000	35.0000	4.0000	3.0000	4.0000
	75	1.0000	48.5000	5.0000	6.0000	5.2500

a Il existe de multiples modes

Personnage allemand vivant plus connu

Valide		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
	Gerhard Schröder	6	20.0	20.0	20.0
	Boris Becker	4	13.3	13.3	33.3
	Claudia Schiffer	3	10.0	10.0	43.3
	Franz Beckenbauer	6	20.0	20.0	63.3
	Helmut Kohl	6	20.0	20.0	83.3
	Juliane Behrisch	5	16.7	16.7	100.0
	Total	30	100.0	100.0	

groupe d'âge

Valide		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
	<= 10	1	3.3	3.3	3.3
	10< <=20	4	13.3	13.3	16.7
	20< <=30	7	23.3	23.3	40.0
	30< <=40	7	23.3	23.3	63.3
	40< <=50	4	13.3	13.3	76.7
	50< <=60	2	6.7	6.7	83.3
	60< <=70	3	10.0	10.0	93.3
	70<	2	6.7	6.7	100.0
	Total	30	100.0	100.0	

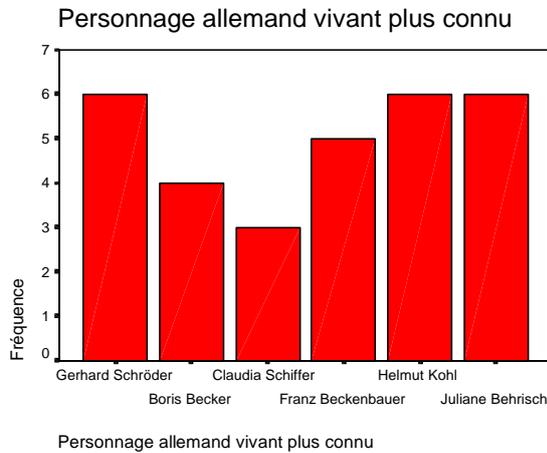


Figure 16: Diagramme pour groupes d'âges

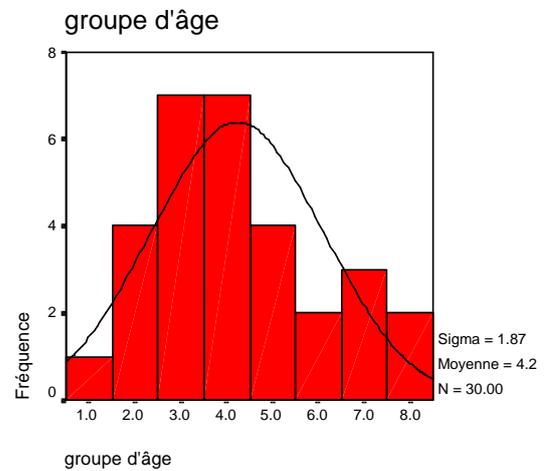


Figure 17: Histogramme pour Personnages

..

Par l'analyse fréquence Vous constatez le suivant:

La distribution de l'âge des répondeurs diffère dans les critères suivants de la distribution normale: Aplatissement et Asymétrie sont légèrement positive (voyez l'explication dans ce chapitre). Même ainsi la groupe des répondeurs est assez jeune. Des tableaux de fréquence vous constatez le fait que la compagnie Bayer est inconnu parmi les enquêtés est les personnages les plus populaires sont les politiciens et Juliane Behrisch.

5.2. L'analyse explorative

Explorer permet d'obtenir des statistiques et des affichages pour toutes vos observations ou pour des groupes d'observations distincts. Vous pouvez ainsi obtenir des boxplots, des graphiques tiges et feuilles, des histogrammes, des tests de normalité, des estimations de position robustes, des tableaux de fréquences, ainsi que d'autres statistiques et graphiques explicatifs.

Vous pouvez, par exemple, étudier par des graphique la distribution des réponses aux questions 3. et 4. du questionnaire exemplaire et comparer les réponses selon groupe d'âge et sexe. Pour chacun des groupes, vous pouvez voir si les réponses suivent une distribution plus ou moins normale(pour ce type de question cela est sans signification), et si les quatre variances son égales.

Menu: [Analyse /Statistiques descriptives/ Explorer]

Langage de commandes: *EXAMINE*

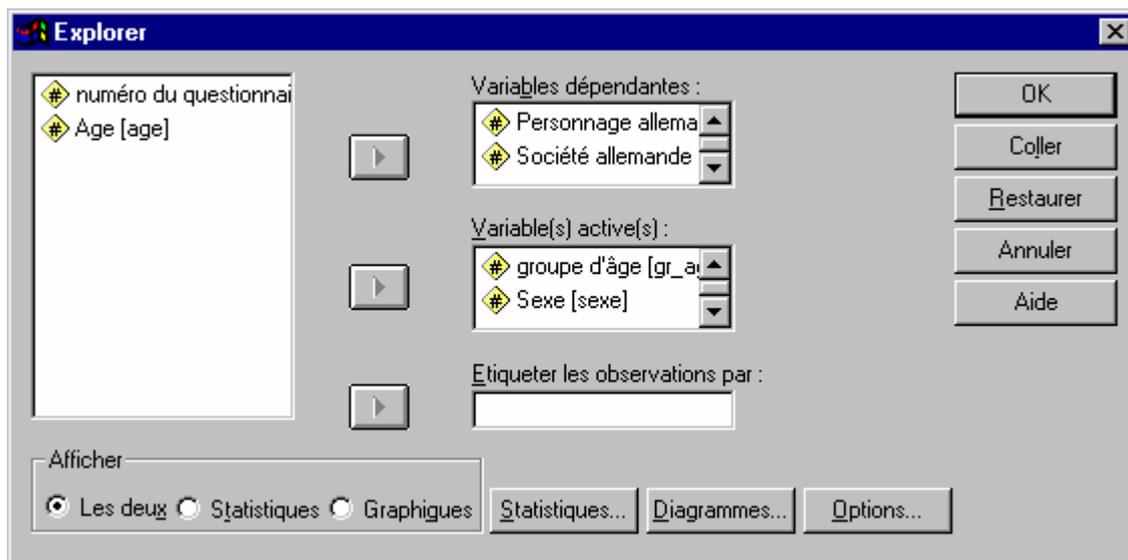


Figure 18: Explorer

Faites passer au moins une variable dans la zone de liste Variable(s) dépendante(s). Les statistiques et les graphiques qui résulteront de cette commande décriront la/les distributions des variables de cette liste. Vous pouvez faire passer un ou plusieurs critères de regroupement dans la Variable(s) active(s). Les valeurs des variables actives définissent les groupes d'observations, ce qui fait l'objet d'une description séparée.

Sélectionnez l'un des choix de la zone Afficher pour choisir si les résultats seront des graphiques, des statistiques, ou les deux (tout). Ces choix activent les boutons de commande

Graphiques et Statistiques:

[Diagrammes] vous permet de gérer les résultats sous forme de graphiques.

[Statistiques] vous permet de demander des caractéristiques et un listage d'observations avec les 5 plus petites valeurs et les 5 plus grandes.

[Options] vous permet de déterminer le traitement des valeurs manquantes.



Figure 19: Explore les données

Exemple: l'enquête exemplaire du CdeM

Appliquons la fonction aux données du questionnaire exemplaire. Les résultats des opérations en boites de dialogue seront affichés dans un fichier de syntaxe, un fichier de résultat et un fichier graphique.

```
*****
* Syntax file      : ex05_2.sps
* Data File charged : ques03.sav
* Output File     : res_05_2.lst
* Graphics File   : gra_05_3.jpg,gra_05_4.jpg
* Comments       :
* Comments       :
* Comments       :
* Comments       : K54
* Author         : CDG-FZ München/klaus röder
*****
* Syntax File
GET
FILE='C:\Roder_K\Roder_K53\Quest03.sav'.
EXAMINE
```

```

VARIABLES=soalp_co BY sexe
/PLOT BOXPLOT STEMLEAF
/COMPARE GROUP
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
EXAMINE
VARIABLES=pealp_co soalp_co BY gr_age sexe
/PLOT BOXPLOT STEMLEAF NPLOT
/COMPARE GROUP
/PERCENTILES(5,10,25,50,75,90,95) HAVERAGE
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
EXAMINE
VARIABLES=pealp_co BY gr_age
/PLOT BOXPLOT STEMLEAF NPLOT
/COMPARE GROUP
/MESTIMATORS HUBER(1.339) ANDREW(1.34) HAMPEL(1.7,3.4,8.5) TUKEY(4.685)
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.

```

*** Résultats (partiel):**

Tests de normalité

	groupe d'âge	Kolmogorov-Smirnov		Signification	Shapiro-Wilk		Signification
		Statistique	ddl		Statistique	ddl	
Personnage allemand vivant plus connu	10< <=20	.151	4	.			
	20< <=30	.313	7	.037	.784	7	.037
	30< <=40	.243	7	.200	.800	7	.048
	40< <=50	.441	4	.			
	60< <=70	.385	3	.			

* Il s'agit d'une borne inférieure de la signification réelle.

a Correction de signification de Lilliefors

b Personnage allemand vivant plus connu est une constante lorsque groupe d'âge = <= 10. Elle a été omise.

c Personnage allemand vivant plus connu est une constante lorsque groupe d'âge = 50< <=60. Elle a été omise.

d Personnage allemand vivant plus connu est une constante lorsque groupe d'âge = 70< . Elle a été omise.

La procédure Explorer produit des résumés statistiques et des affichages graphiques pour toutes vos observations ou séparément pour des groupes d'observations. Il existe plusieurs raisons pour utiliser la procédure Explorer : le filtrage de données, l'identification des valeurs éloignées, la description, la vérification d'hypothèses et la caractérisation des différences

parmi les sous populations (groupes d'observations). Le filtrage de données peut vous indiquer les valeurs inhabituelles, les valeurs extrêmes, les trous dans les données ou d'autres particularités. L'exploration des données peut vous aider à déterminer si les techniques statistiques que vous envisagez d'utiliser pour l'analyse de vos données sont appropriées. L'exploration peut indiquer que vous avez besoin de transformer les données si la technique nécessite une répartition gaussienne. Vous pouvez également choisir d'utiliser des tests non paramétriques.

Diagrammes et statistiques. moyenne, médiane, moyenne tronquée à 5 %, erreur standard, variance, Ecart-type, minimum, maximum, intervalle, intervalle interquartile, asymétrie et aplatissement avec leurs erreurs standard, intervalle de confiance pour la moyenne (et niveaux de confiance spécifiés), centiles, Huber, Andrews, Hampel, Tukey, les cinq plus grandes et cinq plus petites valeurs, le Kolmogorov-Smirnov avec un niveau de signification Lilliefors pour tester la normalité, et la statistique Shapiro-Wilk. Boîtes à moustaches, diagrammes en tige et feuille, histogrammes, diagrammes de répartition gaussienne, et dispersion/niveaux avec le test de Levene et les transformations.

Boîtes à moustaches : ces alternatives contrôlent l'affichage de boîtes à moustaches quand vous avez plus d'une variable dépendante. Niveaux de critère génère un affichage séparé pour chaque variable dépendante. Dans un affichage, les boîtes à moustache sont données pour chacun des groupes définis par une variable active. Dépendantes génère un affichage séparé pour chaque groupe défini par une variable active. Dans un affichage, les boîtes à moustaches s'affichent côte à côte pour chaque variable dépendante. Cet affichage est particulièrement utile lorsque les différentes variables représentent une seule caractéristique mesurée à des moments différents.

Caractéristique : le groupe caractéristiques vous permet de choisir les diagrammes Tige et feuilles et les Histogrammes.

Diagrammes feuille et tige

...

Société allemande jouissant du plus grand prestige Stem-and-Leaf Plot for SEXE= masculin

Frequency Stem & Leaf

6.00	1 . 000000
4.00	2 . 0000
.00	3 .
1.00	4 . 0
1.00	5 . 0
6.00	6 . 000000

Stem width: 1.00
Each leaf: 1 case(s)

Personnage allemand vivant plus connu Stem-and-Leaf Plot for
GR_AGE= 30< <=40

Frequency	Stem & Leaf
4.00	0 . 1144
3.00	0 . 666

Stem width: 10.00
Each leaf: 1 case(s)

...

Graphes de répartition gaussiens avec tests : affiche les diagrammes de probabilité gaussienne et les résidus. La statistique de Kolmogorov-Smirnov avec un niveau de signification Lilliefors pour le test de normalité s'affiche. La statistique Shapiro-Wilk est calculée pour des échantillons avec 50 observations ou moins.

Dispersion/niveau avec test de Levene : contrôle les transformations de données pour les diagrammes de dispersion par niveau. Pour tous les diagrammes de dispersion par niveau, la pente de la ligne de régression et les tests de Levene portant sur l'homogénéité de la variance s'affichent. Si vous sélectionnez une transformation, les tests de Levene sont basés sur les données transformées. Si aucune variable active n'est sélectionnée, les diagrammes de dispersion par niveau ne sont pas produits. Estimation d'exposants produit un diagramme des logs naturels des intervalles interquartile opposés au logs naturels des médianes pour toutes les cellules, en même temps qu'une estimation de la transformation de l'exposant pour arriver à des variances égales dans les cellules. Un diagramme de dispersion par niveau aide à déterminer l'exposant pour qu'une transformation stabilise (rende plus égales) les variances entre groupes. Transformation Exposant : vous permet de sélectionner une des alternatives de l'exposant, en suivant éventuellement les recommandations de l'estimation de l'exposant et de produire les diagrammes des données transformées. L'intervalle interquartile et la médiane des données transformées sont dessinés. Sans transformation produit des diagrammes de données brutes. Ceci est équivalent à une transformation avec une puissance de 1.

Boîtes à moustaches

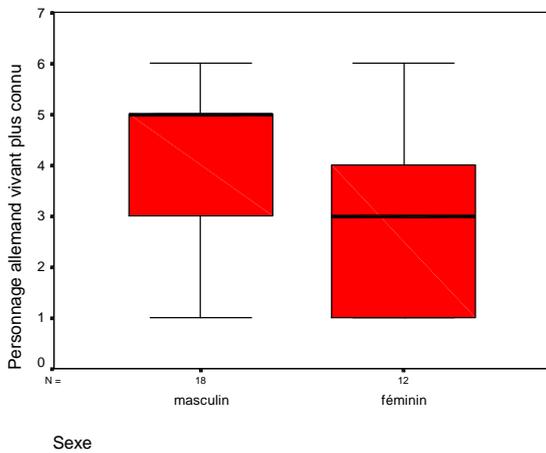


Figure 20: Personnages par sexe

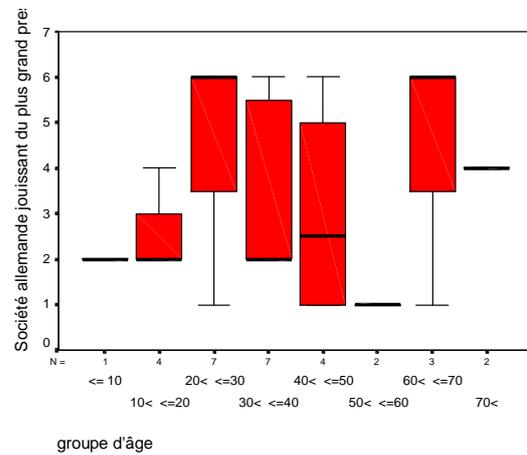


Figure 21: Personnages par groupes d'âges

L'analyse montre le comportement des différents groupe d'âge très divergent, comme les réponse des femmes et des hommes diverge aussi. Les statistiques des variables indiquent rien de nouveau. La hypothèse que la distribution de `pralp_co` sur `gr_age = 4` soit normale peut être rejeté. Les graphes indiqués dans le fichier résultats montre les valeurs observés comparés avec une ligne théorique indiquée par la distribution normale. L'analyse explorative part de l'analyse univariée (une variable est analysée) à l'analyse multivariée (l'interaction entre plusieurs variables).

5.3. Les tableaux croisé

Tableaux croisés sert à compter le nombre d'observations ayant différentes combinaisons de valeurs de deux variables ou plus, à calculer des statistiques principales et à effectuer des tests. Par exemple, si vous voulez savoir s'il existe un lien entre l'âge et la connaissance d'un personnage. Ou encore, vous pouvez étudier la relation qui existe entre la connaissance d'un personnage en faisant la distinction entre hommes et femmes.

Les variables utilisées pour former les catégories sur lesquelles porteront les calculs doivent présenter un nombre limité de valeurs distinctes. Utilisez la commande *RECODER* pour définir des fourchettes des valeurs.

Menu: [Analyse /Statistiques descriptives/ Tableaux croisés...]

Langage de commandes: *CROSSTABS*

Faites passer une ou plusieurs variables dans la zone Ligne(s), et une ou plusieurs variables dans la zone Colonne(s). Les catégories d'une variable de ligne forment les lignes du tableau, et les catégories d'une variable de colonne forment ses colonnes. Un tableau est produit pour chaque variable de ligne combinée à chaque variable de colonne. Les Couches optionnelles, décrites ci-après, permettent de spécifier des variables de contrôle.

[Statistiques] vous permet de demander des statistiques bivariées relatives aux variables de ligne et de colonne.

[Cellules] vous permet de gérer le contenu des cellules constituant le tableau.

[Format] vous permet de gérer l'étiquetage et l'aspect des tableaux.

Sélectionnez [Supprimer les tableaux] pour voir uniquement les statistiques, sans les tableaux proprement dits.

Exemple: l'enquête exemplaire du CdeM

Appliquons la fonction aux données du questionnaire exemplaire. Les résultats des opérations en boîtes de dialogue seront affichés dans un fichier de syntaxe, un fichier de résultat et un fichier graphique. —

```
*****
* Syntax file      : ex05_3.sps
* Data File charged : ques03.sav
* Output File     : res_05_3.lst
* Graphics File   :
* Comments       :
* Comments       :
* Comments       :
* Comments       : K54
* Author         : CDG-FZ München/klaus röder
*****
* Syntax File
GET FILE='C:\Roder_K\Roder_K54\que03_2.sav'.
COMPUTE gr_age2=0.
IF (gr_age < 3) gr_age2 = 1 .
IF (gr_age >=3 & gr_age <5) gr_age2 = 2.
IF (gr_age >=5 & gr_age <6) gr_age2 = 3 .
IF (gr_age >=6) gr_age2 = 4 .
EXECUTE .
CROSSTABS
  /TABLES=pealp_co BY gr_age2
```

```

/FORMAT= AVALUE NOINDEX BOX LABELS TABLES
/STATISTIC=CHISQ
/CELLS= COUNT .
SAVE OUTFILE='C:\Roder_K\Roder_K54\que05_1.sav'.
/COMPRESSED.
EXECUTE .

```

* Résultats (partiel):.

Tableaux croisés

Récapitulatif du traitement des observations

	Observations		Manquante		Total	
	N	Pour cent	N	Pour cent	N	Pour cent
Personnage allemand vivant plus connu * GR_AGE2	30	100.0%	0	.0%	30	100.0%

Tableau croisé Personnage allemand vivant plus connu * GR_AGE2

Effectif

	GR_AGE2	Total				
		1.00	2.00	3.00	4.00	
Personnage allemand vivant plus connu	Gerhard Schröder	1	3		2	6
	Boris Becker	1		3		4
	Claudia Schiffer	1			2	3
	Franz Beckenbauer	1	3	1		5
	Helmut Kohl		3		3	6
	Juliane Behrisch	1	5			6
Total		5	14	4	7	30

Tests du Khi-deux

	Valeur	ddl	Signification asymptotique (bilatérale)
Khi-deux de Pearson	29.289	15	.015
Rapport de vraisemblance	32.069	15	.006
Association linéaire par linéaire	.489	1	.484
Nombre d'observations valides	30		

a 24 cellules (100.0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de .40.

Khi-deux

Le khi-deux de Pearson est la forme la plus couramment utilisée; le khi-deux du rapport de vraisemblance est basé sur la théorie du maximum de vraisemblance. Le khi-deux de Mantel-Haenszel (Association linéaire par linéaire), mesure de liaison des variables ordinales, est également affiché par cette option. Dans l'exemple le coefficient Pearson indique que la hypothèse que les deux variables sont indépendantes peut être acceptée avec la probabilité de 1.5 % donc la dépendance est très vraisemblable. Bien pour l'exemple le khi-deux de Mantel-Haenszel est plus approprié car la variable PEALP_CO est une variable ordinale.

Exercice: Enquête Camerounaise Auprès des Ménages (ECAM):

Dans l'annexe Vous trouverez la description de l'Enquête Camerounaise Auprès des Ménages. Les données principales Vous trouverez sous le nom CHS01_02.SAV le fichier individuel et CHS02_01.SAV le fichier du ménage et le fichier combiné CHS03_01.SAV.

Utilisez SPSS pour manipuler et analyser les données:

1. Calculez les fréquence de la scolarité des enquêtes, excluez les enfants (< 18 ans). Utilisez le variables Plus haut diplôme obtenu et Année de fin de scolarité définitive et la valeur Jamais attendu
2. Créez une variable groupe d'âge (avec 4 valeurs: 0-15,16-30,31-45,45+), analysez la scolarité des enquêtes par groupe d'âge et par sexe (explorative). Utilisez la variable Plus haut diplôme obtenu.
3. Analysez l'indépendance entre sexe et scolarité des enquêtes. Utilisez les variables Plus haut diplôme obtenu et Année de fin de scolarité définitive et la valeur Jamais attendu.
4. Analysez l'indépendance entre dépenses pour la santé et la scolarité des chefs de ménages . Utilisez le variables S1Q18 et Année de fin de scolarité définitive
5. Analysez la fréquence des employés dans le secteur privé et publique avec les tableaux croisés. Utilisez les variables S2V6 et groupe d'âge.
6. Analysez les dépenses pour la santé dans le secteur privé et publique avec les tableaux croisés. Utilisez les variables S1Q18 et S2V6 et sexe. Utilisez les couches différentes pour la variable sexe (un tableaux ou deux tableaux).

6. Les fichiers de syntaxe

- 4 Vous devez apprendre l'essentiel d'utilisation des fichiers de syntaxe. Vous devez apprendre également l'interface des commandes (en tapant les commandes dans les fichiers de syntaxe) comme les menus de SPSS.

6.1. Considérations générales

Presque toutes les fonctionnalités du langage de commandes SPSS ont été préservées dans les boîtes de dialogue. Certaines options ne sont toutefois pas accessibles à partir des boîtes de dialogue. Les utilisateurs de SPSS avertis trouveront peut-être le langage de commandes plus efficace.

Dans SPSS pour WINDOWS vous exécutez le langage de commandes à partir d'une fenêtre de syntaxe. Vous pouvez ouvrir une fenêtre de syntaxe), taper les commandes, et les exécuter. Pour éviter de taper, vous pouvez coller dans une fenêtre de commande la syntaxe de la commande de la plupart des boîtes de dialogue à l'aide du bouton Coller. Si vous voulez, vous pouvez modifier la commande dans la fenêtre de syntaxe avant de l'exécuter.

Le bouton **[Exécuter]** d'une fenêtre de syntaxe exécute la/les commandes contenant la sélection. Pour exécuter un groupe de commandes, sélectionnez-les avec la souris ou au clavier avant de sélectionner **[Exécuter]**.

Lorsque des fonctions significatives ne sont accessibles qu'à partir du langage de commandes, la rubrique d'Aide de la boîte de dialogue affiche ces fonctions. Des schémas de syntaxe sont aussi accessibles en ligne. Pour obtenir des renseignements complets sur le langage de commandes SPSS, consultez le Manuel de référence SPSS.

Les avantages de l'utilisation es fichiers de syntaxe en bref:

- Répétition rapide des commandes et des sections des commandes
- Documentation des commandes et de l'analyse statistique
- Utilisation des fichiers de syntaxe des versions anciennes de SPSS.

6.2. Utiliser la fenêtre de syntaxe

Menu: [Fichier/Nouveau/Syntaxe SPSS]

Ouvrez une fenêtre de texte dont les options vous permettront de créer et de soumettre la syntaxe d'une commande SPSS. Le bouton **[Coller]** des boîtes de dialogue vous permettra de coller la syntaxe de la commande dans la fenêtre de syntaxe désignée

Menu: [Fichier/Ouvrir/Syntaxe SPSS]

Sélectionnez ou entrez le nom du fichier de texte devant être ouvert dans une fenêtre de syntaxe ou de résultats. La liste de fichiers vous permet de choisir le fichier à ouvrir. La liste Répertoires vous permet de choisir un répertoire. Vous pouvez sélectionner le type de fichier que vous souhaitez ouvrir dans la liste Type de fichier. Vous pouvez sélectionner un lecteur en choisissant l'une des options de la liste Lecteurs. Vous pouvez également taper le nom de fichier complet dans la zone Nom de fichier.

Menu: [Fichier/Enregistrer]

Enregistrez le fichier dans la fenêtre active. Si la fenêtre active contient un fichier de données, de graphiques ou de texte, celui-ci est enregistré sous son nom existant. Si le fichier n'a pas été nommé, cette commande est identique à *Enregistrer sous*. Les fichiers de données sont enregistrés sous le format SPSS. Si le même nom de fichier figure dans deux fenêtres différentes, utilisez *Enregistrer sous* ou *Fermer* et non *Enregistrer* pour éviter le risque de remplacement (écrasement) d'un fichier par un autre.

Si vous enregistrez un fichier de texte (à partir d'une fenêtre de résultats ou de syntaxe) existant déjà sur le disque, l'ancienne version sera renommée avec l'extension bak.

Menu: [Fichier/Ferme]

Fermez la fenêtre active, si cela est possible.

6.3. Edition d'un fichier de syntaxe

Vous pouvez modifier le texte d'une fenêtre de résultats ou d'une fenêtre de syntaxe avec la souris ou au clavier:

Les caractères que vous tapez apparaissent au point d'insertion, que vous pouvez déplacer avec la souris ou avec les touches de direction du clavier.

Sélectionnez du texte avec le bouton de gauche de la souris (pour des lignes entières) ou le bouton de droite de la souris (pour des sections de lignes, ou pour des rectangles). Il est conseillé d'utiliser le bouton gauche pour copier du texte dans d'autres applications.

L'accélérateur **[F8]** (Étendre la sélection) vous permet de balayer une sélection avec le curseur ou les touches de direction. Il existe d'autres accélérateurs pour sélectionner du texte.

Sélectionner du texte à partir du clavier

Fonction	Touche
Sélectionner le texte jusqu'au début du document	[MAJUSCULE]+[CTRL]+[ORIGINE]
Sélectionner le texte jusqu'à la fin du document	[MAJUSCULE]+[CTRL]+[FIN]
Sélectionner le caractère à gauche	[MAJUSCULE]+[GAUCHE]
Sélectionner le mot à gauche	[MAJUSCULE]+[CONTROLE]+[GAUCHE]
Sélectionner le caractère à droite	[MAJUSCULE]+[DROITE]
Sélectionner le mot à droite	[MAJUSCULE]+[CONTROLE]+[DROITE]
Sélectionner le texte jusqu'au début de la ligne	[MAJUSCULE]+[ORIGINE]
Sélectionner le texte jusqu'à la fin de la ligne	[MAJUSCULE]+[FIN]

Le menu [Edition] comporte les commandes [Couper, Copier, Coller, Rechercher] et [Remplacer]. Ces commandes ont aussi des accélérateurs. La commande [Annuler] des fenêtres de texte n'est pas accessible dans cette version. Utiliser les commandes d'édition à partir du clavier

Fonction	Touche
Rechercher	[F5]
Remplacer	[MAJUSCULE]+[F5]
Copier la sélection	[CTRL]+[C]
Couper la sélection	[CTRL]+[X]
Coller la sélection	[CTRL]+[V]
Supprimer la sélection	[SUPPR]
Définir un terme (dans la fenêtre de résultats)	[CTRL]+[G]
Arrondir les nombres dans la zone sélectionnée	[CTRL]+[0]

6.4. Exécuter les commandes

Le bouton [Exécuter] d'une fenêtre de syntaxe exécute la/les commandes contenant la sélection. Pour exécuter un groupe de commandes, sélectionnez-les avec la souris ou au clavier avant de sélectionner [Exécuter/Tout].

Lorsque des fonctions significatives ne sont accessibles qu'à partir du langage de commandes, la rubrique d'Aide de la boîte de dialogue affiche ces fonctions. Des schémas de syntaxe sont aussi accessibles en ligne. Pour obtenir des renseignements complets sur le langage de commandes SPSS, consultez le Manuel de référence SPSS.

6.5. La syntaxe des commandes

Les schémas décrivent la syntaxe du langage de commandes SPSS.

Appelez la fenêtre d'aide [syntaxe des commandes], cliquez sur la commande souhaitée.

Note: Les lignes d'un schéma de syntaxe ne reviennent pas automatiquement à la ligne sur votre écran. Si le schéma est trop large pour la fenêtre d'aide courante, une barre de défilement apparaît au bas de la fenêtre. Vous pouvez élargir ou réduire la fenêtre d'aide de manière à pouvoir lire les schémas de syntaxe sans les faire défiler.

Interpréter les schémas de syntaxe

Le caractère d'arrêt de la commande (point) n'apparaît pas dans le schéma de syntaxe.

Les éléments en majuscules correspondent à des mots-clés. Vous devez les entrer comme indiqué; les formes abrégées sont autorisées à condition d'éviter toute ambiguïté. Les éléments en minuscules correspondent à vos spécifications.

Les éléments [entre crochets] sont optionnels.

Des parenthèses, apostrophes et guillemets sont nécessaires là où ils figurent, sauf s'ils sont entre crochets. D'une manière générale, les doubles apostrophes ont la même valeur que les guillemets.

Les accolades {} indiquent un choix entre des éléments.

Pour listever, on lira: liste de noms de variables.

Exemple: l'enquête exemplaire du CdeM

Appliquons un schéma de syntaxe, ceci pour la commande *Descriptives*

```
DESCRIPTIVES [VARIABLES=] nomvar[(zname)] [nomvar...]
  [/MISSING={VARIABLE**} [INCLUDE]]
             {LISTWISE }
  [/SAVE]
  [/FORMAT={LABELS** } {NOINDEX**} {LINE**}]
           {NOLABELS } {INDEX } {SERIAL}
  [/STATISTICS=[DEFAULT**] [MEAN**] [MIN**] [SKEWNESS]]
              [STDDEV** ] [SEMEAN] [MAX**] [KURTOSIS]
              [VARIANCE ] [SUM ] [RANGE] [ALL]
  [/SORT=[ {MEAN } ] [ { (A) } ] ]
         {SMEAN } { (D) }
         {STDDEV }
         {VARIANCE}
         {KURTOSIS}

         {SKEWNESS}
         {RANGE }
         {MIN }
         {MAX }
         {SUM }
         {NAME }
```

**Option par défaut si la sous-commande n'est pas spécifiée.

Ceci le fichier de syntaxe pour l'application de la commande

```
* fichier syntaxe      : ex06_1.sps
* fichier données début: Que03_2.sav
* fichier données fin  :
* fichier résultats    :
* fichier graphique    :
* commentaire         :
* commentaire          : K54
* auteur               : CDG-FZ München/klaus röder
```

*****.

* Syntaxe:.

GET

FILE='C:\Roder_K\Roder_K54\Que03_2.sav'.

DESCRIPTIVES

VARIABLES=gr_age pealp_co soalp_co

/FORMAT=LABELS INDEX

/STATISTICS=MEAN STDDEV VARIANCE RANGE MIN MAX

/SORT=MEAN (A) .

Exercice: Enquête Camerounaise Auprès des Ménages (ECAM):

Dans l'annexe Vous trouverez la description de l'Enquête Camerounaise Auprès des Ménages. Les données principales Vous trouverez sous le nom CHS01_02.SAV le fichier individuel et CHS02_01.SAV le fichier du ménage et le fichier combiné CHS03_01.SAV.

Utilisez SPSS pour manipuler et analyser les données:

Utilisez la touche [Coller] pour le début. Continuez avec l'éditeur de texte après. Utilisez la commande *Descriptives* pour afficher la variance, l'asymétrie et l'aplatissement des variables suivantes:

1. Age
2. Coût de la consultation (Dépenses pour la santé)

Utilisez la commande *Crosstabs* pour afficher le tableau croisé avec les variables

3. Coût de la consultation (Dépenses pour la santé). Utilisez les groupes de dépenses. Si ces groupes n'existent pas, créez les avec l'utilisation des fichiers de syntaxe.
et
Activité Principale

7. La présentation des données

- 4 L'analyse statistique produit les résultats. Ces résultats doivent être présentés au consommateur des statistiques d'une forme compréhensive, informative et attrayante. C'est à dire le statisticien doit présenter ses résultats appropriés pour l'utilisateur. Certaines méthodes de présentation demandent des connaissances statistiques préalables. Une présentation dans un magazine quotidien adresse un autre public que dans une publication spécialisée. Moins d'explications en texte sera toujours préférable. Les graphiques auront toujours la préférence, si les données peuvent être présentées de cette manière, même en supplément aux présentations plus traditionnelles. Ce chapitre s'occupe des méthodes de présentation des données en SPSS, listage, tabulation et présentation graphique. Vous devez connaître les possibilités et les raisons pour lesquelles les méthodes sont choisies et comment elles sont effectuées.

7.1. Les tableaux récapitulatifs

7.1.1. Tableaux de bord en lignes

produit des tableaux de bord dans lesquels différentes statistiques élémentaires sont disposées en lignes. Les listes d'observations sont également disponibles, avec ou sans statistiques élémentaires.

Menu: [Analyse/Rapports/Tableaux de bord en lignes]

Langage de commandes: *REPORT*

[Variables en colonnes]. Donne la liste des variables de tableau pour lesquelles vous voulez obtenir des listes d'observations ou des statistiques élémentaires et contrôle le format d'affichage des Variables en colonnes.

[Variables de ventilation]. Donne la liste des variables de ventilation optionnelles qui divisent le tableau de bord en groupes et contrôlent les statistiques élémentaires et les formats d'affichage des colonnes de ventilation. Pour les variables de ventilation multiples, il y aura un groupe séparé pour chaque modalité de chaque variable de ventilation à l'intérieur des modalités de la variable de ventilation précédente dans la liste. Les variables de ventilation doivent être des variables qualitatives discrètes qui divisent les observations en un nombre limité de modalités significatives. Les valeurs individuelles de chaque variable de ventilation apparaissent, triées, dans une colonne séparée à gauche des Variables en colonnes.

[Tableau de bord]. Contrôle les caractéristiques globales du tableau de bord, y compris les statistiques élémentaires globales, l'affichage des valeurs manquantes, la numérotation des pages et les titres.

[Afficher les observations]. Affiche les valeurs effectives (ou les étiquettes de valeur) des variables de

Variables en colonnes pour chaque observation. Cela produit une liste, qui peut être nettement plus longue qu'un tableau de bord. [Aperçu]. N'affiche que la première page du tableau de bord. Cette option est utile pour avoir un aperçu du format de votre tableau sans traiter le tableau entier.

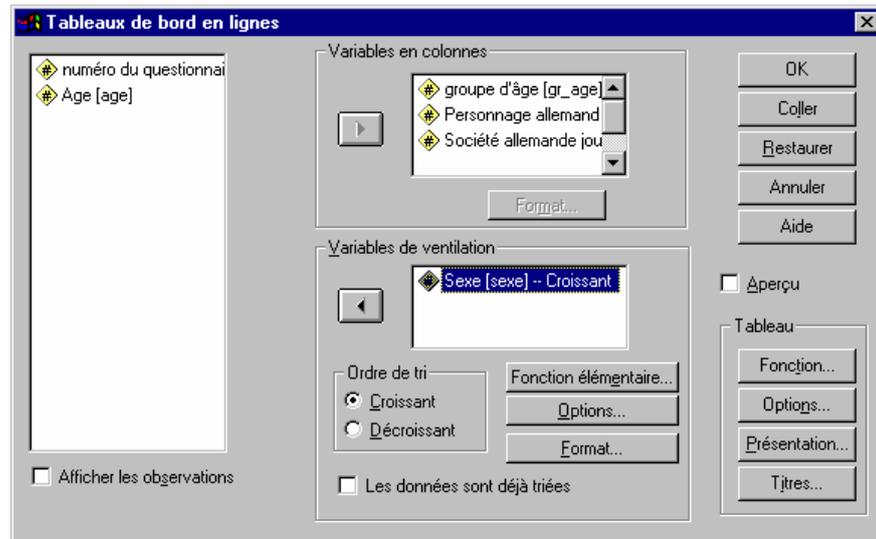


Figure 22: Tableaux de bord en lignes

[Les données sont déjà triées]. Pour les rapports avec variables de ventilation, le fichier de données doit être trié par valeurs des variables de ventilation avant de générer le tableau de bord. Si votre fichier de données est déjà trié par valeurs des variables de ventilation, vous pouvez gagner du temps de traitement en sélectionnant cette option. Cette option est particulièrement utile après avoir vu un aperçu du tableau.

Fonctions élémentaires du Tableau

Les deux boîtes de dialogue Fonctions élémentaires contrôlent l'affichage des statistiques élémentaires pour les groupes de ventilation et pour l'ensemble du tableau de bord. Fonctions élémentaires contrôle les statistiques de sous-groupe pour chaque modalité définie par la ou les variables de ventilation. Fonctions élémentaires Finales contrôle les statistiques globales affichées à la fin du tableau de bord.

Les statistiques élémentaires disponibles sont la somme des valeurs, la moyenne des valeurs, la valeur minimale, la valeur maximale, le nombre d'observations (effectif), le pourcentage d'observations situées au-dessus ou en dessous d'une valeur spécifiée, le pourcentage d'observations compris à l'intérieur d'un intervalle donné de valeurs, l'écart-type, l'aplatissement, la variance et l'asymétrie.

Options de Ventilation de Tableau de Bord

Options de Ventilation contrôle l'espacement et la pagination des informations de modalité de ventilation.

Gestion des pages : Contrôle l'espacement et la pagination des modalités de la variable de ventilation sélectionnée. Vous pouvez spécifier un nombre de lignes vides entre les modalités de ventilation ou commencer chaque modalité de ventilation sur une nouvelle page.

Lignes à sauter avant fonctions élémentaires. Contrôle le nombre de lignes vides entre les étiquettes ou les données des modalités de ventilation et les statistiques élémentaires. Cette option est particulièrement utile pour les tableaux de bords combinés incluant des listes d'observations individuelles et des statistiques élémentaires pour les modalités de ventilation ; dans ces tableaux, vous pouvez insérer des espaces entre les listes d'observations et les statistiques élémentaires.

Les boîtes de dialogue [Format] contrôlent les titres et largeurs des colonnes, l'alignement du texte et l'affichage des valeurs de données ou des étiquettes de valeur. Format colonne de données contrôle le format des Variables en colonnes du côté droit de la page du tableau de bord. Format colonne de ventilation contrôle le format des Colonnes de ventilation du côté gauche.

Titre de la colonne. Pour la variable sélectionnée, contrôle le titre de la colonne. Les titres longs sont automatiquement ajustés dans la colonne. Utilisez la touche Entrée pour insérer manuellement des sauts de lignes aux endroits où vous voulez ajuster les titres.

Position des valeurs dans la colonne. Pour la variable sélectionnée, contrôle l'alignement des valeurs de données ou des étiquettes de données dans la colonne. L'alignement des valeurs ou des étiquettes n'affecte pas l'alignement des titres de colonnes. Vous pouvez soit indenter le contenu de la colonne d'un nombre de caractères donné, soit centrer le contenu de la colonne.

Contenu de la colonne. Pour la variable sélectionnée, contrôle l'affichage soit des valeurs de données, soit des étiquettes de valeur définies. Les valeurs de données sont affichées pour toutes les valeurs qui ne possèdent pas d'étiquette de valeur définie. (Non disponible pour les Variables en colonnes dans les Tableaux de bord en colonnes)

[Options] du Tableau de bord contrôle le traitement et l'affichage des valeurs manquantes et la numérotation des pages du tableau de bord.

Exclure les observations avec des valeurs manquantes : Elimine (du tableau de bord) toute observation avec des valeurs manquantes pour l'une des variables du tableau de bord.

Représentation des valeurs manquantes : Vous permet de spécifier le symbole représentant les valeurs manquantes dans le fichier de données. Ce symbole ne peut comporter qu'un seul caractère et sert à représenter les valeurs manquantes système et manquantes utilisateur .

Paginer à partir de : Vous permet de spécifier un numéro pour la première page du tableau de bord.

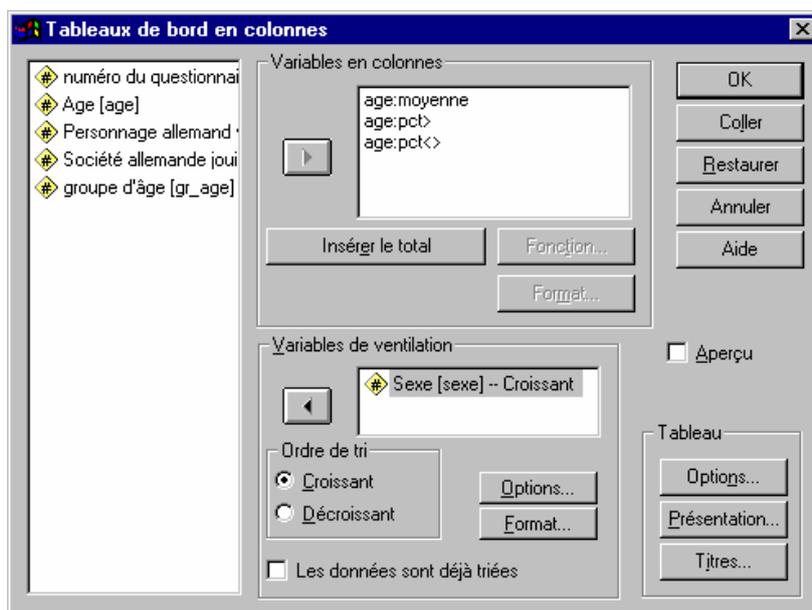
7.1.2. Tableaux de bord en colonnes

crée des tableaux dans lesquels les différentes fonctions élémentaires apparaissent les unes sous les autres (en colonnes).

Menu: [Analyse/Rapports/Tableaux de bord en colonnes]

Langage de commandes: *REPORT*

Dans un tableau, les colonnes de données présentent des statistiques récapitulant vos variables, tandis que les colonnes de ventilation (à gauche) affichent les valeurs individuelles des variables de ventilation, ce qui permet d'identifier les statistiques des groupes dans les colonnes de



données. Vous pouvez voir un exemple de tableau de bord en colonnes, afin de le comparer avec un tableau de bord en lignes.

Tableaux de bord en colonnes produit des tableaux de bord dans lesquels différentes statistiques élémentaires apparaissent en colonnes séparées.

Figure 23: Tableau de bord en colonnes

Exemple : Une société possédant une chaîne de magasins conserve des dossiers sur les employés comprenant le salaire, l'ancienneté et le service où chaque employé travaille. Vous pourriez générer

un tableau de bord fournissant des statistiques élémentaires sur les salaires (par exemple moyenne, minimum, maximum) pour chaque division.

Variables en colonnes. Fournit la liste des variables du tableau de bord pour lesquelles vous voulez des statistiques élémentaires et contrôle le format d'affichage et les statistiques élémentaires affichées pour chaque variable.

Variables de ventilation. Fournit la liste des variables de ventilation optionnelles qui divisent le tableau de bord en groupes et contrôle les formats d'affichage des colonnes de ventilation. Pour les variables de ventilation multiples, il y aura un groupe séparé pour chaque modalité de chaque variable de ventilation à l'intérieur des modalités de la variable de ventilation précédente dans la liste. Les variables de ventilation doivent être des variables qualitatives discrètes qui divisent les observations en un nombre limité de modalités significatives.

Tableau de bord. Contrôle les caractéristiques globales du tableau de bord, y compris l'affichage des valeurs manquantes, la numérotation des pages et les titres.

Aperçu. N'affiche que la première page du tableau de bord. Cette option est utile pour avoir un aperçu du format de votre tableau sans traiter le tableau entier.

Les données sont déjà triées. Pour les rapports avec variables de ventilation, le fichier de données doit être trié par valeurs des variables de ventilation avant de générer le tableau de bord. Si votre fichier de données est déjà trié par valeurs des variables de ventilation, vous pouvez gagner du temps de traitement en sélectionnant cette option. Cette option est particulièrement utile après avoir vu un aperçu du tableau.

Faites passer une variable au moins dans la zone de liste Variables en colonnes. Elle apparaîtra tout d'abord accompagnée de la statistique Somme. (Pour obtenir une statistique différente, sélectionnez le couple variable:statistique dans la zone Variables en colonnes et cliquez sur Fonction.) Une colonne est générée pour chaque couple variable:statistique de la zone Variables en colonnes; cette colonne comprend les statistiques principales indiquées pour la variable, calculées à partir des observations de chaque groupe. La zone de groupe Variables en colonnes comprend également un jeu de boutons permettant de gérer le format et le contenu des colonnes de données: Pour afficher une colonne contenant la somme, la moyenne, le rapport ou autre calculé(e) sur les colonnes désignées, sélectionnez Insérer le total. Vous verrez apparaître une variable appelée *Total* dans la zone de liste Variables en colonnes, sous la variable sélectionnée courante de cette zone. Total étant sélectionné dans la zone Variables en colonnes, cliquez sur **[Fonction]** pour désigner le mode de calcul du total de la colonne. Pour modifier la statistique dans un couple variable:statistique figurant dans la zone de liste Variables en colonnes, choisissez le couple puis sélectionnez **[Fonction]**. Pour spécifier le format de sortie d'une variable ou d'un total, sélectionnez cette variable ou ce total dans la zone de liste Variables en colonnes et sélectionnez Format. Les Variables de

ventilation définissent les sous-groupes d'observations sur lesquels porteront les statistiques principales. Les valeurs individuelles de chaque variable de ventilation apparaissent, triées, dans une colonne distincte située à gauche des colonnes de données. Sélectionnez une ou plusieurs variables qualitatives comme variables de ventilation, et faites les passer dans la zone Variables de ventilation, dans l'ordre voulu. Vous pouvez trier les groupes définis par les variables de ventilation dans l'ordre croissant ou décroissant. Sélectionnez la variable dans la zone Variables de ventilation, puis sélectionnez *Croissant* ou *Décroissant* dans la zone Ordre de tri. L'ordre de tri choisi pour chaque variable est indiqué à côté du nom de celle-ci, dans la zone Variables de ventilation, par la lettre A (croissant) ou D (décroissant). Les variables chaîne sont triées dans l'ordre alphabétique normal ou inverse. Si les données sont déjà triées, sélectionnez *Les données sont déjà triées* afin que SPSS ne les trie pas à nouveau. Cela s'avère souvent pratique lorsque l'on affine le format d'un tableau. Le premier tableau généré (avec un ensemble donné de variables de ventilation) trie le fichier, et si vous sélectionnez cette option les tableaux suivants pourront être produits plus rapidement. Les boutons de commandes situés dans la zone de groupe Variables de ventilation affectent l'espacement et l'aspect des valeurs affichées dans la/les colonnes de ventilation. Sélectionnez une variable dans la zone de liste Variables de ventilation puis cliquez sur:

[Options] pour gérer les sauts de page et les lignes vierges avant les sous-totaux.

[Format] pour gérer l'affichage des valeurs ou des étiquettes de valeur des colonnes de ventilation.

Sélectionnez **[Aperçu]** pour visualiser une page type du tableau produit selon les spécifications courantes de cette boîte de dialogue. Cette option ne génère pas le tableau complet.

La zone de groupe Tableau vous permet de demander des statistiques générales et de gérer les sauts de page, la présentation, et les titres:

Sélectionnez **[Options]** pour gérer le traitement des valeurs manquantes, le tri des données et la pagination.

Sélectionnez **[Présentation]** pour personnaliser la présentation du tableau.

Sélectionnez **[Titres]** pour spécifier un texte de titre et de pied de page pour le haut et le bas de chaque page du tableau.

Exemple: l'enquête exemplaire du CdeM

Appliquons la fonction aux données du questionnaire exemplaire. Les résultats des opérations en boites de dialogue seront affichés dans un fichier de syntaxe, un fichier de résultat et un fichier graphique.

```
*****
* Syntax file      : ex07_1.sps
* Data File charged : que03_2.sav
* Output File     : res07_1.txt
* Graphics File   :
* Comments       :
* Comments       :
* Comments       :
* Comments       : K54
* Author         : CDG-FZ München/klaus röder
*****
* Syntax File
GET
FILE='C:\Roder_K\Roder_K54\Que03_2.sav'.
SORT CASES BY sexe (A) .
Report
/FORMAT= CHWRAP(ON) PREVIEW(OFF) CHALIGN(BOTTOM) UNDERSCORE(ON)
ONEBREAKCOLUMN(OFF) CHDSPACE(1) SUMSPACE(0) AUTOMATIC NOLIST
BRKSPACE(0)
PAGE(1) MISSING:' ' LENGTH(1, 59) ALIGN(LEFT) TSPACE(1) FTSPACE(1)
MARGINS(1,83)
/TITLE=
RIGHT 'Page )PAGE'
/VARIABLES
gr_age (VALUES) (RIGHT) (OFFSET(0)) (20)
pealp_co (VALUES) (RIGHT) (OFFSET(0)) (17)
soalp_co (VALUES) (RIGHT) (OFFSET(0)) (26)
/BREAK sexe (LABELS) (LEFT) (OFFSET(0))(13)(SKIP(1))
/SUMMARY SUM( gr_age) SKIP(0) SUM( pealp_co ) SUM( soalp_co ) 'Somme'
/SUMMARY MEAN( gr_age) MEAN( pealp_co ) MEAN( soalp_co ) 'Moyenne' .
SORT CASES BY sexe (A) .
Report
/FORMAT= CHWRAP(ON) BRKSPACE(-1) SUMSPACE(0) AUTOMATIC
PREVIEW(OFF) CHALIGN(BOTTOM) CHDSPACE(1)
UNDERSCORE(ON) ONEBREAKCOL(OFF)
PAGE(1) MISSING:' ' LENGTH(1, 59)ALIGN(LEFT) TSPACE(1) FTSPACE(1)
MARGINS(1,48)
/TITLE=
RIGHT 'Page )PAGE'
/VARIABLES
age 'Age' 'Moyenne' (RIGHT) (OFFSET(0)) (13)
$dummy01 (DUMMY) 'Age' '> 30' (RIGHT) (OFFSET(0)) (8)
$dummy02 (DUMMY) 'Age' 'Entre 30 et 40' (RIGHT) (OFFSET(0)) (8)
/BREAK sexe (LABELS) (LEFT) (OFFSET(0)) (SKIP(1))(13)
/SUMMARY MEAN(age) SKIP(0) ADD( PGT(30)(age)) ($dummy01(PCT)(1)) ADD(
PIN(30,40)(age)) ($dummy02(PCT)(1)).
```

SORT CASES BY sexe (A) .

Tableau de bord en lignes

Sexe	Personnage allemand vivant groupe d'âge	Personnage allemand vivant plus connu	Société allemande jouissant du plus grand prestige
masculin			
Somme	80.00	74.00	59.00
Moyenne	4.44	4.11	3.28
féminin			
Somme	47.00	35.00	45.00
Moyenne	3.92	2.92	3.75

Tableau de bord en colonnes

Sexe	Age		
	Age Moyenne	Age > 30	Entre 30 et 40
masculin	39.44	66.7%	16.7%
féminin	35.92	50.0%	33.3%

7.2. Le module Tableaux

Le module supplémentaire **Tableaux** est une combinaison des fonctions de listage, des tableaux croisés et des statistiques descriptives. La caractéristique principale quand même est la présentation prêt pour la publication des tableaux des données analysées ou simplement affichées par SPSS.

7.2.1. Tableaux de base

crée des tableaux de qualité publication avec tris croisés et statistiques par sous-groupes. Cette commande permet de créer des présentations emboîtées. Les mêmes statistiques sont calculées pour toutes les variables récapitulées dans le tableau.

Menu :

[Analyse/Tabuler/Tableaux
statistiques de base]

Langage de commandes:
TABLES

Faites passer une ou plusieurs variables dans la zone de liste *Principales caractéristiques*, ou dans l'une quelconque des trois zones de liste *Sous-groupes*, ou dans les deux. Au moins une variable doit être spécifiée dans l'une quelconque de ces zones. Si vous spécifiez une variable pour

Principales caractéristiques, vous devez sélectionner Statistiques pour indiquer quelle(s) statistique(s) vous souhaitez voir.

Pour obtenir des tableaux de fréquences ou des tableaux croisés simples comportant des effectifs ou des pourcentages, laissez vide la zone de liste *Principales caractéristiques* et faites passer un ou plusieurs critères de regroupement dans les zones de liste *Sous-groupes*.

Pour obtenir des statistiques sur des sous-groupes, faites passer la/les variables pour lesquelles vous souhaitez obtenir des statistiques dans la zone de liste *Principales caractéristiques*, et le/les critères de regroupement dans les zone de liste *Sous-groupes*. Sélectionnez

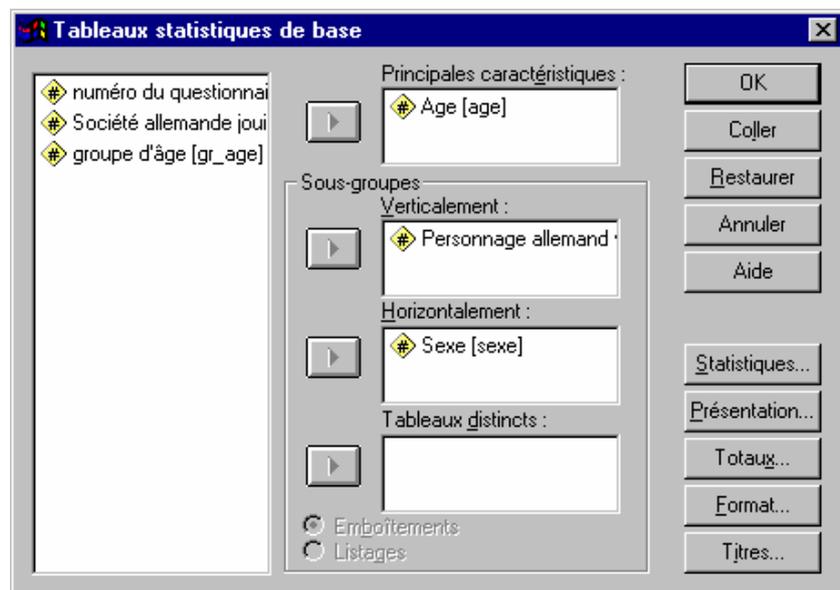


Figure 24: Tableaux de base

[Statistiques] pour spécifier les statistiques voulues. Veillez à sélectionner des statistiques appropriées pour les caractéristiques.

Pour obtenir des statistiques générales, faites passer la/les variables pour lesquelles vous voulez obtenir des statistiques dans la zone de liste *Principales caractéristiques* et sélectionnez

[Statistiques] pour spécifier les statistiques voulues. Veillez à sélectionner des statistiques appropriées pour les caractéristiques.

Les trois zones de liste *Sous-groupes* vous permettent de définir les cellules dans lesquelles seront affichés des sous-groupes d'observations. Les dispositions suivantes peuvent être combinées:

Verticalement sur la page (en lignes séparées).

Horizontalement sur la page (en colonnes séparées).

La combinaison de *Verticalement* et *Horizontalement* génère un affichage en tableau croisé. Sur des Tableaux distincts.

Lorsque vous spécifiez plusieurs critères de regroupement dans une dimension, vous pouvez choisir entre les options *Emboîtements* et *Listages*. (La commande *Tableaux généraux* permet de gérer plus précisément les relations entre les critères de regroupement multiples d'une même dimension.)

Sélectionnez **[Statistiques]** pour spécifier les statistiques qui apparaîtront dans le tableau. Pour obtenir des statistiques sur des caractéristiques, sélectionnez tout d'abord les caractéristiques dans la zone de liste *Principales caractéristiques*. Par défaut, s'il existe des caractéristiques, leurs moyennes sont affichées; dans le cas contraire, l'effectif est affiché.

Sélectionnez **[Présentation]** pour spécifier la dimension Caractéristiques, la dimension Statistiques, la position relative des caractéristiques et des sous-groupes dans la dimension Caractéristiques, et pour supprimer les étiquettes des critères de regroupement. Si vous avez entré une variable dans l'une des zones de liste *Sous-groupes*, vous pouvez sélectionner **[Totaux]** pour demander des statistiques globales (après les statistiques de sous-groupes) pour le/les critères de regroupement.

Sélectionnez **[Format]** pour spécifier les marges, les tailles de colonnes, les bordures, l'aspect des cellules vides et des statistiques non définies, et les caractères au niveau desquels les lignes

peuvent être coupées dans les étiquettes.

Sélectionnez **[Titres]** pour spécifier les lignes de texte pour l'en-tête, le pied de page et l'angle supérieur gauche du tableau.

7.2.2. Tableaux généraux

offre la gestion la plus précise du contenu et des formats des tableaux au prix de spécifications plus complexes. Différentes statistiques peuvent être affichées pour les différentes variables.

La commande *Tableaux généraux* permet d'emboîter certaines variables et d'en lister d'autres.

Menu: [Analyse/Tabuler/Tableaux généraux]

Langage de commandes: *TABLES*

Spécifiez *Les étiquettes des statistiques s'affichent* pour déterminer où seront affichées les étiquettes des statistiques des cellules. (Ces choix sont parfois non disponibles, lorsqu'un seul emplacement est possible.)

Faites passer une ou plusieurs variables dans une ou plusieurs des zones de liste Lignes, Colonnes ou Strates. Vous pouvez soit utiliser les variables normales de la liste source située en haut à gauche, soit utiliser des vecteurs de réponses multiples.

Pour que les boutons OK et Coller soient disponibles:

1. Le côté du tableau ne doit pas être vierge. Vous devez introduire une variable dans la zone de liste Ligne, à moins que vous n'ayez spécifié que les étiquettes des statistiques doivent être affichées sur le côté.
2. Dans les tableaux à plusieurs pages, le haut également ne doit pas être vierge. Si vous introduisez une variable dans la zone de liste Strates, vous devez également en introduire une dans la zone de liste Colonnes, à moins que vous n'ayez spécifié que les étiquettes des statistiques doivent être affichées en haut. (Après Ligne et Colonne, la Couche est la troisième dimension des tableaux).

Lorsqu'une variable apparaît en surbrillance dans la zone Lignes, Colonnes ou Strates, vous pouvez utiliser les options de Variable sélectionnée, à droite, pour modifier sa fonction dans le tableau:

Crit. de regroupement identifie un critère de regroupement dont chaque catégorie forme une ligne ou une colonne dans le tableau, ou un sous-tableau séparé. Vous pouvez afficher des effectifs ou divers pourcentages pour les catégories d'un critère de regroupement.

Caractéristique identifie une caractéristique pour laquelle des statistiques principales apparaissent dans le tableau. La première caractéristique que vous spécifiez définit la dimension Caractéristiques.

La statistique principale par défaut est la moyenne, et n'est pas étiquetée; vous pouvez l'étiqueter en sélectionnant la caractéristique et en utilisant la boîte de dialogue secondaire **[Modifier les statistiques]**.

Lorsqu'une dimension Caractéristiques est spécifiée, les mots *Dimension Caractéristiques* apparaissent au-dessus de la liste de variables de cette dimension, et vous ne pouvez pas sélectionner Caractéristique pour une variable se trouvant dans une autre dimension.

Sélectionnez *Omettre* l'étiquette pour supprimer l'étiquette de la variable sélectionnée.

Dans chacune de ces listes, les variables multiples sont listées dans leur dimension dans le tableau, sauf si vous les emboîtez. Le bouton **[Emboîter]** vous permet d'emboîter une variable sélectionnée dans les catégories de la variable qui la précède dans la liste correspondant à sa dimension. Le bouton **[Déboîter]** vous permet de déboîter une variable emboîtée.

[Modifier les statistiques] vous permet de spécifier quelles statistiques seront affichées pour la variable sélectionnée. Vous ne pouvez pas affecter des statistiques à une variable dans laquelle une autre variable est emboîtée. Une fois que vous avez affecté des statistiques à une variable, elles deviennent les statistiques par défaut des autres variables de même type dans la même dimension. Cela facilite l'affectation des mêmes statistiques à plusieurs variables dans la dimension statistiques.

Insérer le total vous permet d'ajouter des statistiques en total après les statistiques de groupe de la variable sélectionnée. Vous pouvez modifier l'étiquette qui sera utilisée pour les statistiques en total dans la zone *Etiquette total*: Les mêmes statistiques sont normalement affichées pour un total comme pour la variable totalisée, mais en sélectionnant le total et en cliquant sur **[Modifier les statistiques]**, vous pouvez spécifier des statistiques différentes pour la ligne ou la colonne contenant la/les statistiques en total.

La zone de liste Statistiques montre les statistiques disponibles. La zone de liste Statistiques dans les cellules indique les statistiques qui seront affichées si vous sélectionnez **[Poursuivre]**.

Dans **Tableaux de base**, les statistiques spécifiées ici s'appliquent à l'ensemble du tableau. Vous devez sélectionner une ou plusieurs statistiques des caractéristiques si vous avez fait passer des variables dans la zone de liste *Principales caractéristiques*, et l'une des statistiques des critères de regroupement dans le cas contraire.

Dans **Tableaux généraux**, les statistiques spécifiées ici s'appliquent à la ou aux variables qui étaient sélectionnées lorsque vous avez accédé à cette boîte de dialogue, et la liste des statistiques disponibles dépend du type de cette variable: *critère de regroupement*, *caractéristique*, ou *vecteur de réponses multiples*.

[Ajouter] transfère la statistique sélectionnée dans la zone de liste Statistiques à la zone de liste Statistiques dans les cellules en utilisant le format affiché sous les listes. Sélectionnez tout d'abord une statistique dans la zone de liste Statistiques, modifiez son format si vous le souhaitez, puis sélectionnez **[Ajouter]**.

[Remplacer] permet de remplacer le format de la statistique sélectionnée dans Statistiques dans les cellules par celui affiché sous les listes. Sélectionnez tout d'abord une statistique dans la zone de liste Statistiques dans les cellules, modifiez son format si vous le souhaitez, puis sélectionnez **[Remplacer]**.

[Supprimer] permet de supprimer la statistique sélectionnée dans Statistiques de cellules.

[Format]. Modifiez ces choix après avoir sélectionné une statistique, mais avant de sélectionner **[Ajouter]** ou **[Remplacer]**.

[Format] vous permet de sélectionner le format d'affichage de la statistique à partir d'un menu de formats génériques. Les zones *Longueur* et *Décimales* permettent d'entrer des spécifications de format supplémentaires. Ainsi, le format numérique standard ddd.dd n'affiche pas nécessairement 2 positions décimales, mais respecte la valeur de Décimales. **Nota:** Vous ne pouvez pas spécifier une longueur d'une valeur supérieure à celle de

largeur de la colonne établie dans la boîte de dialogue *Format*. Pour modifier la largeur de colonne, cliquez sur **[Format]**, dans la boîte de dialogue principale.

Etiquette vous permet de spécifier l'étiquette qui décrira la statistique dans l'en-tête de la ligne, de la colonne ou de la couche.

Les choix de *Tri par effectif* vous permettent optionnellement de réorganiser les cellules du tableau en fonction de l'effectif présent dans chaque cellule.

Aucun, l'option par défaut, laisse les cellules des tableaux dans l'ordre des valeurs des critères de regroupement.

Dans *Tableaux généraux*, le tri *par effectif* transforme automatiquement la dimension courante en dimension Statistiques.

Au bas de cette boîte de dialogue, SPSS affiche la largeur courante de la colonne *Etiquettes* ligne (talon) et de chacune des Colonnes du tableau même, en fonction des spécifications entrées dans la boîte de dialogue secondaire *Format*.

Dans la boîte de dialogue principale:

Sélectionnez **[Format]** pour spécifier les marges, les dimensions de page, les bordures, et le traitement des cellules et données manquantes.

Sélectionnez **[Titres]** pour spécifier les lignes de texte pour l'en-tête, le bas de page et l'angle supérieur gauche du tableau.

7.2.3. Les tableaux de réponses multiples

produisent des tableaux croisés et des tableaux de fréquence de base dans lesquels au moins une variable est un vecteur multiréponses. Vous n'êtes pas obligé de définir un vecteur multiréponses pour utiliser cette procédure, mais vous obtiendrez probablement de meilleurs

résultats avec des Tableaux statistiques de base si vous n'avez pas besoin d'utiliser de vecteur multiréponses.

Menu: [Analyse/Tabuler/Tableaux de réponses multiples]

Langage de commandes: *TABLES*

Pour créer un tableau, vous devez placer au moins une variable ou un vecteur multiréponses dans une au moins des trois dimensions du tableau :

Vecteurs multiréponses : Liste les vecteurs multiréponses actuellement définis. Cliquez sur Vecteurs multiréponses... pour définir un nouveau vecteur.

Un vecteur de réponses multiples est, d'une manière générale, une façon de coder les réponses à une série de questions, soit sous forme de dichotomies multiples, soit sous forme de catégories multiples. Cette expression est parfois synonyme de "vecteurs de catégories multiples".

7.2.4. Tableaux de fréquences

crée des tableaux de distribution de fréquences de variables qualitatives. Des tableaux de fréquences disposés côte à côte sont disponibles pour les variables multiples qui présentent les mêmes catégories ainsi que pour des sous-groupes d'observations.

Menu: [Analyse/Tabuler/Tableaux de fréquences]

Langage de commandes: *TABLES*

Faites passer une variable au moins dans la zone de liste Fréquences pour. Les variables multiples doivent normalement avoir les mêmes codages de catégorie et étiquettes de valeur. (Si des étiquettes de valeur différentes existent pour une même valeur, SPSS utilisera l'étiquette pour la variable spécifiée en premier dans cette liste.)

Pour obtenir des tableaux de fréquences pour des Sous-groupes, faites passer une ou plusieurs variables dans la zone de liste appropriée:

Dans chaque tableau affiche les fréquences des sous-groupes côte à côte dans un seul tableau.

Tableaux distincts affiche un tableau de fréquences séparé pour chaque sous-groupe.

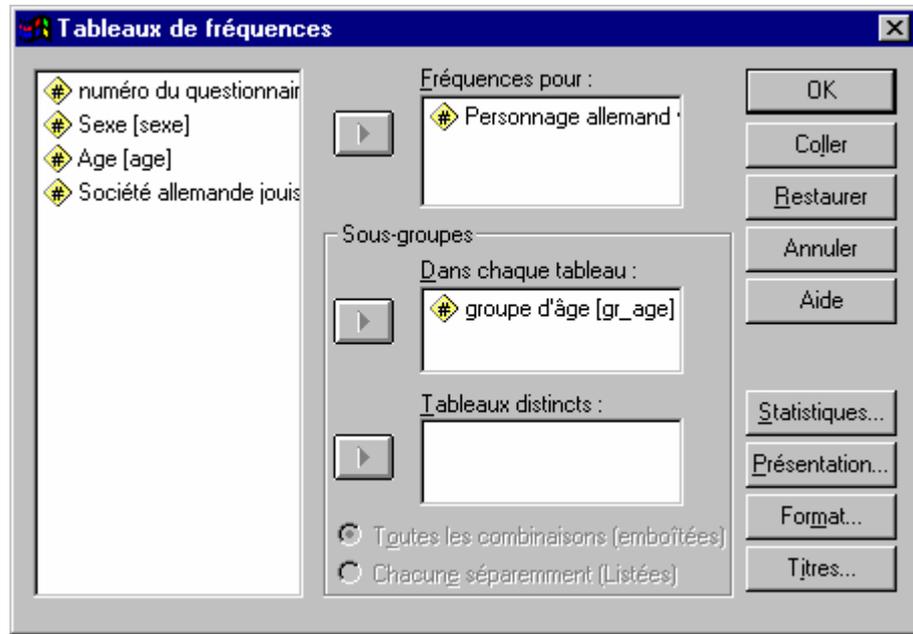


Figure 25: Tableaux de fréquences

Si vous faites passer deux variables ou davantage dans l'une quelconque des zones de liste Sous-groupes, choisissez entre *Emboîtements* et *Listages*.

Sélectionnez **[Statistiques]** pour choisir les statistiques du tableau de fréquences.

Sélectionnez **[Présentation]** pour spécifier la présentation du tableau de fréquences.

Sélectionnez **[Format]** pour spécifier les marges, les dimensions de page, les bordures, et le traitement des cellules et données manquantes.

Sélectionnez **[Titres]** pour spécifier les lignes de texte pour l'en-tête, le pied de page et l'angle supérieur gauche des tableaux.

Appliquons les transformations aux données du questionnaire exemplaire. Les résultats seront affichés dans un fichier de syntaxe.

Exemple: l'enquête exemplaire du CdeM

```
*****
* fichier syntaxe: ex07_2.sps
* fichier données début: QUE03_2.sav
```

```
* fichier données fin :
* fichier résultats : RES07_2.spo
* fichier graphique : -----
* commentaire :
* commentaire : K54
* auteur : CDG-FZ München/klaus röder
*****
GET
FILE='C:\Roder_K\Roder_K54\que03_2.sav'.
* Basic Tables.
TABLES
/FORMAT BLANK MISSING('.')
/OBSERVATION age
/TABLES pealp_co > age
BY sexe > (STATISTICS).
* General Tables.
TABLES
/FORMAT BLANK MISSING('.')
/GBASE=CASES
/TABLE=gr_age BY soalp_co > (STATISTICS) BY sexe .
* Table of Frequencies.
TABLES
/FORMAT BLANK MISSING('.') /TABLES
(LABELS) BY
gr_age > pealp_co
/STATISTICS
CPCT ((PCT7.1) '%' :gr_age ) .
```

Tableaux de base

Personnage allemand vivant plus connu		Age	Sexe	
			masculin	féminin
Gerhard Schröder	Age	23.00	50.00	
Boris Becker	Age	44.33	13.00	
Claudia Schiffer	Age	47.00	.	
Franz Beckenbauer	Age	44.00	28.25	
Helmut Kohl	Age	41.17	.	
Juliane Behrisch	Age	30.50	27.50	

Tableaux generaux

Sexe féminin

groupe d'âge	Société allemande jouissant du plus grand prestige Mercedes- Benz				
		Carl-Duisberg Gesellschaft	MAN	BMW	Siemens
<= 10					
10< <=20		2	1		
20< <=30		1		1	1

30< <=40	2	2
40< <=50		
50< <=60		
60< <=70		
70<	2	

Sexe masculin

		Société allemande jouissant du plus grand prestige				
		Mercedes-Benz	Carl-Duisberg Gesellschaft	MAN	BM W	Siemens
groupe d'âge	<= 10			1		
	10< <=20			1		
	20< <=30	1				3
	30< <=40			2	1	
	40< <=50	2			1	1
	50< <=60	2				
	60< <=70	1				2
70<						

Tableaux de fréquences

	groupe d'âge <= 10	10< <=20	20< <=30	30< <=40	40< <=50	50< <=60	60< <=70	70<
	Personnage allemand vivant plus connu %							
Gerhard Schröder		25.0%	14.3%	28.6%				100.0%
Boris Becker		25.0%			75.0%			
Claudia Schiffer		25.0%					66.7%	
Franz Beckenbauer		25.0%	14.3%	28.6%	25.0%			
Helmut Kohl			42.9%			100.0%	33.3%	
Juliane Behrisch	100.0%		28.6%	42.9%				

Page laissée vide intentionnellement

7.3. Les graphiques

Les graphiques représentent la forme de présentations des données plus attractive et plus persuasif. Leur utilisation est plutôt recommandable si l'analyse statistique peut compter avec peu de temps pour présenter la thèse d'analyste. La présentation des résultats dans les médias comme la télévision est inimaginable sans l'utilisation des diagrammes. Les graphiques, quand même ne sont pas auto-explicatives, l'analyste doit s'occuper de choisir le diagramme plus approprié pour présenter les résultats et doit prendre soin de commenter les graphiques de façon nécessaire, soit avec des outils graphiques (légendes etc.) ou avec du texte simple. Vous devez apprendre dans ce chapitre l'utilisation des graphiques, leur modification et leur importance dans le contexte de la présentation de Vos données.

Il existe trois façons de créer des diagrammes et des graphiques dans SPSS:

1. Les commandes du menu **Graphes**, dans le menu principal, produisent des diagrammes ou des graphiques. La barre de menus principale est disponible lorsque l'éditeur de données, une fenêtre de résultats, ou une fenêtre de syntaxe est active.
2. Certaines procédures statistiques permettent de créer des graphiques. Par exemple des boxplots avec la procédure *EXPLORE* ou des nuages de points à partir de *Régression linéaire*.
3. Au moyen des procédures indiquées dans le sous-menu [Interactif] du menu Graphes, vous pouvez, avant de créer le diagramme, en spécifier de nombreuses options ainsi que des variables. Par la suite, le diagramme peut être modifié dans le Viewer. Lorsque vous sélectionnez un type de diagramme dans le sous-menu Interactif du menu Graphes, la boîte de dialogue correspondante s'affiche avec l'onglet Affecter les variables sélectionné.

Il existe deux manières principales de modifier les graphes

1. Double-cliquez sur le diagramme souhaité. Le diagramme s'affiche dans une fenêtre. Vous pouvez modifier le diagramme par l'intermédiaire des menus, de la barre d'outils ou encore en double-cliquant sur l'objet à modifier. Pour sélectionner un objet du diagramme, cliquez dessus. Par exemple, cliquer sur le titre de la légende permet de l'entourer d'une zone de sélection.
2. Pour modifier un diagramme interactif, double-cliquez sur le diagramme pour l'activer. Puis double-cliquez sur l'élément dont vous souhaitez modifier les attributs. Par exemple, pour afficher des valeurs sur un bâton d'un diagramme, double-cliquez sur ce bâton. La boîte de dialogue Bâtons s'affiche. Dans le groupe Etiquettes des bâtons, cliquez sur Effectif ou sur Valeurs. Comme ce sont les ventes totales qui sont représentées par les bâtons de ce diagramme, cliquez sur Valeurs. Les valeurs des ventes totales sont maintenant affichées en haut de chaque bâton du diagramme.

7.3.1. Création d'un diagramme

1. Choisissez le type de diagramme désiré dans le menu [Graphes].
2. Cliquez sur l'icône du type de diagramme désiré.
3. Choisissez l'option de récapitulatif des données qui convient le mieux à vos données.
4. Cliquez sur Définir.

7.3.2. Les principaux diagrammes

Bâtons
Dispersion
Courbes
Histogramme
Aires
P-P gaussien
Secteurs
Q-Q gaussien
Cycle
Séquence
Pareto
Autocorrélations
Contrôle
Corrélations croisées
Boîte
Spectral
Barre d'erreurs

Les options disponibles dépendent du type de diagramme. Pour plus d'informations sur le type de diagramme courant, cliquez sur Aide dans la boîte de dialogue.

7.3.3. Edition d'un diagramme

Double-cliquez sur le diagramme dans le Navigateur.

Le diagramme est activé dans la fenêtre Diagrammes, où vous pouvez l'éditer en utilisant les menus et la barre d'outils. Quelques conseils généraux :

Cliquez une fois sur un objet (une barre ou une ligne, par exemple) pour le sélectionner. Vous pouvez ensuite modifier les attributs tels que les couleurs ou le motif de remplissage de l'objet.

Double-cliquez sur un objet pour ouvrir une boîte de dialogue qui vous permet de le modifier.

7.3.4. Création de diagrammes interactifs depuis le menu Graphes

Au moyen des procédures indiquées dans le sous-menu Interactif du menu Graphes, vous pouvez, avant de créer le diagramme, en spécifier de nombreuses options ainsi que des variables. Par la suite, le diagramme peut être modifié dans le Viewer. Lorsque vous sélectionnez un type de diagramme dans le sous-menu Interactif du menu Graphes, la boîte de dialogue correspondante s'affiche avec l'onglet Affecter les variables sélectionné.

Pour créer un diagramme, il suffit d'affecter des variables, mais vous pouvez également parcourir les autres onglets si vous désirez personnaliser votre diagramme.

Liste source: les variables disponibles pour le type de diagramme figurent dans la liste source. Vous pouvez modifier leur ordre d'affichage dans la liste en cliquant avec le bouton droit de la souris sur une variable et en choisissant l'option Trier par nom, Trier par fichier, ou Trier par type dans le menu contextuel. Vous pouvez également choisir Afficher les étiquettes des variables ou Afficher les noms des variables. Dans le menu contextuel, ces entrées et dimensions de données remplacent les éléments Qu'est-ce que c'est? et Informations sur les variables apparaissant dans les autres boîtes de dialogue du système.

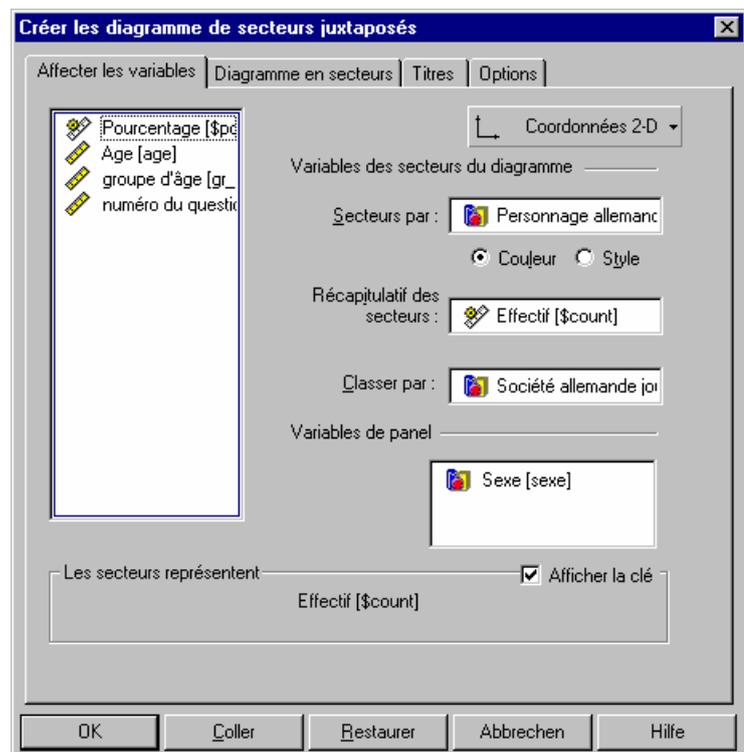


Figure 26: Création d'un graphe interactif

Affectation de variables: vous pouvez affecter les variables des axes en les faisant glisser depuis la liste source et en les déposant sur les destinations des axes. Vous pouvez également faire glisser les variables d'une destination à une autre. Si une variable est déposée sur une

destination à laquelle une autre variable est déjà affectée, les deux variables permutent. Vous pouvez également affecter des variables multiples à chaque axe en cliquant avec la souris tout en maintenant la touche [Ctrl] enfoncée.

Fonctions élémentaires : une variable d'échelle affectée à un axe des dépendantes peut comporter une fonction élémentaire destinée à de nombreux types de diagrammes. Par exemple, la fonction élémentaire Moyennes sera souvent appropriée pour la variable SALAIRE. La fonction élémentaire pour une variable d'échelle peut être spécifiée par une sélection dans une liste de fonctions élémentaires, apparaissant dans la partie inférieure de la boîte de dialogue.

Dernière étape de la création du diagramme : dès qu'un nombre minimum de variables est affecté, le diagramme peut être créé et affiché dans le Viewer. Le bouton OK est toujours activé. Si vous cliquez sur OK et que le programme détermine que d'autres informations sont nécessaires, une boîte de message vous indique les éléments requis. Avant de cliquer sur OK, vous souhaitez peut-être affiner le diagramme ; pour cela, sélectionnez les options qui se trouvent sous les autres onglets de la boîte de dialogue.

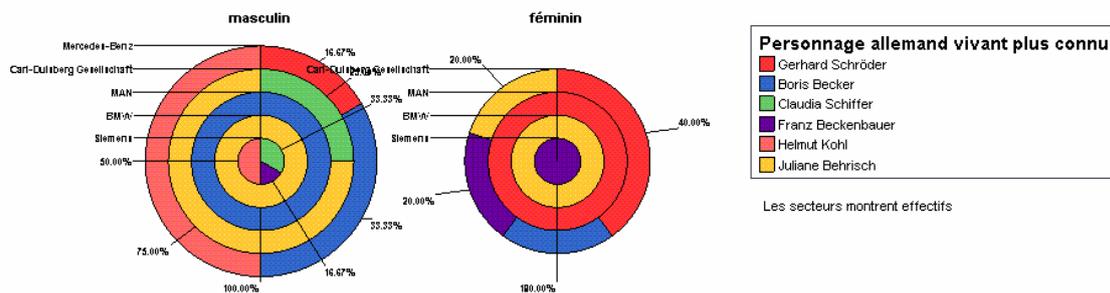


Figure 27: Le Résultat - les secteurs juxtaposés

Exercice: Enquête Camerounaise Auprès des Ménages (ECAM):

Dans l'annexe Vous trouverez la description de l'Enquête Camerounaise Auprès des Ménages. Les données principales Vous trouverez sous le nom CHS01_02.SAV le fichier individuel et CHS02_01.SAV le fichier du ménage et le fichier combiné CHS03_01.SAV.

Utilisez SPSS pour manipuler et analyser les données:

Listez les données présentant:

1. L'effectif des niveaux atteints plus élevé d'éducation
2. Les moyennes des dépenses pour les médicaments par groupe d'âge et sexe
3. L'effectif des niveaux atteints plus élevé d'éducation par sexe du chef de ménage

Présentez les données en tableaux :

1. L'occupation principale (fréquence)
2. L'occupation principale par niveau atteint plus élevé d'éducation
3. L'occupation principale par niveau atteint plus élevé d'éducation et sexe

Présentez les données en diagrammes:

1. Les groupes d'âges (comparez l'histogramme avec la distribution normale)
2. Les motifs de l'absence de l'école (graphiques en secteurs)
3. L'occupation principale par niveau atteint plus élevé d'éducation et sexe
4. Le niveau atteint plus élevé d'éducation par sexe du chef de ménage

Page laissée vide intentionnellement

8. Comparaison entre les possibilités de SPSS et EXCEL

Microsoft EXCEL propose un ensemble d'outils d'analyse de données, appelé Utilitaire d'analyse, que vous pouvez utiliser pour éviter certaines étapes lors du développement d'analyses statistiques ou techniques complexes. Lorsque vous utilisez l'un de ces outils, vous devez fournir les données et les paramètres pour chaque analyse. L'outil utilise alors les fonctions macro statistiques ou techniques appropriées, puis affiche les résultats dans une table de sortie. Certains outils génèrent des graphiques en plus de tables de sortie.

Pour afficher la liste des outils d'analyse disponibles, cliquez dans le menu Outils sur Utilitaire d'analyse. Si la commande Utilitaire d'analyse ne figure pas dans le menu Outils, exécutez le programme d'installation pour installer l'Utilitaire d'analyse. Une fois l'Utilitaire d'analyse installé, sélectionnez-le dans le Gestionnaire de macros complémentaires. Pour plus d'informations sur l'installation de l'Utilitaire d'analyse, cliquez sur .

L'utilisation de ces outils implique une bonne connaissance des domaines statistiques ou techniques pour lesquels vous souhaitez développer des analyses.

Remarque : Microsoft EXCEL propose de nombreuses fonctions supplémentaires de feuille de calcul statistiques, financières et techniques. Pour obtenir la liste des fonctions de feuille de calcul disponibles, cliquez dans la barre de formule sur Zone de formule, puis sur la flèche pointant vers le bas dans Fonction .

Tandis que SPSS est un logiciel spécialisée pour l'analyse des statistique, EXCEL est un logiciel pour presque tout les taches de calculer, analyse et présentation des données.

Bien sur que la gamme des fonctions statistique de EXCEL est plus limitée en comparaison avec SPSS, la facilite d'utilisation et la incorporation facile dans les logiciel de WINDOWS, peut être la caractéristique décisive pour le choix de EXCEL.

Imaginez une présentation avec des donnes statistiques, quand ces présentation doivent être répétés régulièrement :

Exemple : L'enquête semestriel d'agriculture

Avec un lien permanent d'une feuille de calcul EXCEL dans le fichier POWERPOINT Vous pouvez actualiser régulièrement ces présentations

Avec les résultats de SPSS, ceci est plus compliques et les calculs statistiques doivent être répétés régulièrement et copier dans la présentation

Comment appelez Vous les fonctions statistiques en EXCEL ? :

Très simple : Vous choisissez la fonction pour analyser les statistiques descriptives des contenus des cellules

Ceci est très semblable à SPSS

À propos des fonctions statistiques

Les fonctions statistiques de la feuille de calcul permettent d'effectuer une analyse statistique des plages de données. Par exemple, une fonction statistique peut fournir des informations statistiques sur une droite tracée au travers d'un groupe de valeurs, telles que la pente de la ligne et l'intersection des y, ou à propos des points constituant réellement la droite.

8.1. Premier exemple : Les statistiques descriptives

Vous connaissez les fonctions DESCRIPTIVES et EXPLORE en SPSS

1. Chargez un fichier des données en EXCEL (p.e. utilisez la option [Fichier/Exportation] en SPSS
2. Appelez [Outils/ Utilitaire d'analyse/..]

Outil d'analyse Statistiques descriptives

Cet outil d'analyse génère un rapport de statistiques à une seule variable pour les données contenues dans la plage d'entrée, fournissant ainsi des informations sur la tendance centrale et la dispersion de vos données. Pour plus d'informations :

Plage d'entrée

Tapez la référence des cellules contenant la plage de données que vous souhaitez analyser. La référence doit contenir plusieurs plages de données adjacentes disposées en lignes ou en colonnes.

Groupées par

Pour indiquer si les données contenues dans la plage d'entrée sont disposées en lignes ou en colonnes, cliquez sur Lignes ou Colonnes.

Intitulés en première ligne/ Intitulés en première colonne

Si la première ligne de votre plage d'entrée contient des étiquettes, activez la case à cocher Intitulés en première ligne. Si ces étiquettes se trouvent dans la première colonne de votre plage d'entrée, activez la case à cocher Intitulés en première colonne. Désactivez ces cases à cocher si votre plage d'entrée ne contient pas d'étiquette. Dans ce cas, Microsoft EXCEL génère les étiquettes de données appropriées pour la table de sortie.

Niveau de confiance pour la moyenne

Activez cette case à cocher si vous souhaitez inclure une ligne dans la table de sortie pour le niveau de confiance de la moyenne. Dans la zone, tapez le niveau de confiance voulu. Par exemple, la valeur 95 % calcule le niveau de confiance de la moyenne avec une signification de 5 %.

K ème maximal

Activez cette case à cocher si vous souhaitez inclure une ligne dans la table de sortie pour la plus grande valeur de chaque plage de données. Dans la zone, tapez le nombre à utiliser pour k. Si k est égal à 1, cette ligne contient la valeur maximale de la série de données.

K ème minimal

Activez cette case à cocher si vous souhaitez inclure une ligne dans la table de sortie pour la plus petite valeur de chaque plage de données. Dans la zone, tapez le nombre à utiliser pour k. Si k est égal à 1, cette ligne contient la valeur minimale de la série de données.

Plage de sortie

Tapez la référence de la cellule située dans le coin supérieur gauche de la table de sortie. Cet outil génère deux colonnes d'informations pour chaque série de données. La colonne de gauche contient les étiquettes statistiques et la colonne de droite, les statistiques. Microsoft EXCEL écrit une table contenant deux colonnes de statistiques pour chaque colonne ou ligne de la plage d'entrée, selon l'option que vous avez sélectionnée sous Groupées par.

Insérer une nouvelle feuille

Cliquez sur cette option pour insérer une nouvelle feuille de calcul dans le classeur en cours et coller les résultats à partir de la cellule A1 de cette nouvelle feuille de calcul. Pour attribuer un nom à la nouvelle feuille de calcul, tapez le nom souhaité dans la zone appropriée.

Créer un nouveau classeur

Cliquez sur cette option pour créer un classeur et coller les résultats dans une nouvelle feuille de calcul du nouveau classeur.

Rapport détaillé

Activez cette case à cocher si vous souhaitez que Microsoft EXCEL génère un champ pour chacune des statistiques suivantes dans la table de sortie : Moyenne, Erreur type (de la moyenne), Médiane, Mode, Écart type, Variance, Kurtosis, Asymétrie, Plage, Minimum, Maximum, Somme, Nombre, Plus grande valeur (#), Plus petite valeur (#) et Niveau de confiance.

Resume :

- SPSS permet les statistiques descriptives plus détaillées et ils être modifiés individuellement (EXCEL -, SPSS +)
- EXCEL permet plus de flexibilité dans l'utilisation interactive, plus de facilité pour le débutant (EXCEL +, SPSS -)

8.2. Deuxième exemple: Les tableaux croisés

Vous connaissez les fonctions CROSSTABS et TABLES de SPSS
EXCEL permet les tableaux croisés très sophistiqués avec la fonction tableaux croisés dynamiques

Généralement Vous utilisez l'Assistant Tableau croisé dynamique pour créer ces tableaux

Ceci est un tableau typique avec ces éléments typiques:

Fonctions de synthèse pour l'analyse de données

SOMME	La somme des valeurs. C'est la fonction par défaut pour les données source de type numérique.
NB	Le nombre des éléments. La fonction NB est similaire à la fonction HNBVAL des feuilles de calcul. Nb est la fonction par défaut lorsque les données source ne sont pas de type numérique.
MOYENNE	La moyenne des valeurs.
MAX	La valeur la plus élevée.
MIN	La valeur la moins élevée.
PRODUIT	Le produit des valeurs.
NBVAL	Le nombre de lignes contenant des données numériques. La fonction NBVAL est similaire à la fonction NB des feuilles de calcul.
ECARTYPE	Une estimation de l'écart type d'une population pour laquelle l'échantillon correspond à la totalité des données à synthétiser.
ECARTYPEP	L'écart type d'une population qui constitue la totalité des données à synthétiser.
VAR	Une estimation de la variance d'une population pour laquelle l'échantillon correspond à la totalité des données à synthétiser.
VAR.P	La variance d'une population qui constitue la totalité des données à synthétiser.

Vous disposez aussi des fonctions Synthèse de données et calculs dans un tableau croisé dynamique

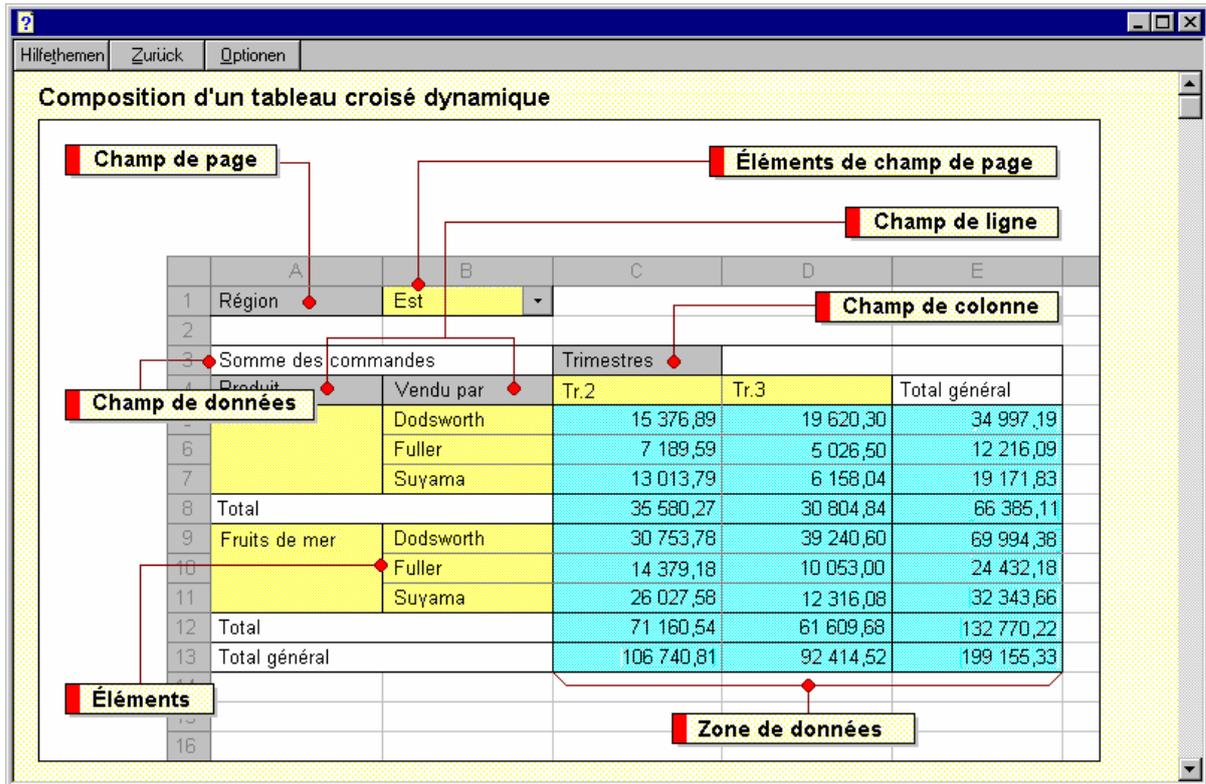


Figure 28: Les éléments d'un tableau croisé dynamique

Vous pouvez personnaliser les champs de données d'un tableau croisé dynamique en utilisant un autre type de calcul, en remplacement ou en complément de la fonction de synthèse utilisée par défaut. Si votre travail nécessite des calculs plus complexes que ceux qui sont effectués par les fonctions de synthèse, vous avez la possibilité de créer vos propres formules.

- Pour effectuer les opérations de base par exemple, Somme, Moyenne, Minimum et Maximum servez-vous d'une fonction de synthèse.
- Pour comparer ou indexer deux champs, ou encore afficher un pourcentage, utilisez un calcul personnalisé..
- Pour créer votre propre formule destinée à exploiter les données du tableau, utilisez un champ calculé ou un élément calculé d'un champ.

Enfin la création des tableaux (est successivement des graphiques) n'est pas plus compliqués que en SPSS, mais les but des deux programmes sont différents :

Pour les tableaux croises nous constatons le suivant :

	SPSS	EXCEL
--	------	-------

Presentation	+	++
Analyse statistique	++	+
Flexibilites de manipuler les donnees	+	++
Quantites des donnees	++	+/-
Integration dans des systemes MICROSOFT	+	++
Integration dans des systemes d'analyse	++	+/-