

UNIVERSITE POLYTECHNIQUE
DE BOBO-DIOULASSO
(U.P.B)

ECOLE SUPERIEURE D'INFORMATIQUE
(E.S.I)

Cycle des Ingénieurs de Travaux
Informatiques
Option analystes-programmeurs
(CITI/AP)

Année académique 2003-2004



SYSTEME DES NATIONS UNIES
(S.N.U)



PROGRAMME DES NATIONS UNIES
POUR LE DEVELOPPEMENT
(PNUD)

PROJET DE FIN D'ETUDES

THEME:

DOCUMENTATION DE BASES DE DONNEES ET
REALISATION DE GUIDES D'UTILISATEURS
POUR LES LOGICIELS *SPSS* ET *STATA*

Dossier d'activité

Présenté par :
Noémie Yétema NIDJERGOU
Yaya SIRIMA

Maître de stage :

M. Laouli SANOU
Ingénieur Informaticien
PNUD

Superviseur :

M. Moussa DAVOU
Ingénieur Informaticien
ONATEL

SOMMAIRE

| | |
|---|------------|
| Introduction ----- | 2 |
| CHAPITRE 1 : Enquête prioritaire 1998 (EPII) ----- | 3 |
| CHAPITRE 2 : Enquête corruption ----- | 73 |
| CHAPITRE 3 : Guide SPSS ----- | 82 |
| CHAPITRE 4 : Guide STATA ----- | 121 |
| Conclusion ----- | 153 |

Introduction

Science de traitement automatique de l'information, des données, l'informatique se révèle sous deux aspects : celui merveilleux de la technologie avec fiabilité, performance, intégration, convivialité des systèmes de gestion des bases de données disponibles sur le marché ; l'autre moins avenant avec une incompréhension des possibilités offerte par les logiciels et des difficultés liées à leur exploitation.

Le présent rapport a pour objectif de présenter les activités menées dans le but de présenter l'aspect merveilleux de l'exploitation des bases de données. Il contiendra comme prévus une documentation des bases de données disponibles et les guides d'utilisateurs des logiciels SPSS et STATA.

CHAPITRE 1

ENQUETE PRIORITAIRE 1998

SOMMAIRE

| | |
|---|-----------|
| I. Description des différents fichiers..... | 5 |
| <i>I.1. Ep2fiid : Fichier Identification des ménages</i> | <i>5</i> |
| <i>I.2. Ep2fi0 : Fichier de description du chef du ménage(CM)</i> | <i>10</i> |
| <i>I.3. Ep2fi1 : Fichier de description de la liste des ménages.</i> | <i>13</i> |
| <i>I.4. Ep2fi2 : Fichier santé</i> | <i>16</i> |
| <i>I.5. Ep2fi3 : Fichier Education</i> | <i>19</i> |
| <i>I.6. Ep2fi4 : Fichier Alphabétisation et Migration</i> | <i>22</i> |
| <i>I.7. Ep2fi5a : Fichier Chômage et Emploi principal</i> | <i>24</i> |
| <i>I.8. Ep2fi5b : Fichier Emploi secondaire</i> | <i>30</i> |
| <i>I.9. Ep2fi5c : Fichier Emploi précédent</i> | <i>33</i> |
| <i>I.10. Ep2fi6 : Fichier Logement et confort</i> | <i>36</i> |
| <i>I.11. Ep2fi7a : Fichier Bétail</i> | <i>42</i> |
| <i>I.12. Ep2fi7b : Fichier Production campagne 98/99</i> | <i>44</i> |
| <i>I.13. Ep2fi7c : Fichier Production campagne 98/99</i> | <i>48</i> |
| <i>I.14. Ep2fi8 : Fichier Entreprise non agricole</i> | <i>51</i> |
| <i>I.15. Ep2fi9 : Fichier Accès aux services de base les plus proche</i> | <i>54</i> |
| <i>I.16. Ep2fi10a : Fichier Dépense scolaire de l'année précédente</i> | <i>60</i> |
| <i>I.17. Ep2fi10b : Fichier Dépenses alimentaires des 30 derniers jours</i> | <i>64</i> |
| <i>I.18. Ep2fi10c : Fichier Dépenses sur les produits non alimentaires</i> | <i>65</i> |
| <i>I.19. Ep2fi11 : Fichier Revenu des ménages</i> | <i>67</i> |
| <i>I.20. Ep2fi12 : Fichier Avoir du ménage</i> | <i>69</i> |
| <i>I.21. Ep2fi13 : Fichier Anthropométrie</i> | <i>70</i> |
| II. Remarques et suggestions..... | 68 |

I. Description des différents fichiers

Les enquêtes sur les conditions de vie des ménages représentent des outils de décision d'une importance capitale en terme de politique économiques.

L'analyse des résultats de l'Enquête Prioritaire II (EPII) a pour objectif principalement la description des conditions de vie des ménages à travers l'élaboration d'indicateurs socioéconomiques et démographiques.

La base de donnée de l'EP II est disponible en format ASCII et aussi sous les logiciels SPSS, STATA, SAS. L'organisation des données est identique dans chaque format ; Ainsi, une documentation de cette base de donnée permettra de faciliter l'analyse en mettant à la disposition des analystes une description sommaire des différents fichiers.

I.1. Ep2fiid : Fichier Identification des ménages

Ep2fiid donne une description générale du ménage. Il précise la situation géographique et socio-économique du ménage, ainsi que les conditions dans lesquelles l'enquête s'est déroulée.

◆ Chemin

.. \\DATA\...\Ep2fiid

◆ Nombre d'observations

Nombre d'observations : 4876

◆ Description des variables

g1 : Label : **Région économique**

Type : Numérique ordinal

Taille: 2

Valeur par défaut: *Néant*

Valeurs représentées :

- 1 Ouest
- 2 Nord-Ouest
- 3 Sahel
- 4 Est
- 5 Sud-Ouest
- 6 Centre-Nord
- 7 Centre-Ouest
- 8 Centre
- 9 Nord
- 10 Centre-Est

g2 : Label : **Province**

Type : Numérique ordinal

Taille: 2

Valeur par défaut: *Néant*

Valeurs représentées :

- 1 Bam
- 2 Bazèga
- 3 Bougouriba
- 4 Boulgou
- 5 Boulkiemdé
- 6 Comoé
- 7 Ganzourgou
- 8 Gnagna
- 9 Gourma
- 10 Houet
- 11 Kadiogo
- 12 Kéné Dougou
- 13 Kossi
- 14 Kouritenga
- 15 Mouhoun
- 16 Nahouri
- 17 Namentenga
- 18 Oubritenga
- 19 Oudalan
- 20 Passoré
- 21 Poni
- 22 Sanguié
- 23 Sanmatenga
- 24 Séno
- 25 Sissili
- 26 Soum
- 27 Sourou
- 28 Tapoa
- 29 Yatenga
- 30 Zoundwéogo

g3 : Label : **Département**

Type : Numérique ordinal

Taille: 2

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

g4 : Label : **Milieu de résidence**

Type : Numérique ordinal

Taille: 1

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

1 Urbain

2 Rural

g5 : Label : **N°ZD**

Type : Numérique

Taille: 3
Valeur par défaut: *Néant*
Valeurs représentées : *Néant*

g6 : Label : **Ville/Secteur**
Type : Numérique ordinal
Taille: 1
Valeur par défaut: *Néant*
Valeurs représentées : *Néant*

g8 : Label : **N°Ménage dans la Zone de Dénombrement(ZD)**
Type : Numérique
Taille: 2
Valeur par défaut: *Néant*
Valeurs représentées : *Néant*

g7 : Label : **N°Concession**
Type : Numérique
Taille: 3
Valeur par défaut: *Néant*
Valeurs représentées : *Néant*

q10 : Label : **Nombre de ménage dans concession**
Type : Numérique ordinal
Taille: 2
Valeur par défaut: *Néant*
Valeurs représentées : *Néant*

q12 : Label : **Résultat interview**
Type : Numérique ordinal
Taille: 1
Valeur par défaut: *Néant*
Valeurs représentées :
1 Entièrement rempli
2 Incomplet
9 Non déterminé(ND)

q13 : Label : **Initialement tiré ?**
Type : Numérique ordinal
Taille: 1
Valeur par défaut: *Néant*
Valeurs représentées :
1 Oui
2 De remplacement

q14 : Label : **Sélectionné pour anthropologie.**
Type : Numérique ordinal
Taille: 1
Valeur par défaut: *Néant*
Valeurs représentées :

- 1 Oui
- 2 Non
- 9 N.D

q15 : Label : **N°Questionnaire dans ménage**

Type : Numérique ordinal

Taille: 2

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

q16eng : Label : **Code enquêteur**

Type : Numérique

Taille: 2

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

q17eng : Label : **Jour de l'enquête**

Type : Numérique

Taille: 2

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

q18eng : Label : **Mois de l'enquête**

Type : Numérique ordinal

Taille: 2

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

q19eng : Label : **Observation**

Type : Numérique ordinal

Taille: 1

Valeur par défaut: *Néant*

Valeurs représentées :

- 1 Aucun Problème
- 2 Des réticences
- 3 Des dissimulations
- 9 N.D

gse : Label : **Groupe socio-économique**

Type : Numérique ordinal

Taille: 2

Valeur par défaut: *Néant*

Valeurs représentées :

- 1 Salariés du public
- 2 Salariés du privé structurel
- 3 Indépendants ou employeurs non agricoles
- 4 Aides familiaux, bénévoles et apprentis
- 5 Agriculteurs de coton
- 6 Agriculteurs vivriers

- 7 Chefs de ménage inactifs
- 8 Salariés du privé non structurés
- 9 Chômeurs

pond : Label : **Coefficient de pondération**

Type : Numérique

Taille: 7

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

loyer : Label : **Loyer**

Type : Numérique

Taille: 9

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

tail : Label : *Néant*

Type : Numérique

Taille: 7

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

I.2. Ep2fi0 : Fichier de description du chef de ménage(CM)

Ep2fi0 donne une description du chef de ménage, des responsables du ménage et des conditions d'interview, pour l'évaluation de la validité des réponses.

◆ Chemin

.. \DATA\ ... \ Ep2fi0

◆ Nombre d'observations

Nombre d'observations : 8478

◆ Description des variables

g5 : Idem que dans Ep2fiid

g8 : Idem que dans Ep2fiid

s0q1 : Label : **Ethnie /Nationalité du CM**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation :

| | |
|----|-------------|
| 0 | Etranger |
| 1 | Mossi |
| 2 | Dioula |
| 3 | Peulh |
| 4 | Bissa |
| 5 | Gourmantché |
| 6 | Gouroussi |
| 7 | Bobo |
| 8 | Samo |
| 9 | Sénoufo |
| 10 | Lobi |
| 11 | Dagari |
| 12 | Gouin |
| 13 | Autres |

s0q2 : Label : **Religion du chef de ménage**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

| | |
|---|----------------|
| 1 | Musulman |
| 2 | Catholique |
| 3 | Protestant |
| 4 | Animiste |
| 5 | Autre religion |

6 Sans religion

s0q3 : Label : **Existence du ménage il y'a douze (12) mois**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

1 oui

2 non

9 N.D

s0q4 : Label : **Durée d'existence**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation : *Néant*

s0q5 : Label : **Nombre de couple**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation : *Néant*

s0q6 : Label : **Présence du chef de ménage**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

1 CM présent

2 CM absent

9 N.D

s0q7 : Label : **Responsable en cas d'absence du CM**

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation : *Néant*

s0q8 : Label : **Répondant**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation : *Néant*

s0q9 : Label : **Langue utilisée**

Type : numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation :

- | | |
|----|---------------------------|
| 1 | Moré |
| 2 | Dioula |
| 3 | Fulfuldé |
| 4 | Bissa |
| 5 | Gulmancéma |
| 6 | Gourounsi |
| 7 | Bobo |
| 8 | Samo |
| 9 | Sénoufo |
| 10 | Lobi |
| 11 | Dagara |
| 12 | Gouin |
| 13 | Autres langues du Burkina |
| 14 | Français |

s0q10 : Label : **Appel à un interprète**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

- | | |
|---|-----|
| 1 | oui |
| 2 | non |

I.3. Ep2fi1 : Fichier de description de la liste des ménages.

Ep2fi1 renvoie la liste des différents membres du ménage et des informations sur chacun d'eux. Il prend en compte le type de résidence (permanent ou alternatif) et retient uniquement ceux qui sont présent dans le ménage au moment de l'interview.

◆ Chemin

.. \DATA\...\ Ep2fi1

◆ Nombre d'observations

Nombre d'observations : 63509

◆ Description des variables

g5 : Idem que dans Ep2fiid

g8 : Idem que dans Ep2fiid

s1q1 : Label : **Numéro d'ordre**

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation : *Néant*

s1q3 : Label : **Situation de résidence**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

- | | |
|---|----------------------|
| 1 | Résidence Permanente |
| 2 | Résidence Alternatif |

s1q4 : Label : **Lien de parenté**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation :

- | | |
|---|-------------------|
| 1 | CM |
| 2 | Conjoint(e) |
| 3 | Fils/Fille |
| 4 | Frère/Sœur |
| 5 | Neveu/Nièce |
| 6 | Petit-fils/ Fille |
| 7 | Père/Mère |

s1q5 : Label : Situation matrimoniale

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

| | |
|---|-----------------------|
| 1 | Célibataire |
| 2 | Marié(e) monogame |
| 3 | Marié(e) polygame |
| 4 | Veuf(ve) |
| 5 | Divorcé(e)/Séparé(e) |
| 6 | Union libre |

s1q6 : Label : Sexe

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

| | |
|---|----------|
| 1 | Masculin |
| 2 | Féminin |

s1q7 : Label : Mois de naissance

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s1q8 : Label : Année de naissance

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s1q9 : Label : Age en mois

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s1q10 : Label : Age en année

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s1q11 : Label : Handicap

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- | | |
|---|----------------|
| 1 | Aucun handicap |
| 2 | Moteur |
| 3 | Manchot |
| 4 | Visuel |
| 5 | Lépreux |
| 6 | Mental |
| 7 | Autres |
| 8 | N.D |

tajust : Label : **Taille ajustée**
Type : Numérique décimal
Taille : 8
Valeur par défaut : *Néant*
Valeurs et représentations : *Néant*

tajuste : Label : *Néant*
Type : Numérique décimal
Taille : 8
Valeur par défaut : Néant
Valeurs et représentations : Néant

I.4. Ep2fi2 : Fichier santé

Ep2fi2 donne une description de l'état de santé de l'individu, les différentes procédures de soins, ainsi que les dépenses engendrées.

◆ Chemin

.. \\DATA\...\ Ep2fi2

◆ Nombre d'observations

Nombre d'observations : 63357

◆ Description des variables

g5 : Idem que dans Ep2fiid

g8 : Idem que dans Ep2fiid

s1q1 : Idem que dans Ep2fi1

s2q2 : Label : **Tombé malade 15 derniers jours**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

| | |
|---|-----|
| 1 | oui |
| 2 | non |
| 9 | N.D |

s2q3 : Label : **Avoir consulté raison de santé**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

| | |
|---|-----|
| 1 | oui |
| 2 | non |
| 9 | N.D |

s2q4 : Label : **Nombre de consultation**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation : *Néant*

s2q5 : Label : **Personne consultée**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation :

- 1 Guérisseur traditionnel/Marabout
- 2 Médecin privé
- 3 Pharmacien privé
- 4 Infirmier privé
- 5 CHN (Centre Hospitalier National)
- 6 CHR (Centre Hospitalier Régional)
- 7 CMA/CM
- 8 CSPS
- 9 Privé / ONG
- 10 Matrone traditionnelle
- 11 Autres

s2q6 : Label : **Personne consultée**

Type : Numérique ordinal

Taille :2

Valeur par défaut : *Néant*

Valeurs et représentation :

- 1 Guérisseur traditionnel/Marabout
- 2 Médecin privé
- 3 Pharmacien privé
- 4 Infirmier privé
- 5 CHN
- 6 CHR
- 7 CMA/CM
- 8 CSPS
- 9 Privé / ONG
- 10 Matrone traditionnelle
- 11 Autres

s2q7 : Label : **Dépense en consultation**

Type : Numérique

Taille :2

Valeur par défaut : *Néant*

Valeurs et représentation : *Néant*

s2q8 : Label : **Dépense en analyses médicales**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs et représentation : *Néant*

s2q9 : Label : **Raison de non-consultation**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

- 1 Eloignement

- 2 Pas d'argent
- 3 Coût élevé
- 4 Pas grave/ bénin
- 5 Automédication
- 6 Mauvais accueil
- 7 Mauvaise qualité
- 8 Autres

s2q10 : Label : **Dépenses en médicament**

Type : Numérique

Taille :9

Valeur par défaut : *Néant*

Valeurs et représentation : *Néant*

s2q11 : Label : **Consultation pour soins préventifs**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

- 0 Pas de consultation préventive
- 1 Visite préventive
- 2 Suivi croissance
- 3 Autres consultations

I.5. Ep2fi3 : Fichier Education

Ce fichier renferme des informations sur le type de formation scolaire ou académique suivie, les établissements fréquentés et le niveau d'instruction.

◆ Chemin

.. \DATA\...\Ep2fi3

◆ Nombre d'observations

Nombre d'observations : 50721

◆ Description des variables

Indiv : Label : *Néant*

Type : Numérique

Taille : 8

Valeur par défaut : *Néant*

Valeurs et représentation : *Néant*

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s1q1 : Idem que dans **Ep2fi1**

s3q2 : Label : **Fréquente en 97/98**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

1 Oui

2 Non

9 N.D

s3q3 : Label : **Classe fréquentée**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation :

1 Maternelle

2 CP

3 CE1

4 CE2

5 CM1

6 CM2

7 6ème

8 5ème

- 9 4ème
- 10 3ème
- 11 2nd
- 12 1ère
- 13 Terminale
- 14 Supérieur
- 15 Professionnel avant CEPE
- 16 Professionnel Secondaire après CEPE
- 17 Professionnel Secondaire après BEPC
- 99 ND

s3q4 : Label : **Type d'établissement**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

- 1 Public
- 2 Privé laïc
- 3 Privé catholique
- 4 Privé protestant
- 5 Medersa
- 6 CFJA
- 7 CEBNF
- 8 Cours du soir

s3q5 : Label : **Fréquentait en 96/97**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

- 1 Oui
- 2 Non
- 9 ND

s3q6 : Label : **Classe fréquentée 96/97**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation :

- 1 Maternelle
- 2 CP
- 3 CE1
- 4 CE2
- 5 CM1
- 6 CM2
- 7 6ème
- 8 5ème
- 9 4ème
- 10 3ème

- 11 2nd
- 12 1ère
- 13 Terminale
- 14 Supérieur
- 15 Professionnel avant CEPE
- 16 Professionnel Secondaire après CEPE
- 17 Professionnel Secondaire après BEPC
- 99 ND

s3q7: Label : **Dernière classe achevée**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentation :

- 0 Aucune classe
- 1 Maternelle
- 2 CP
- 3 CE1
- 4 CE2
- 5 CM1
- 6 CM2
- 7 6ème
- 8 5ème
- 9 4ème
- 10 3ème
- 11 2nd
- 12 1ère
- 13 Terminale
- 14 Supérieur
- 15 Professionnel avant CEPE
- 16 Professionnel Secondaire après CEPE
- 17 Professionnel Secondaire après BEPC
- 99 ND

s3q8: Label : **Raison**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentation :

- 1 Renvoyé
- 2 Fin des études
- 3 Veut travailler
- 4 Coût élevé
- 5 Pas nécessaire
- 6 Pas d'école/éloigné
- 7 Manque de cantine
- 8 Maladie
- 9 ND

I.6. Ep2fi4 : Fichier Alphabétisation et Migration

Ce fichier est rempli pour les personnes ayant au moins 10 ans. Il renferme des informations sur la langue d'alphabétisation ainsi que sur les causes et caractéristiques de migration des individus.

◆ Chemin

.. \DATA\...\Ep2fi4

◆ Nombre d'observations

Nombre d'observations : 42119

◆ Description des variables

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s1q1 : Idem que dans **Ep2fi1**

s4q2 : Label : **Langue**

Type: Numérique ordinal

Taille: 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 0 Ne sais pas lire/écrire
- 1 Français
- 2 Moré
- 3 Dioula
- 4 Fulfudé
- 5 Gulmacéma
- 6 Autres langues du BF
- 7 Autres langues étrangères.

s4q3 : Label : **Avoir résidé hors du pays 6 mois ou plus**

Type: Numérique ordinal

Taille: 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Oui
- 2 Non
- 9 ND

s4q4 : Label: **Durée en années**

Type: Numérique ordinal

Taille: 2

Valeur par défaut : *Néant*

Valeurs et représentations :

0 Moins d'un an

88 Natif

99 Ne sais pas

s4q5 : Label : **Province/Pays**

Type: Numérique

Taille: 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s4q6 : Label : **Milieu**

Type: Numérique ordinal

Taille: 1

Valeur par défaut : *Néant*

Valeurs et représentations :

1 Zone rurale

2 Chef lieu province rurale

3 Chef lieu province urbaine

4 Autres villes

9 ND

s4q7 : Label : **Principal motif**

Type: Numérique ordinal

Taille: 1

Valeur par défaut : *Néant*

Valeurs et représentations :

1 Recherche emploi

2 Affectation

3 A suivi la famille

4 Mariage

5 Poursuite études

6 Recherche terres cultivables

7 Raison de santé

8 Retour chez soi

9 Autres

s4q8 : Label : **Recherche d'emploi**

Type: Numérique ordinal

Taille: 1

Valeur par défaut : *Néant*

Valeurs et représentations :

1 Oui

2 Non

9 ND

I.7. Ep2fi5a : Fichier Chômage et Emploi principal

Ep2fi5a précise si l'individu est occupé ou au chômage. Il donne aussi le statut d'occupation (entre autre employeur, apprentis, etc.) et les critères de l'emploi principal que l'individu occupe. Il précise aussi le montant du revenu généré par l'emploi indiqué.

◆ Chemin

.. \DATA\...\ Ep2fi5a

◆ Nombre d'observations

Nombre d'observations : 42136

◆ Description des variables

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s1q1 : Idem que dans **Ep2fi1**

s51q2 : Label : **Statut d'occupation**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Occupé
- 2 Inoccupé déjà travaillé.
- 3 Inoccupé jamais travaillé.
- 4 Elève/étudiant
- 5 Femme au foyer
- 6 Retraité/rentier
- 7 Autres inactifs
- 9 ND

s51q3 : Label : **Avoir exercé une activité les sept derniers jours**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Oui
- 2 Non
- 9 ND

s51q4 : Label : **Avoir cherché du travail**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Oui
- 2 Non
- 9 ND

s51q5 : Label : **Raison**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Pas de qualification
- 2 Pas d'emploi
- 3 Malade
- 4 Attend une réponse
- 5 Autres
- 9 ND

s51q6 : Label : **Raison**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Nouvelle insertion
- 2 Licencié raison économique
- 3 Liquidation entreprise
- 4 Restructuration
- 5 Privatisation
- 6 Départ volontaire
- 7 Autres

s51q7 : Label : **Prêt à prendre ce travail**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Oui
- 2 Non
- 9 ND

s51q8 : Label : **Emploi occupé le plus**

Type : Numérique

Taille : 3

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s51q9 : Label : **Branche d'activité**

Type : Numérique ordinal
Taille : 2
Valeur par défaut : *Néant*
Valeurs et représentations : *Néant*

s51q10 : Label : Catégorie socioprofessionnelle

Type : Numérique ordinal
Taille : 2
Valeur par défaut : *Néant*
Valeurs et représentations :
1 Cadre direct. patron
2 Cadre supérieur
3 Cadre moyen
4 Ouvrier qualifié
5 Ouvrier semi-qualifié
6 Manœuvre
7 Apprenti
8 Aide-familial
9 Indépendant
10 Autres

s51q11 : Label : Situation dans la profession

Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs et représentations :
1 Indépendant
2 Salarié secteur public
3 Salarié secteur privé
4 Employeur
5 Aide familial
6 Apprenti/stagiaire
7 Bénévole

s51q12 Label : Nature de l'emploi

Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs et représentations :
1 Permanent
2 Saisonnier
3 Journalier
4 Autre temporaire
9 ND

s52q13 : Label : Droit à une pension

Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Oui
- 2 Non
- 9 ND

s52q14 : Label : **Type de sécurité**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 CARFO
- 2 CNSS
- 3 Assurances
- 4 CARFO+Assurances
- 5 CNSS+Assurances
- 9 ND

s52q15 : Label : **Droit à des congés payés**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Oui
- 2 Non

s52q16 : Label : **Montant du revenu**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s52q16u : Label : **Unité de temps**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Jour
- 2 Semaine
- 3 Mois
- 4 Année
- 9 ND

s52q17 : Label : **Nombre de mois**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s52q18 : Label : Principal statut d'occupation

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Occupé
- 2 Inoccupé travaillé
- 3 Inoccupé jamais trav.
- 4 Elève/étudiant
- 5 Femme au foyer
- 6 Retraité/rentier
- 7 Autre inactif
- 9 ND

s52q19 : Label : Situation dans la profession

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Indépendant
- 2 Salarié sect. public
- 3 Salarié sect. privé
- 4 Employeur
- 5 Aide familial
- 6 Apprenti/stagiaire
- 7 Bénévole
- 9 ND

s52q20 : Label : Emploi

Type : Numérique

Taille : 3

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s52q21 : Label : Branche d'activité

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations :

s52q22 : Label : Nombre de mois

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s52q23 : Label : Montant du revenu

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s52q24 : Label : **Evolution du revenu monétaire**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 4 Non concerné
- 9 ND

I.8. Ep2fi5b : Fichier Emploi secondaire

Ce fichier décrit l'emploi secondaire occupé par l'individu, son statut dans cet emploi et le revenu engendré par cet emploi.

◆ **Chemin**

.. \\DATA\...\ Ep2fi5b

◆ **Nombre d'observations**

Nombre d'observations : 32520

◆ **Description des variables**

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s53q1 : Label : **Numéro d'ordre**

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s53q2 : Label : **Autres activités**

Type : Numérique

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s53q3 : Label : **Emploi**

Type : Numérique

Taille : 3

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s53q4 : Label : **Branche d'activité**

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s53q5 : Label : **Nature de l'emploi**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

1 Permanent

2 Saisonnier

- 3 Journalier
- 4 Autre temporaire
- 9 ND

s53q6 : Label : Situation dans la profession

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Indépendant
- 2 Salarié
- 3 Employeur
- 4 Aide familial
- 5 Apprenti/stagiaire
- 6 Bénévole
- 9 ND

s53q7 : Label : Montant du revenu

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s53q7u : Label : Unité de temps

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Jour
- 2 Semaine
- 3 Mois
- 4 Année
- 9 ND

s53q8 : Label : Nombre de mois

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s53q9 : Label : Montant du revenu

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s53q10 : Label : Evolution du revenu

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 4 Non concerné
- 9 ND

I.9. Ep2fi5c : Fichier Emploi précédent

Il s'agit du dernier emploi occupé avant l'emploi principal. Ce fichier renferme une description de cet emploi et du statut de l'individu dans cet emploi.

◆ **Chemin**

.. \DATA\...\Ep2fi5c

◆ **Nombre d'observations**

Nombre d'observations : 36468

◆ **Description des variables**

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s54q1 : Label : **Numéro d'ordre**

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s54q2 : Label : **Autres emplois**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Exer. act, un autre
- 2 N exerc. pas, exerça
- 3 Toujours le même
- 4 N'a jamais exercé
- 9 ND

s54q3 : Label : **Dernier emploi**

Type : Numérique

Taille : 3

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s54q4 : Label : **Nombre d'année**

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s54q5 : Label : **Branche d'activité**

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s54q6 : Label : **Nature de l'emploi**

Type : Numérique Ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Permanent
- 2 Saisonnier
- 3 Journalier
- 4 Autre temporaire
- 9 ND

s54q7 : Label : **Situation dans la profession**

Type : Numérique Ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Indépendant
- 2 Salarié secteur public
- 3 Salarié secteur privé
- 4 Employeur
- 5 Aide familial
- 6 Apprenti/stagiaire
- 7 Bénévole
- 9 ND

s54q8 : Label : **Droit à une pension**

Type : Numérique Ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Oui
- 2 Non
- 9 ND

s54q9 : Label : **Droit à des congés payés**

Type : Numérique

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Oui
- 2 Non
- 9 ND

s54q10 : Label : **Raison du départ de l'emploi**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Salaire trop bas
- 2 Licenciement/compression
- 3 Fermeture de l'entreprise.
- 4 Démission
- 5 Promotion
- 6 Sanction
- 7 Environnement hostile
- 8 Raison de santé
- 9 Autre

I.10. Ep2fi6 : Fichier Logement et confort

Ep2fi6 renferme les conditions de logement et d'accès aux services de bases. Il décrit les types de bâtiments de logement, les moyens d'approvisionnement en eau et en énergie.

◆ **Chemin**

.. \DATA\ ... \ Ep2fi6

◆ **Nombre d'observations**

Nombre d'observations : 8478

◆ **Description des variables**

g5 : Idem que dans Ep2fiid

g8 : Idem que dans Ep2fiid

s6q1 : Label : **Ménage existait**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Oui
2 Non

s6q2 : Label : **Evolution nombre de personnes**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Augmenté
2 Diminué
3 Inchangé
9 ND

s6q3 : Label : **Zone de résidence**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Zone lotie
2 Zone non lotie

s6q4 : Label : **Nature des murs**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Immeuble à appartement.
- 2 Villa
- 3 Maison individuelle
- 4 Bâtiment à plusieurs logements
- 5 Maisons traditionnelles.
- 6 Autres

s6q5 : Label : **Matériau des murs extérieurs**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Béton, Pierres, parpaings
- 2 Briques cuites/compressés
- 3 Banco
- 4 Paille
- 5 Autre

s6q6 : Label : **Matériau de la toiture**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Béton
- 2 Tôle
- 3 Tuile
- 4 Terre battue
- 5 Paille
- 6 Autres

s6q7 : Label : **Matériel du sol**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Carreau
- 2 Ciment
- 3 Terre battue/sable
- 4 Autres

s6q8 Label : **Type de lieu d'aisance**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Chasse avec fosse

- 2 Latrines à fosses
- 3 Latrines ordinaires
- 4 Dans la nature
- 5 Autres

s6q9 : Label : **Type de sanitaire**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Baignoire
- 2 Douche avec robinet
- 3 Douche sans robinet
- 4 Autres

s6q10 : Label : **Evacuation des ordures**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Poubelle
- 2 Tas d'immondices public
- 3 Tas d'immondices individuel
- 4 Fosse
- 5 Bac
- 6 Rue

s6q11 : Label : **Nombre de pièces**

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s6q12 : Label : **Nombre de pièces**

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s6q13 : Label : **Statut d'occupant**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Propriétaire
- 2 Location simple
- 3 Location vente
- 4 Logé gratuitement
- 5 Autres

s6q14 : Label : **Statut d'occupant**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Propriétaire
- 2 Location simple
- 3 Location vente
- 4 Logé gratuitement
- 5 Autres

s6q15 : Label : **Mode d'approvisionnement en eau de boisson**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Rivière, cours d'eau
- 2 Puits ordinaires
- 3 Puits busés
- 4 Forage
- 5 Fontaine publique
- 6 Robinet intérieur propre
- 7 Robinet intérieur partagé
- 8 Autres

s6q16 : Label : **Mode d'approvisionnement en eau de boisson**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Rivière, cours d'eau
- 2 Puits ordinaires
- 3 Puits busés
- 4 Forage
- 5 Fontaine publique
- 6 Robinet intérieur propre
- 7 Robinet intérieur partagé
- 8 Autres

s6q17 : Label : **Energie pour éclairage**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Bois
- 2 Pétrole
- 3 Gaz
- 4 Électricité/solaire
- 5 Bougie
- 6 Charbon de bois

- 7 Torche à pile
- 8 Autre

s6q18 : Label : Energie pour éclairage

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Bois
- 2 Pétrole
- 3 Gaz
- 4 Electricité/solaire
- 5 Bougie
- 6 Charbon de bois
- 7 Torche à pile
- 8 Autre

s6q19 : Label : Energie pour la cuisine

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Bois
- 2 Pétrole
- 3 Gaz
- 4 Electricité/solaire
- 5 Bougie
- 6 Charbon de bois
- 7 Torche à pile
- 8 Autre

s6q20 : Label : Energie pour la cuisine

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Bois
- 2 Pétrole
- 3 Gaz
- 4 Electricité/solaire
- 5 Bougie
- 6 Charbon de bois
- 7 Torche à pile
- 8 Autre

s6q21 : Label : Existence pièce aménagée pour cuisine

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

1 Oui

2 Non

s6q22 : Label : **Possède le téléphone**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

1 Oui

2 Non

s6q23 : Label : **Possède le téléphone**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

1 Oui

2 Non

I.11. Ep2fi7a : Fichier Bétail

Ce fichier renferme l'ensemble du bétail possédé par le ménage. Il s'agit du bétail appartenant à l'ensemble du ménage y compris les résidents absents.

◆ **Chemin**

.. \DATA\...\Ep2fi7a

◆ **Nombre d'observations**

Nombre d'observations : 38 168

◆ **Description des variables**

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s71q1 : Label : *Néant*

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

1 Oui

2 Non

s71q2 : Label : **Nombre de têtes**

Type : Numérique

Taille : 3

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

s71q3 : Label : **Evolution**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

1 Augmenté

2 Inchangé

3 Diminué

4 N'existait pas

9 ND

s71q4 : Label : **Raison principale**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Vente
- 2 Mortalité
- 3 Cérémonie
- 4 Autre
- 9 ND

s71q0 : Label : *Néant*

Type : Numérique ordinal

Taille : 8

Valeur par défaut : *Néant*

Valeurs et représentations : *Néant*

I.12. Ep2fi7b : Fichier Production campagne 98/99

Ep2fi7b renferme les informations concernant la production agricole 98/99 des ménages. Il renvoi les moyens mis en œuvre, les résultats obtenus, le mode d'écoulement des produits.

◆ **Chemin**

.. \DATA\ ... \Ep2fi7b

◆ **Nombre d'observations**

Nombre d'observations : 75798

◆ **Description des variables**

g5 : Idem que dans Ep2fiid

g8 : Idem que dans Ep2fiid

s72q1: Label : **Le ménage a produit**

Type: Numérique ordinal

Taille: 1

Valeur par défaut: Néant

Valeurs représentées :

1 Oui

2 Non

9 ND

s72q2: Label : **Production 97/98**

Type: Numérique

Taille: 5

Valeur par défaut: Néant

Valeurs représentées : Néant

s72q2u: Label : **Unité d'observation**

Type: Numérique ordinal

Taille: 2

Valeur par défaut: Néant

Valeurs représentées : Néant

s72qf: Label : **Forme du produit**

Type: Numérique ordinal

Taille: 1

Valeur par défaut: Néant

Valeurs représentées :

1 Grain

2 Epis

3 Coque

9 ND

s72q3: Label : Le ménage a-t-il vendu

Type: Numérique ordinal

Taille: 1

Valeur par défaut: Néant

Valeurs représentées :

1 Oui

2 Non

9 ND

s72q4: Label : Quantité vendue

Type: Numérique ordinal

Taille: 5

Valeur par défaut: Néant

Valeurs représentées : Néant

s72q4u: Label : Unité d'observation

Type: Numérique

Taille: 2

Valeur par défaut: Néant

Valeurs représentées : Néant

s72q4f: Label : Forme du produit

Type: Numérique ordinal

Taille:1

Valeur par défaut: Néant

Valeurs représentées :

1 Grain

2 Epis

3 Coque

9 ND

s72q5: Label : Principal débouché

Type: Numérique ordinal

Taille: 1

Valeur par défaut: Néant

Valeurs représentées :

1 Marché

2 Négociant

3 Coopérant

4 Off. commerce

5 Autre

9 ND

s72q6: Label : Utilisation de semences sélectionnées

Type: Numérique ordinal

Taille: 1

Valeur par défaut: Néant

Valeurs représentées :

- 1 Oui
- 2 Non
- 9 ND

s72q7: Label : **Bénéfice de service d'encadrement.**

Type: Numérique ordinal

Taille: 1

Valeur par défaut: Néant

Valeurs représentées :

- 1 Oui
- 2 Non
- 9 ND

s72q8: Label : **Utilisation d'engrais**

Type: Numérique ordinal

Taille: 1

Valeur par défaut: Néant

Valeurs représentées :

- 1 Oui
- 2 Non
- 9 ND

s72q9 : Label : **Raison**

Type: Numérique ordinal

Taille: 1

Valeur par défaut: Néant

Valeurs représentées :

- 1 Trop cher
- 2 Non disponible
- 3 Pas nécessaire
- 4 Pas de crédit engrais
- 5 Ne sais pas utiliser
- 6 Autre
- 9 ND

s72q10: Label : **Crédit agricole sollicité**

Type: Numérique ordinal

Taille: 1

Valeur par défaut: Néant

Valeurs représentées :

- 1 Oui
- 2 Non
- 9 ND

s72q11: Label : **Obtention du crédit agricole**

Type: Numérique ordinal

Taille: 1
Valeur par défaut: Néant
Valeurs représentées :
1 Oui
2 Non
9 ND

s72q12: Label : **Nombre de personnes employées**

Type: Numérique
Taille: 2
Valeur par défaut: Néant
Valeurs représentées : Néant

s72q13: Label : **Type de matériel agricole**

Type: Numérique ordinal
Taille: 1
Valeur par défaut: Néant
Valeurs représentées :
1 Traditionnel
2 Charrue
3 Tracteur
4 Autres
9 ND

s72q0: Label : *Néant*

Type: Numérique ordinal
Taille: 8
Valeur par défaut: Néant
Valeurs représentées : Néant

I.13. Ep2fi7c : Fichier Production campagne 98/99

Ep2fi7c contient des informations complémentaires sur la production agricole 98/99. Il s'agit de déterminer les productions par type et par quantités, ainsi que les facteurs ayant influencés la production.

◆ **Chemin**

.. \\DATA\...\Ep2fi7c

◆ **Nombre d'observations**

Nombre d'observations : 75438

◆ **Description des variables**

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s732q1 : Label : **Le ménage a-t-il produit**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

1 Oui

2 Non

9 ND

s732q2 : Label : **Quantité de la production**

Type : Numérique

Taille : 5

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s732q2u : Label : **Unité d'observation**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s732q2f : Label : **Forme du produit**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

1 Grain

2 Epis

3 Coque

9 ND

s732q3 : Label : **Nombre d'employés**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s732q4 : Label : **Evolution de la superficie**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

1 Augmenté

2 Inchangé

3 Diminué

8 Pas de culture 96/97

9 ND

s732q5 : Label : **Evolution de la superficie**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

1 Augmenté

2 Inchangé

3 Diminué

9 ND

s732q61 : Label : **1^{ère} raison**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées :

1 Baisse de crédit

2 Moyen d'écoulement.

3 Manque main d'oeuvre

4 Pas d'engrais

5 Baisse prix de vente

6 Conditions climatiques.

7 Déprédations

8 Manque de semences

9 Manque d'encadrement

10 Autre

s732q62 : Label : **2^{ème} raison**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Baisse de crédit
- 2 Moyen d'écoulement.
- 3 Manque main d'œuvre
- 4 Pas d'engrais
- 5 Baisse prix de vente
- 6 Conditions climatiques.
- 7 Déprédations
- 8 Manque de semences
- 9 Manque d'encadrement
- 10 Autre

s732q63 : Label : 3^{ème} raison

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Baisse de crédit
- 2 Moyen d'écoulement.
- 3 Manque main d'œuvre
- 4 Pas d'engrais
- 5 Baisse prix de vente
- 6 Conditions climatiques.
- 7 Déprédations
- 8 Manque de semences
- 9 Manque d'encadrement
- 10 Autre

s732q0 : Label : *Néant*

Type : Numérique ordinal

Taille : 8

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

I.14. Ep2fi8 : Fichier Entreprise non agricole

Ce fichier renferme les informations relatives aux principales entreprises non agricoles du ménage et leurs facteurs de fonctionnements.

◆ **Chemin**

.. \DATA\ ... \ Ep2fi8

◆ **Nombre d'observations**

Nombre d'observations : 4876

◆ **Description des variables**

g5 : Idem que dans Ep2fiid

g8 : Idem que dans Ep2fiid

s8q1 : Label : **Activité de l'entreprise**
Type : Numérique
Taille : 2
Valeur par défaut : *Néant*
Valeurs représentées : *Néant*

s8q2 : Label : **Membre du ménage**
Type : Numérique
Taille : 2
Valeur par défaut : *Néant*
Valeurs représentées : *Néant*

s8q3 : Label : **Durée d'existence de l'entreprise.**
Type : Numérique
Taille : 2
Valeur par défaut : *Néant*
Valeurs représentées : *Néant*

s8q4 : Label : **Nombre de mois**
Type : Numérique ordinal
Taille : 2
Valeur par défaut : *Néant*
Valeurs représentées : *Néant*

s8q5 : Label : **Actuellement en activité**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Oui

2 Non

s8q6 : Label : **Nombre de personnes**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s8q7 : Label : **Nombre de personnes**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s8q8 : Label : **Dépense totale**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s8q9 : Label : **Chiffre d'affaire**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s8q10 : Label : **Impôts et taxes**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s8q11 : Label : **Evolution du revenu**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

1 Amélioré

2 Inchangé

3 Dégradé

8 N'existait pas

s8q12 : Label : **Raison principale**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

1 Accès au crédit

2 Manque d'équipement

- 3 Manque de main d'œuvre
- 4 Hausse des prix de matière première
- 5 Difficultés d'approvisionnement.
- 6 Clientèle
- 7 Taxes
- 8 Autres

s8q0 : Label : Néant
Type : Numérique ordinal
Taille : 8
Valeur par défaut : *Néant*
Valeurs représentées : *Néant*

I.15. Ép2fi9 : Fichier Accès aux services de base les plus proche

Ce fichier renferme les informations relatives à l'accès aux services de base les plus proches. Il s'agit de l'école primaire, l'école secondaire, le service de santé publique et le marché le plus proche. Ces informations correspondent à une série de quatre questions posées pour chaque service d'où la redondance des labels.

◆ Chemin

.. \DATA\ ... \ Ep2fi9

◆ Nombre d'observations

Nombre d'observations : 8478

◆ Description des variables

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s90q1\$01 : Label : **Temps nécessaire**
 Type : Numérique ordinal
 Taille : 1
 Valeur par défaut : Néant
 Valeurs représentées :

- 1 Disponibilité à la maison
- 2 Moins d 1/2 heures
- 3 1/2 à 1 heure
- 4 1 heure à 2 heures
- 5 2 h à 1/2 journée
- 6 1/2 j à 1 j
- 7 Plus d'une journée
- 8 N'a pas accès
- 9 ND

s90q1\$02 Label : **Temps nécessaire**
 Type : Numérique ordinal
 Taille : 1
 Valeur par défaut : *Néant*
 Valeurs représentées :

- 1 Disponibilité à la maison
- 2 Moins d 1/2 heure
- 3 1/2 à 1 heure
- 4 1 heure à 2 heures
- 5 2 h à 1/2 journée
- 6 1/2 j à 1 j
- 7 Plus d'une journée
- 8 N'a pas accès

9 ND

s90q1\$03

Label : **Temps nécessaire**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Disponibilité à la maison
- 2 Moins d 1/2 heure
- 3 1/2 à 1 heure
- 4 1 heure à 2 heures
- 5 2 h à 1/2 journée
- 6 1/2 j à 1 j
- 7 Plus d'une journée
- 8 N'a pas accès
- 9 ND

s90q1\$04

Label : **Temps nécessaire**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Disponibilité à la maison
- 2 Moins d 1/2 heure
- 3 1/2 à 1 heure
- 4 1 heure à 2 heures
- 5 2 h à 1/2 journée
- 6 1/2 j à 1 j
- 7 Plus d'une journée
- 8 N'a pas accès
- 9 ND

s90q1\$05

Label : **Temps nécessaire**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

- 1 Disponibilité à la maison
- 2 Moins d 1/2 heure
- 3 1/2 à 1 heure
- 4 1 heure à 2 heures
- 5 2 h à 1/2 journée
- 6 1/2 j à 1 j
- 7 Plus d'une journée
- 8 N'a pas accès
- 9 ND

s90q2\$01 : Label : **Moyen le plus utilisé**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 A pied
2 Bicyclette
3 Motocycle
4 Automobile
5 Autre
9 ND

s90q2\$02 : Label : **Moyen le plus utilisé**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 A pied
2 Bicyclette
3 Motocycle
4 Automobile
5 Autre
9 ND

s90q2\$03 : Label : **Moyen le plus utilisé**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 A pied
2 Bicyclette
3 Motocycle
4 Automobile
5 Autre
9 ND

s90q2\$04 : Label : **Moyen le plus utilisé**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 A pied
2 Bicyclette
3 Motocycle
4 Automobile
5 Autre
9 ND

s90q2\$05 : Label : **Moyen le plus utilisé**
Type : Numérique ordinal

Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 A pied
2 Bicyclette
3 Motocycle
4 Automobile
5 Autre
9 ND

s90q3\$01 : Label : **Des membres utilisent ce service**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Oui
2 Non
9 ND

s90q3\$02 : Label : **Des membres utilisent ce service**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Oui
2 Non
9 ND

s90q3\$03 : Label : **Des membres utilisent ce service**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Oui
2 Non
9 ND

s90q3\$04 : Label : **Des membres utilisent ce service**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Oui
2 Non
9 ND

s90q3\$05 : Label : **Des membres utilisent ce service**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Oui
2 Non
9 ND

s90q4\$01 : Label : **Raison**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Trop cher
2 Trop loin
3 Mauvaise qualité
4 Autres
5 Non concerné

s90q4\$02 : Label : **Raison**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Trop cher
2 Trop loin
3 Mauvaise qualité
4 Autres
5 Non concerné

s90q4\$03 : Label : **Raison**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Trop cher
2 Trop loin
3 Mauvaise qualité
4 Autres
5 Non concerné

s90q4\$04 : Label : **Raison**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :
1 Trop cher
2 Trop loin

- 3 Mauvaise qualité
- 4 Autres
- 5 Non concerné

s90q4\$05 : Label : **Raison**
Type : Numérique ordinal
Taille : 1
Valeur par défaut : *Néant*
Valeurs représentées :

- 1 Trop cher
- 2 Trop loin
- 3 Mauvaise qualité
- 4 Autres
- 5 Non concerné

I.16. Ep2fi10a : Fichier Dépense scolaire de l'année précédente

Ep2fi10a renferme l'ensemble des dépenses, frais scolaires, frais transports, frais médicaux et autres frais entrant dans le cadre de la vie scolaire.

◆ **Chemin**

.. \DATA\...\Ep2fi10a

◆ **Nombre d'observations**

Nombre d'observations : 8478

◆ **Description des variables**

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s91q11 : Label : **Dépenses frais de scolarité**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s91q21 : Label : **Dépenses livres et fournitures**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s91q31 : Label : **Frais de transport scolaire**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s91q41 : Label : **Cotisation des parents d'élève**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s91q51 : Label : **Autres contributions scolaires**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s91q12 : Label : Evolution de cette dépense

Type : Numérique ordinal

Taille : 1

Valeur par défaut : Néant

Valeurs représentées :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 8 Le ménage n'existait pas
- 9 ND

s91q22 : Label : Evolution de cette dépense

Type : Numérique ordinal

Taille : 1

Valeur par défaut : Néant

Valeurs représentées :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 8 Le ménage n'existait pas
- 9 ND

s91q32 : Label : Evolution de cette dépense

Type : Numérique ordinal

Taille : 1

Valeur par défaut : Néant

Valeurs représentées :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 8 Le ménage n'existait pas
- 9 ND

s91q42 : Label : Evolution de cette dépense

Type : Numérique ordinal

Taille : 1

Valeur par défaut : Néant

Valeurs représentées :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 8 Le ménage n'existait pas
- 9 ND

s91q52 : Label : Evolution de cette dépense

Type : Numérique ordinal

Taille : 1
Valeur par défaut : Néant
Valeurs représentées :
1 Augmenté
2 Inchangé
3 Diminué
8 Le ménage n'existait pas
9 ND

s91q61 : Label : **Frais de consultation**

Type : Numérique
Taille : 9
Valeur par défaut : Néant
Valeurs représentées : *Néant*

s91q71 : Label : **Frais d'analyses médicales**

Type : Numérique
Taille : 9
Valeur par défaut : Néant
Valeurs représentées : *Néant*

s91q81 : Label : **Dépenses en médicaments**

Type : Numérique
Taille : 9
Valeur par défaut : Néant
Valeurs représentées : *Néant*

s91q91 : Label : **Frais d'hospitalisation**

Type : Numérique
Taille : 9
Valeur par défaut : Néant
Valeurs représentées : *Néant*

s91q101 : Label : **Autres services médicaux**

Type : Numérique
Taille : 9
Valeur par défaut : Néant
Valeurs représentées : *Néant*

s91q62 : Label : **Evolution de cette dépense**

Type : Numérique ordinal
Taille : 1
Valeur par défaut : Néant
Valeurs représentées :
1 Augmenté
2 Inchangé
3 Diminué
8 Le ménage n'existait pas
9 ND

s91q72 : Label : **Evolution de cette dépense**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : Néant

Valeurs représentées :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 8 Le ménage n'existait pas
- 9 ND

s91q82 : Label : **Evolution de cette dépense**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : Néant

Valeurs représentées :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 8 Le ménage n'existait pas
- 9 ND

s91q92 : Label : **Evolution de cette dépense**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : Néant

Valeurs représentées :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 8 Le ménage n'existait pas
- 9 ND

s91q102 : Label : **Evolution de cette dépense**

Type : Numérique ordinal

Taille : 1

Valeur par défaut : Néant

Valeurs représentées :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 8 Le ménage n'existait pas
- 9 ND

I.17. Ep2fi10b : Fichier Dépenses alimentaires des 30 derniers jours

Ce fichier regroupe toutes les dépenses effectuées au cours des 30 derniers jours pour la consommation alimentaire. Il renferme également la valeur des cadeaux et des produits auto consommés.

♦ **Chemin**

.. \\DATA\...\ Ep2fi10b

♦ **Nombre d'observations**

Nombre d'observations : 288 252

♦ **Description des variables**

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s92c0 : Label : *Néant*

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s92c1 : Label : **Le ménage a consommé ce produit.**

Type : Numérique Ordinal

Taille : 1

Valeur par défaut : *Néant*

Valeurs représentées :

1 Oui

2 Non

9 ND

s92c2 : Label : *Néant*

Type : Numérique

Taille : 8

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s92c3 : Label : **Valeur des cadeaux reçus et consommés.**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

s92c4 : Label : **Valeur des produits auto-consommés.**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs représentées : *Néant*

I.18. Ep2fi10c : Fichier Dépenses sur les produits non alimentaires

Ep2fi10c renferme le montant et les variations des dépenses non alimentaires effectuées par le ménage.

◆ **Chemin**

.. \\DATA\...\ Ep2fi10c

◆ **Nombre d'observations**

Nombre d'observations : 262 818

◆ **Description des variables**

g5 : Idem que dans Ep2fiid

g8 : Idem que dans Ep2fiid

s93c0 : Label : *Néant*

Type : Numérique

Taille : 2

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

s93c1 : Label : **Le ménage a dépensé**

Type : Numérique ordinal

Taille : 1

Valeur par défaut: *Néant*

Valeurs représentées :

1 Oui

2 Non

9 ND

s93c2 : Label : *Néant*

Type : Numérique

Taille : 8

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

s93c3 : Label : **Evolution de la dépense**

Type : Numérique

Taille : 1

Valeur par défaut: *Néant*

Valeurs représentées :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 4 N'existait pas
- 9 ND

s93c4 : Label : **Valeur d'autoconsommation**

Type : Numérique

Taille : 9

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

I.19. Ep2fi11 : Fichier Revenu des ménages

Ep2fi11 donne les informations liées au type et au montant de revenus des ménages. Il indique aussi les variations perçus dans le montant de ce revenu.

◆ Chemin

.. \DATA\...\ Ep2fi11

◆ Nombre d'observations

Nombre d'observations : 296 730

◆ Description des variables

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s94c0 : Label : *Néant*

Type : Numérique

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Arachide et sous produits
- 2 Coton et sous produits
- 3 Mil/sorgho et sous produits
- 4 Riz et sous produits
- 5 Maïs et sous produits
- 6 Niébé et sous produits
- 7 Fonio
- 8 Sésame
- 9 Igname, patate
- 10 Autres cultures
- 11 Cultures maraîchères
- 12 Cultures fruitières
- 13 Karité (fruit et amende)
- 14 Néré et sous produits
- 15 Autres produits de la cueillette
- 16 Vente de bovins
- 17 de caprins
- 18 d'ovins
- 19 de porcins
- 20 Asins
- 21 Volailles
- 22 Autres animaux
- 23 Produits d'animaux
- 24 Salaires du public et parapublic
- 25 Salaire du privé moderne

- 26 Autres salaires du secteur privé
- 27 Revenus des activités non agricoles
- 28 Loyers perçus, rentes, dividendes
- 29 Transferts reçus villes Burkina
- 30 Transferts reçus Reste du Burkina
- 31 Transferts reçus de Côte d'Ivoire
- 32 Transferts reçus de France
- 33 Transferts reçus autres pays
- 34 Retraite, allocation familiales,...
- 35 Autres sources

s94c1 : Label : **Le ménage a tiré un revenu**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Oui
- 2 Non
- 9 ND

s94c2 : Label : **Montant du revenu**

Type : Numérique

Taille : 9

Valeur par défaut : *Néant*

Valeurs et représentations : Néant

s94c3 : Label : **Evolution du revenu**

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs et représentations :

- 1 Augmenté
- 2 Inchangé
- 3 Diminué
- 8 N'existait pas

I.20. Ep2fi12 : Fichier Avoir du ménage

Ce fichier renvoi les informations relatives au type de biens possédé par le ménage.

◆ Chemin

.. \DATA\...\ Ep2fi12av

◆ Nombre d'observations

Nombre d'observations : 194 994

◆ Description des variables

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s95c0 : Label : *Néant*

Type : Numérique ordinal

Taille : 2

Valeur par défaut : *Néant*

Valeurs représentées : Néant

s95c1 : Label : **Le ménage possède ce bien**

Type : Numérique ordinal

Taille :

Valeur par défaut : *Néant*

Valeurs représentées :

1 Oui

2 Non

9 ND

s95c2 : Label : **Nombre d'années**

Type : Numérique ordinal

Taille :

Valeur par défaut : *Néant*

Valeurs représentées :

0 Moins d'un an

99 Ne sais pas

s95c3 : Label : **Le ménage possédait ce bien**

Type : Numérique ordinal

Taille :

Valeur par défaut : *Néant*

Valeurs représentées :

1 Oui

- 2 Non
- 8 Non concerné
- 9 ND

I.21. Ep2fi13 : Fichier Anthropométrie

Elle concerne les enfants âgés de 6 mois à 59 mois. Cette section, qui fait l'objet de mesures directes de taille et de poids, ne concerne qu'un sous-échantillon de ménages.

◆ Chemin

.. \DATA\...\Ep2fi13

◆ Nombre d'observations

Nombre d'observations : 4712

◆ Description des variables

g5 : Idem que dans **Ep2fiid**

g8 : Idem que dans **Ep2fiid**

s96c1 : Label : **Numéro d'ordre**

Type : Numérique

Taille :2

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

s96c2 : Label : **Numéro d'ordre mère**

Type : Numérique

Taille :2

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

s96c3 : Label : **Age**

Type : Numérique

Taille :2

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

s96c4 : Label : **Numéro d'ordre père**

Type : Numérique

Taille :2

Valeur par défaut: *Néant*

Valeurs représentées : *Néant*

s96c5 : Label : **Poids**

Type : Numérique ordinal
Taille :2
Valeur par défaut: *Néant*
Valeurs représentées : *Néant*

s96c52 : Label : *Néant*
Type : Numérique ordinal
Taille :1
Valeur par défaut: *Néant*
Valeurs représentées : *Néant*

s96c6 : Label : **Taille(cm)**
Type : Numérique
Taille :3
Valeur par défaut: *Néant*
Valeurs représentées : *Néant*

s96c7 : Label : **L'enfant a fait une diarrhée**
Type : Numérique ordinal
Taille :1
Valeur par défaut: *Néant*
Valeurs représentées :
1 Oui
2 Non
9 ND

II. Remarques et suggestions

- L'organisation de la base sous forme de fichiers correspondant à chaque section du profil de pauvreté améliore la lisibilité des données ainsi que les analyses statistiques.
- La disponibilité d'un rapport d'analyse sur l'EPII avec les différents indicateurs de niveau de vie des ménages permet de mieux comprendre les données disponibles et les critères d'organisation des données.
- D'autre part, l'absence de documents relatifs aux méthodes de calculs des différents indicateurs ne facilite pas les vérifications des résultats établies.
- Dans l'optique de permettre une bonne exploitation des bases de données, il serait très intéressant de disposer pour chaque enquête aussi bien, les résultats que les méthodes mis en œuvres pour les obtenir.
- De plus, il faut veiller à ce que les données de la base ne soient altérées par les manipulations. Les manipulations ne devront s'effectuer que sur une copie de la base brute.

CHAPITRE 2

ENQUETE CORRUPTION

SOMMAIRE

| | |
|---|-----------|
| I. Description des variables | 75 |
| I.1. Identification du ménage | 75 |
| I.2. Caractéristiques socio-économiques et démographiques du répondant..... | 75 |
| I.3. Définition de la corruption | 76 |
| I.4. Perception de la corruption | 76 |
| I.5. Corruption et genre | 77 |
| I.6. Les formes et pratiques en matière de corruption | 77 |
| <i>I.6.1. Les secteurs d'activités et la corruption.....</i> | <i>78</i> |
| <i>I.6.2 Corruption et gouvernance économique.....</i> | <i>79</i> |
| <i>I.6.3 Corruption et gouvernance démocratique.....</i> | <i>80</i> |
| I.7. Les causes et effets de la corruption..... | 81 |
| I.8. Institution de lutte contre la corruption | 83 |
| II. Remarques et suggestions | 84 |

I. Description des variables

La base de donnée de l'enquête corruption est réalisée par le groupe national de réflexion sur le développement humain durable. Elle est structurée en un seul fichier et Elle renferme 525 observations.

Pour la description de la base, les variables seront regroupées en section.

I.1. Identification du ménage

Cette section détermine l'identifiant du ménage ainsi que sa situation géographique.

Numque : Label : **Numéro du questionnaire rempli**

Type : Caractère nominal

Taille :6

Valeur par défaut : *Néant*

Prov : Label : **Province**

Type : Caractère nominal

Taille : 2

Valeur par défaut : *Néant*

Dept : Label : **Département**

Type : Caractère nominal

Taille :2

Valeur par défaut : *Néant*

Secteur : Label : **Secteur/Village**

Type : Caractère nominal

Taille :6

Valeur par défaut : *Néant*

Ident : Label : **Identifiant**

Type : Caractère nominal

Taille :10

Valeur par défaut : *Néant*

I.2. Caractéristiques socio-économiques et démographiques du répondant.

Dans cette section, nous avons une description des caractéristiques du répondant, son niveau d'instruction et sa situation professionnelle et économique.

Sexe : Label : **Sexe**

Type : Caractère nominal

Taille : 5

Valeur par défaut : *Néant*

Age : Label : **Age**

Type : Caractère nominal
 Taille :2
 Valeur par défaut : *Néant*

Instruc : Label : **Niveau d'instruction du répondant**

Type : Caractère nominal
 Taille :5
 Valeur par défaut : *Néant*

Catprof : Label : **Catégorie socioprofessionnelle**

Type : Caractère nominal
 Taille :4
 Valeur par défaut : *Néant*

Sitprof : Label : **Situation dans la profession dans l'activité principale**

Type : Caractère nominal
 Taille :4
 Valeur par défaut : *Néant*

I.3. Définition de la corruption

Il s'agit de recueillir la définition que le répondant donne de la corruption. Est-ce l'utilisation des ressources publiques à des fins personnelles, un arrangement, une altération ou une dépravation des valeurs au niveau d'un individu ou autre chose ?

Cordefi : Label : **Définition de la corruption**

Type : Caractère nominal
 Taille :4
 Valeur par défaut : *Néant*

I.4. Perception de la corruption

Il s'agit de la perception que le répondant a de la corruption : son ampleur, sa gravité, la situation actuelle par rapport au passé.

Corampl : Label : **Ampleur de la corruption**

Type : Caractère nominal
 Taille :4
 Valeur par défaut : *Néant*

Corprofo : Label : **Profondeur de la corruption**

Type : Caractère nominal
 Taille :4
 Valeur par défaut : *Néant*

Corrgl : Label : **Corruption = règle de conduite**

Type : Caractère nominal
 Taille :4
 Valeur par défaut : *Néant*

Corevo : Label : **Corruption avant**
Type : Caractère nominal
Taille :4
Valeur par défaut : *Néant*

Coregi : Label : **Corruption dans la sous-region**
Type : Caractère nominal
Taille :4
Valeur par défaut : *Néant*

I.5. Corruption et genre

Cette section établit le lien entre la corruption et le genre en précisant le genre pour lequel elle est plus développée.

Corgenr : Label : **Corruption et genre**
Type : Caractère nominal
Taille :5
Valeur par défaut : *Néant*

Corhom : Label : **Corruption chez les hommes**
Type : Caractère nominal
Taille :5
Valeur par défaut : *Néant*

Corfem : Label : **Corruption chez les femmes**
Type : Caractère nominal
Taille :5
Valeur par défaut : *Néant*

I.6. Les formes et pratiques en matière de corruption

Dans cette section, il s'agit de préciser les formes et les pratiques de corruption la plus répandue.

Povin : Label : **Corruption par pots de vin**
Type : Caractère nominal
Taille :5
Valeur par défaut : *Néant*

Influen : Label : **Corruption par trafic d'influence**
Type : Caractère nominal
Taille :5
Valeur par défaut : *Néant*

Donarg : Label : **Corruption par don d'argent**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Defond : Label : **Corruption par détournement de fonds**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Recrut : Label : **Corruption par tricherie dans le recrutement du personnel**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Escrog : Label : **Corruption par escroquerie**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Devenq : Label : **Corruption par dévoiement des enquêtes internes**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

1.6.1. Les secteurs d'activités et la corruption.

Cette section permet de classer la corruption par secteurs d'activités.

Corsant : Label : **Corruption dans le secteur de santé**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Coredu : Label : **Corruption dans le secteur de l'éducation**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Cordoua : Label : **Corruption au niveau de la douane**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Coradm : Label : **Corruption dans l'administration publique**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Corpolic : Label : **Corruption au niveau de la police et de la gendarmerie**

Type : Caractère nominal

Taille :5
Valeur par défaut : *Néant*

Cormpu : Label : **Corruption dans les marchés publics**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Cormed : Label : **Corruption au niveau des médias**

Type : Caractère nominal

Taille :4

Valeur par défaut : *Néant*

Corjus : Label : **Corruption de la justice**

Type : Caractère nominal

Taille :4

Valeur par défaut : *Néant*

Cormair : Label : **Corruption dans les mairies**

Type : Caractère nominal

Taille :4

Valeur par défaut : *Néant*

Corimp : Label : **Corruption au niveau des impôts**

Type : Caractère nominal

Taille :4

Valeur par défaut : *Néant*

Corpoli : Label : **Corruption dans l'activité politique**

Type : Caractère nominal

Taille :4

Valeur par défaut : *Néant*

1.6.2 Corruption et gouvernance économique

Ici, il faut établir le lien entre la corruption et la gouvernance économique. La privatisation ou le programme d'ajustement structurel (PAS) sont-ils des facteurs aggravants ?

Corpriv : Label : **Corruption et privatisation**

Type : Caractère nominal

Taille :4

Valeur par défaut : *Néant*

Corpas : Label : **Corruption et Programme d'ajustement structurel (PAS)**

Type : Caractère nominal

Taille :4

Valeur par défaut : *Néant*

1.6.3 Corruption et gouvernance démocratique

Cette section définit le rôle et l'impact de la gouvernance démocratique dans la lutte contre la corruption.

Corgov1 : Label : **Lutte contre la corruption par le gouvernement**

Type : Caractère nominal

Taille :4

Valeur par défaut : *Néant*

Corgov2 : Label : **Encouragement de la corruption par le gouvernement**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Coran : Label : **Lutte contre la corruption par l'assemblée nationale**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Corelec : Label : **Corruption et élections**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Corhpol : Label : **Faire la politique et corruption**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Cordepu : Label : **Député et corruption**

Type : Caractère nominal

Taille :5

Valeur par défaut

Cormaire : Label : **Maire et corruption**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Cormin : Label : **Ministre et corruption**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Corcau : Label : **Causes de la corruption**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Corcons : Label : **Corruption et ses conséquences**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Corcmal : Label : **Conséquences malheureuses de la corruption**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Corlut : Label : **Lutte contre la corruption par des institutions**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Corinst : Label : **Connaissance des institutions de lutte contre la corruption**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

Coreff : Label : **Efficacité des institutions de lutte contre la corruption**

Type : Caractère nominal

Taille :5

Valeur par défaut : *Néant*

QDF1 : Label : **Ressources publique fin personnelle**

Type : Numérique

Taille :8

Valeur par défaut : *Néant*

QDF2 : Label : **Arrangement**

Type : Numérique

Taille :8

Valeur par défaut : *Néant*

QDF3 : Label : **Altération, dépravations des valeurs**

Type : Numérique

Taille :8

Valeur par défaut : *Néant*

QDF4 : Label : **Autres**

Type : Numérique

Taille :8

Valeur par défaut : *Néant*

I.7. Les causes et effets de la corruption

Cette section définit les causes de la corruption dans notre pays et les conséquences engendrées par celle-ci.

Cause1 : Label : **Bas salaire**
Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Cause2 : Label : **Pauvreté**
Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Cause3 : Label : **Impunité**
Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Cause4 : Label : **Ignorance**
Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Cause5 : Label : **Analphabétisme**
Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Cause6 : Label : **Autres**
Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Corcma1 : Label : **Effritement de la société**
Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Corcma2 : Label : **Violence**
Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Corcma3 : Label : **Pauvreté**
Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Corcma4 : Label : **Perte de crédibilité vis à vis des bailleurs**

Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Corcma5 : Label : Délabrement de l'administration

Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Corcma6 : Label : Instabilité du régime

Type : Numérique
Taille :8
Valeur par défaut : *Néant*

I.8. Institution de lutte contre la corruption

Cette section évalue les connaissances que le répondant a des institutions de lutte contre la corruption.

Insti1 : Label : Inspections générales des finances

Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Insti2 : Label : Coordination nationale de lutte contre la fraude

Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Insti3 : Label : Inspection générale d'Etat

Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Insti4 : Label : Cours des comptes

Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Insti5 : Label : Comité national d'éthique

Type : Numérique
Taille :8
Valeur par défaut : *Néant*

Insti6 : Label : Haute autorité de coordination de lutte anti- corruption

Type : Numérique
Taille :8
Valeur par défaut : *Néant*

II. Remarques et suggestions

- ✓ Contrairement à la base de l'enquête prioritaire (EPII), la base corruption est beaucoup moins importante en taille et aussi moins structurée.
- ✓ Les critères de sélections de l'échantillon ne sont pas précisés.
- ✓ Il existe plusieurs variables de type 'caractère'. Ceux ci devront être recodées pour faciliter les analyses en type numérique avec des modalités.

CHAPITRE 3

GUIDE SPSS

SOMMAIRE

| | | |
|-------------|--|------------|
| I. | Présentation générale | 88 |
| 1.1. | Identification | 88 |
| 1.2. | Installation et spécification | 88 |
| 1.3. | Représentation schématique du logiciel SPSS | 89 |
| II. | Le vocabulaire de base | 90 |
| 2.1. | Démarrage et arrêt | 90 |
| 2.2. | Les fenêtres SPSS | 90 |
| 2.3. | La barre de menu | 92 |
| 2.4. | La barre d'outils | 93 |
| 2.5. | L'aide sur SPSS | 94 |
| III. | Structure et gestion des fichiers | 95 |
| 3.1. | Format des données supportées | 95 |
| 3.2. | Ouvrir un fichier de données | 95 |
| 3.3. | Importer des données | 95 |
| 3.4. | Informations sur les fichiers de données | 96 |
| 3.5. | Sauvegarde d'un fichier de données | 97 |
| IV. | Structure et gestion des données | 97 |
| 4.1. | Caractéristiques des variables et entrée des données | 97 |
| 4.2. | Transformation de variables | 98 |
| 4.3. | Suppression de variables | 99 |
| 4.4. | Suppression d'observations ou d'individus | 99 |
| 4.5. | Insérer une variable | 100 |
| 4.6. | Insérer une observation | 100 |
| 4.7. | Trier les observations | 100 |
| 4.8. | Transposition des cas | 100 |
| 4.9. | Fusionner des fichiers de données | 100 |
| 4.10. | Fragmenter un fichier | 101 |
| 4.11. | Sélectionner des observations | 101 |
| V. | Réalisation de statistiques descriptives univariées | 102 |
| 5.1. | Variable quantitative | 102 |
| 5.1.1. | Résumés statistiques | 102 |
| 5.1.2. | Résumés statistiques et graphiques | 103 |

| | | |
|----------|--|-----|
| 5.1.3. | Procédure EXPLORE | 104 |
| 5.2. | Variable qualitative : Table de fréquence et graphiques | 106 |
| VI. | Réalisation de statistiques bivariées | 107 |
| 6.1. | Liaison entre variables quantitatives | 107 |
| 6.1.1. | Nuages de points en deux dimensions | 107 |
| 6.1.2. | Diagrammes en barres pour variables quantitatives | 108 |
| 6.2. | Liaison entre variables qualitatives | 108 |
| 6.2.1. | Tables de contingence | 108 |
| 6.2.2. | Diagrammes en barres | 109 |
| 6.3. | Etude croisée d'une variable quantitative et d'une variable qualitative. | 112 |
| 6.3.1. | Résumés statistiques | 112 |
| 6.3.2. | Réalisation de graphique | 112 |
| 6.3.2.1. | Les diagrammes en barres | 112 |
| 6.3.2.2. | Les Boxplots croisés avec une variable qualitative | 113 |
| 6.3.2.3. | Les Boxplots croisés avec deux qualitatives | 114 |
| 6.3.2.4. | Faire un histogramme sur un sous-groupe de données quantitatives | 114 |
| 6.3.3. | Mise en forme des graphes | 115 |
| VII. | Manipulation de commandes et syntaxe. | 116 |
| 7.1. | Ouverture d'une page de syntaxe. | 116 |
| 7.2. | Règles d'écriture des commandes. | 116 |
| 7.3. | Exécution d'une commande. | 116 |
| 7.4. | Exemple de commandes. | 117 |
| 7.5. | Placer des commandes sur la page de syntaxe. | 117 |
| VIII. | Exemples de programme SPSS | 118 |
| 8.1. | Calcul d'un ratio : Taux d'accès à l'eau potable. | 118 |
| 8.2. | Calcul de l'indice de GINI et tracé de la courbe de Lorenz | 120 |
| IX. | Bibliographie | 123 |
| 9.1. | Les livres | 123 |
| 9.2. | Les sites Internet | 123 |

I.1 Présentation générale

1.1. Identification

L'apprentissage d'un nouveau logiciel, surtout lorsque celui-ci dispose de plusieurs fonctionnalités est une tâche très complexe. Il est donc nécessaire de disposer d'une base vous permettant de le faire aisément.

Ce présent guide de l'utilisateur est réalisé dans le but de vous familiariser avec les nombreuses fonctionnalités du logiciel SPSS 11.0 for Windows. Il vous présentera les fonctions principales ainsi que les procédures de réalisation des analyses les plus courantes.

Le logiciel SPSS (**Statistical Package for Social Sciences**) sous Windows est un outil d'analyse statistique. Il permet principalement de réaliser la codification, la saisie, la manipulation et l'utilisation des données à l'aide des fonctions informatiques et statistiques de base disponibles. Son utilisation exige une certaine familiarité avec l'environnement Windows.

1.2. Installation et spécification

SPSS est disponible sous Windows, Unix et Macintosh. Les fichiers peuvent être échangés d'une plate forme à l'autre sans conversion.

Pour l'installation, il faut disposer d'une licence. L'installation peut se faire à partir du CD-ROM ou à partir du site de SPSS <http://www.spss.com>.

Pour l'installation à partir du CD-ROM.

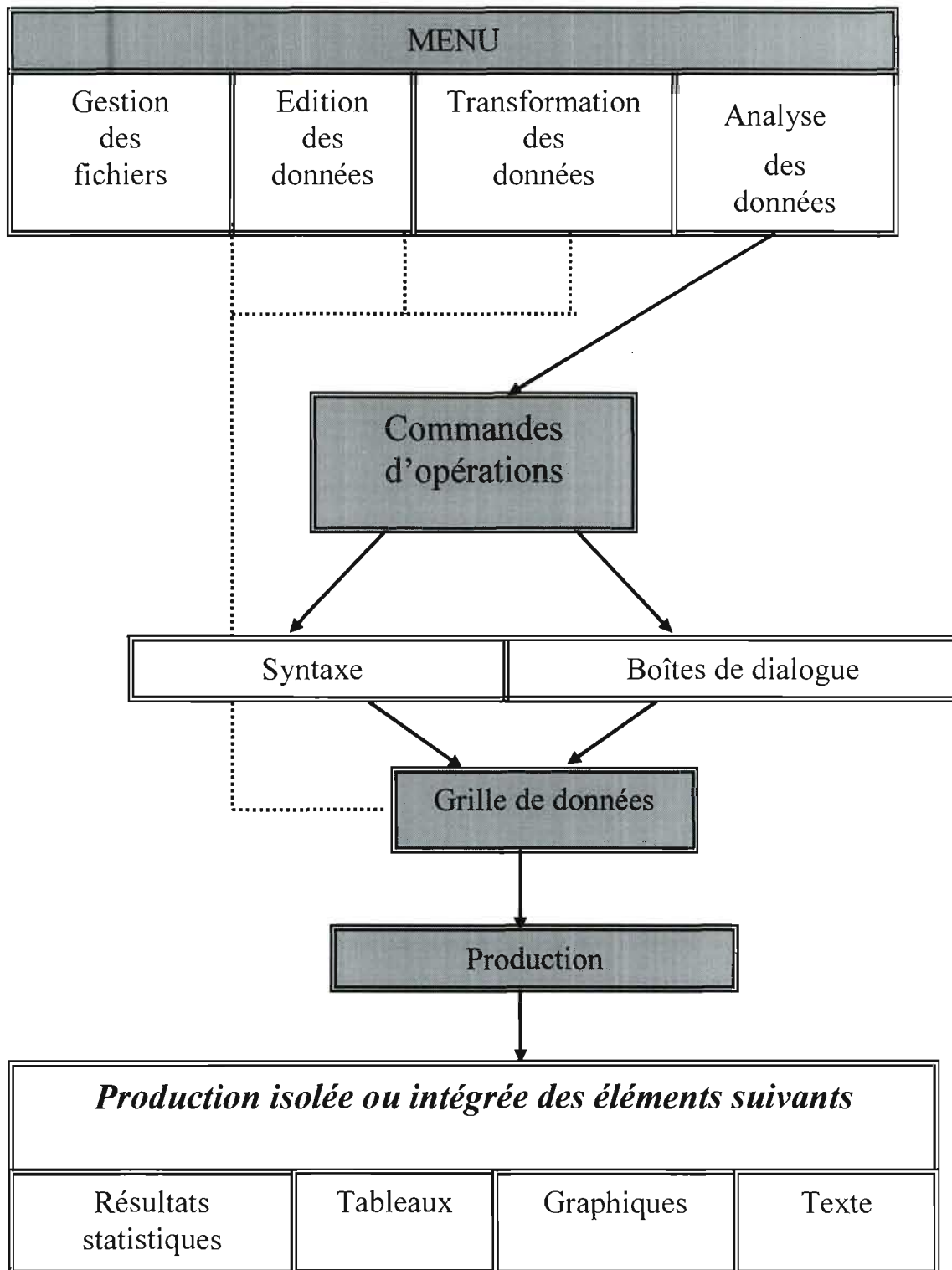
- ✓ Cliquer sur Install SPSS 11.0
- ✓ Suivre pas à pas les instructions de l'installation.

Le logiciel **Statistical Package for the Social Science (SPSS)** est un logiciel de gestion et d'analyse de données statistiques de portée générale.

Il est relativement facile à utiliser, vu que de nombreuses analyses statistiques peuvent être effectuées sans lignes de programmation, par des menus déroulant et des boîtes de dialogue. L'utilisateur peut toutefois avoir accès à des fonctions plus complexes ou des options qui n'apparaissent pas dans les boîtes de dialogue, en utilisant le langage de commande SPSS dans la fenêtre de syntaxe. Ces instructions peuvent être sauvegardées dans un fichier de syntaxe (extension .sps). La barre de menus principale comprend 9 menus dont le menu **Statistics** qui permet de sélectionner les procédures statistiques souhaitées.

La saisie et la gestion des données sont très souples grâce à l'éditeur de données SPSS qui permet de générer ou de manipuler les données d'un fichier (extension .sav). Les résultats d'une analyse statistique sous SPSS apparaissent dans une fenêtre intitulée **Output** (résultat) et peuvent être sauvegardés dans un fichier (extension .lst). Il faut noter que SPSS permet la production d'un ensemble de graphiques hautes résolutions (diagrammes en barres, secteurs, histogrammes, nuages de points,) pouvant être sauvegardés dans un fichier.

1.3.Représentation schématique du logiciel SPSS



I.1 Le vocabulaire de base

2.1. Démarrage et arrêt

Le démarrage de SPSS for Windows se fait selon le cheminement suivant :

Démarrer > Programmes > SPSS for Windows > SPSS for Windows 11.0.

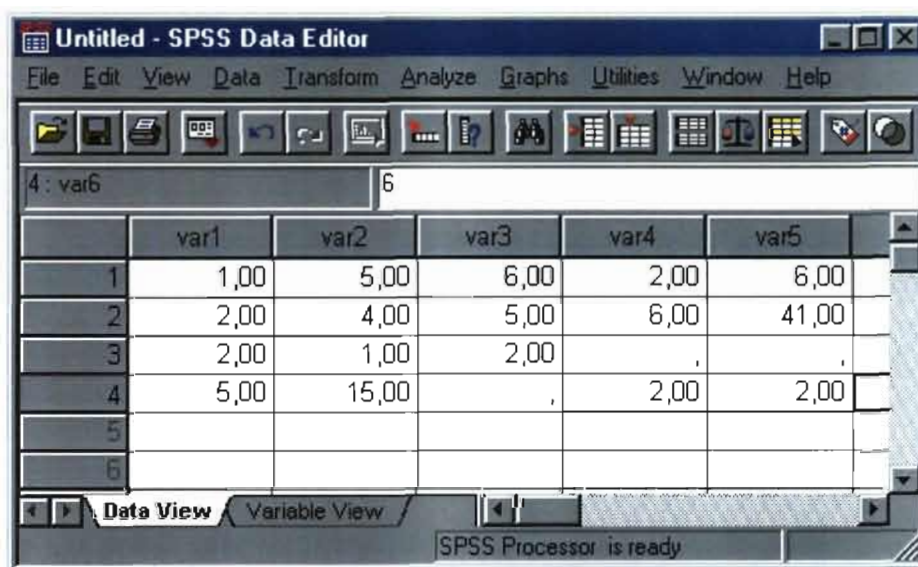
Sur la fenêtre qui s'affiche, précisez ce que vous souhaitez faire en sélectionnant l'option correspondante.

A partir d'une fenêtre SPSS, la procédure pour quitter est : **File > Exit** ou simplement avec le bouton fermer de la barre des menus.

2.2. Les fenêtres SPSS

2.2.1. L'éditeur de données (SPSS Data Editor)

Chaque fois que l'on commence une session de travail avec SPSS, la fenêtre "**Data Editor**" apparaît. C'est dans cette fenêtre que les données figurent. Nous pouvons directement y constituer notre fichier de données ou importer des données d'un autre fichier.

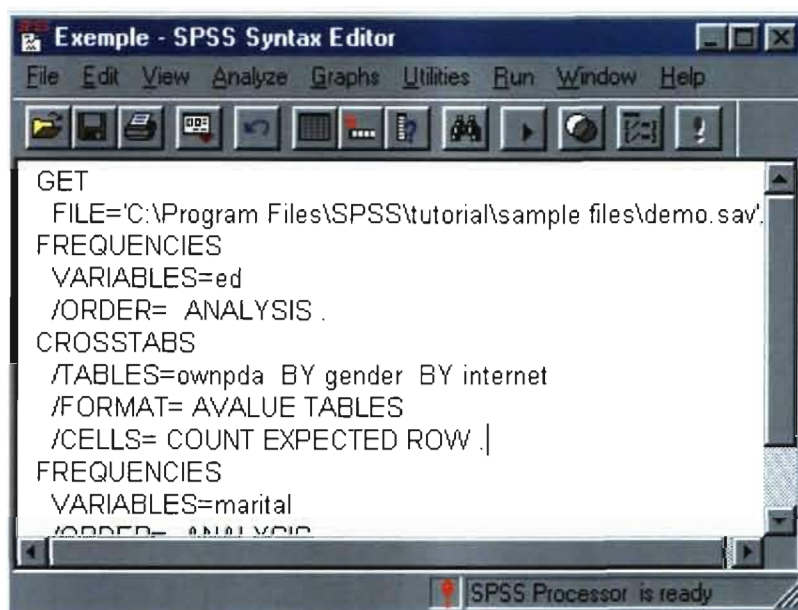


L'éditeur de données est constitué d'un **menu** de fonctions, d'une **barre d'outils** affichant des icônes associées aux fonctions les plus utilisées et d'une **grille de données**.

Les données sont structurées de la façon suivante :

- ✓ Chaque ligne de la grille représente un individu, un cas (case), par exemple l'individu qui répond au questionnaire.
- ✓ Chaque colonne représente les variables, par exemple une question du questionnaire.
- ✓ Chaque cellule contient une donnée et représente l'intersection du cas et de la variable.
- ✓ Les données peuvent être numériques ou alphanumériques.
- ✓ La taille de la grille est fonction du nombre de cas et de variables.

2.2.2. L'éditeur de syntaxe : SPSS Syntax Editor



Pour ouvrir une page de syntaxe à partir du Data Editor : **File > New > Syntax** pour une nouvelle fenêtre ou **File > Open > Syntax** pour choisir un fichier de syntaxe déjà existant.

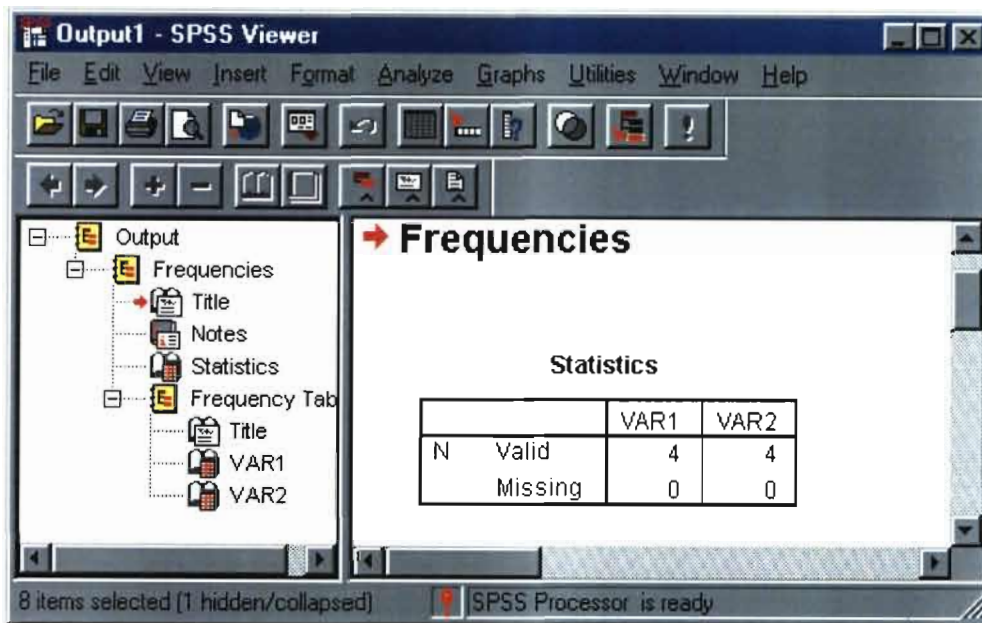
Pour que la page de syntaxe s'ouvre au démarrage de SPSS, à partir du Data Editor, sélectionnez **Options** dans la rubrique **File** du menu. Dans la boîte de dialogue **General** cochez l'option **Open Syntax Windows at start-up**.

Il est possible d'écrire directement des commandes sur la page de syntaxe pour les exécuter par la suite. Les opérations effectuées par le menu peuvent être également sauvegardées sur la page de syntaxe. Elles ne s'inscrivent pas automatiquement, il faut utiliser l'option **Paste** figurant sur les boîtes de dialogue, avant l'exécution, pour coller les commandes dans la page de syntaxe.

Cette page peut être sauvegardée et remaniée à volonté. Elle offre de nombreux avantages : elle permet d'éviter de modifier constamment la grille de données brutes.

La page de syntaxe doit être sauvegardée pour être éventuellement réutiliser avant de sortir de SPSS.

2.2.3. La fenêtre résultats de SPSS (Output).



Comme son nom l'indique, la page des résultats (Output) retient les résultats de chacune des procédures ou analyses exécutées. Elle s'ouvre automatiquement à l'exécution d'une opération avec le résultat. Si le contenu n'est pas effacé après chaque procédure, les résultats s'accumulent.

- ✓ La page de résultats peut être sauvegardée.
- ✓ Chaque tableau individuel peut être copié et collé dans un document géré par un logiciel différent.
- ✓ Un tableau croisé peut être collé dans le tableur Excel ou dans un tableau Word.

2.3. La barre de menu

Elle est constituée d'un ensemble de groupe de fonctions.












- ✓ Le menu **File** permet de gérer les fichiers (par exemple, ouvrir un nouveau fichier, enregistrer, fermer etc.) ;
- ✓ Le menu **Edit** permet d'effectuer les opérations de traitement de texte (par exemple, sélectionner, copier, couper, coller, etc.) ;
- ✓ Le menu **View** permet de définir les options de l'écran (par exemple, les onglets de la barre d'outils) ;






- ✓ Le menu **Data** traite de tout ce qui est en rapport avec la gestion des données (par exemple définir une variable, trier les données, insérer une variable etc.) ;
- ✓ Le menu **Transform** permet d'effectuer des opérations de transformation sur les variables (par exemple, la création d'indice, la recodification des données, etc.) ;
- ✓ Le menu **Analyze** permet de réaliser plusieurs types d'analyses statistiques (entre autres nous avons, les analyses descriptives, les corrélations, les régressions, les tests du Khi-2 etc.) ;
- ✓ Le menu **Graphs** présente tous les types de graphiques que SPSS permet de créer (par exemple, histogrammes, lignes, bar etc.) ;
- ✓ Le menu **Utilities** donne les utilitaires du programme (par exemple, des informations sur les fichiers, des informations sur les variables, etc.) ;
- ✓ Le menu **Windows** permet la gestion des fenêtres, leurs dispositions et le dimensionnement.
- ✓ Le menu **Help** est un outil d'aide à l'utilisation de SPSS.

2.4. La barre d'outils

Elle affiche des icônes associées aux fonctions les plus utilisées.



| | |
|---|---|
|  | Open File pour ouvrir un fichier |
|  | Save pour enregistrer |
|  | Print pour imprimer |
|  | Dialog Recal pour ouvrir les boîtes de dialogues fréquemment utilisées |
|  | Goto Case pour aller à une case particulière selon son numéro |
|  | Variables pour obtenir une description de la variable sélectionnée |
|  | Find pour effectuer une recherche |
|  | Insert Cases pour ajouter des cas |
|  | Insert Variable pour ajouter des variables |

| | |
|---|--|
|  | Split File pour fragmenter un fichier |
|  | Weigh Cases pour appliquer un poids aux cas |
|  | Select Case pour sélectionner des cas |
|  | Value Labels pour obtenir les étiquettes des valeurs saisies |
|  | Use Sets pour utiliser un ensemble ou sous-ensemble particulier |

2.5. L'aide sur SPSS

A chaque niveau d'une séance de travail dans SPSS, il est possible d'obtenir de l'aide par diverses manières.

- ✓ A travers la rubrique **Help** du menu
- ✓ Sur le bouton **Help** des boîtes de dialogues
- ✓ Par un clic droit sur les options pour de brèves explications dans les boîtes de dialogues.
- ✓ En sélectionnant l'option **Run the tutorial** au démarrage.

En plus du tutorial, la rubrique d'aide renferme le **statistics coach** qui vous donne une assistance à la réalisation de vos analyses. Il y a également le **Syntax guide** qui donne une description des commandes ainsi que des exemples d'utilisation.

I.1 Structure et gestion des fichiers

3.1. Format des données supportées

SPSS accepte les données sous les formats suivants :

- ✓ Les tableurs Lotus et Excel ;
- ✓ Les bases de données dBASE et divers formats SQL ;
- ✓ Les fichiers en texte ASCII (.txt) ou délimité par une tabulation, des espaces et virgules ;
- ✓ Les fichiers de données sous format SPSS créés sous d'autres systèmes d'opérations

3.2. Ouvrir un fichier de données

A partir du menu, **File > Open > Data**. A partir de la barre d'outils .

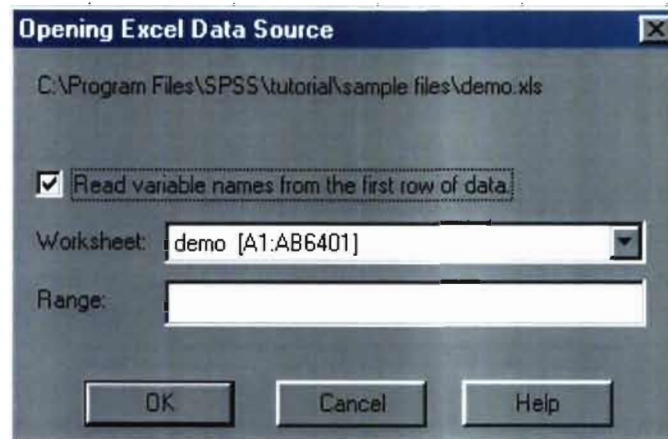
- ✓ Si le fichier a une extension **.sav** (format SPSS), choisissez le et cliquez sur **Open**.
- ✓ Si le fichier est d'un autre format, la lecture des données doit s'accompagner de certaines opérations : Pour les tableurs, vous pouvez lire les noms des variables à partir de la première rangée. Les noms comportant plus de huit (8) caractères seront tronqués. Si deux noms de variables deviennent identiques, l'une sera renommée.

3.3. Importer des données

Pour les formats de donnée supportés par SPSS, l'importation est faite directement dans SPSS. L'importation des formats non supportés peut se faire à travers un logiciel intermédiaire tel EPIDATA.

EXCEL ->SPSS

- ✓ Dans SPSS, faire: **File > Open > Data**
- ✓ Choisir le fichier EXCEL(.xls) à importer.
- ✓ Répondre à la question: le crochet dans la case indique que la première ligne du jeu de données contient les noms des variables.



ASCII-> SPSS

A l'aide de la fonction *Text Wizard*, SPSS vous permet de lire des fichiers délimités par des tabulations, par des espaces, par des virgules ou à champs fixes.

Il y a six étapes à suivre pour lire un fichier texte :

- ✓ Dans SPSS, faire: **File > Open > Read Text Data**
- ✓ Choisir le fichier texte (.txt ou .dat) à importer et cliquer sur Open.
- ✓ Etape 1 de **Wizard**: Répondre à la question (Si le format est bon) ou appuyer sur **Next** pour définir un nouveau format.
- ✓ Etape 2 de **Wizard**: Répondre aux 2 questions et appuyer sur **Next**.
- ✓ Etape 3 de **Wizard**: Répondre aux 3 questions et appuyer sur **Next**.
- ✓ Etape 4 de **Wizard**: Spécifier le délimiteur du fichier texte.
- ✓ Etape 5 de **Wizard**: On peut ici modifier le format du jeu de données à importer
Ex: retirer une variable, mettre le format dollar, etc.
- ✓ Etape 6 de **Wizard**: Répondre aux 2 questions et appuyer sur **Finish**.

Avant de cliquer sur *Terminer* pour exécuter la procédure, vous pouvez conserver ce format pour l'appliquer à un autre fichier de données vous pouvez également conserver la syntaxe des opérations exécutées.

Il faut noter que **Text Wizard** permet d'extraire les données de format texte lorsque le fichier de départ est bien structuré. **Text Wizard** ne corrige pas les erreurs dans les fichiers de données en format texte.

3.4. Informations sur les fichiers de données

En plus des données, chaque fichier dispose des informations concernant les variables qu'il renferme. Ces informations peuvent s'obtenir des deux façons suivantes :

- ✓ Pour la grille active de données, à partir du Menu, faites : **Utilities > File Info**.
- ✓ Pour les autres fichiers, à partir du Menu, faites : **File>Display Data Info**

3.5. Sauvegarde d'un fichier de données

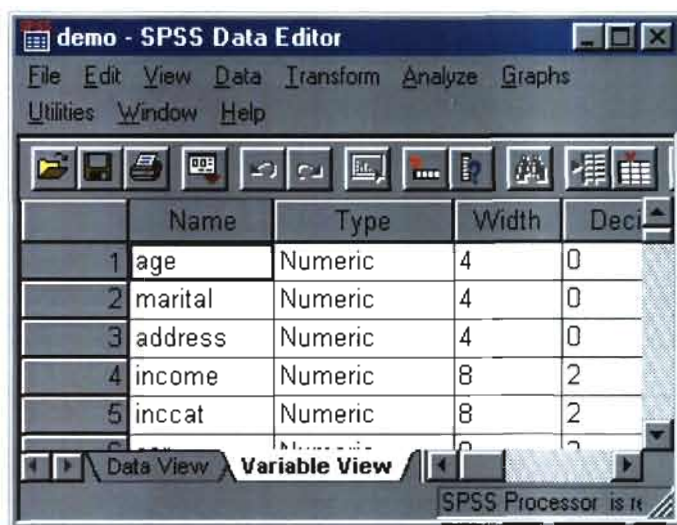
La sauvegarde d'un fichier est une opération très capitale, mais si elle n'est suivie de certaines précautions elle peut être source de désagréments. Dans SPSS, la procédure est : **File > Save** ou **File > Save as**.

Avant toute opération de sauvegarde, il faut vérifier les changements apportés à votre fichier d'origine pour voir s'il y a nécessité de les enregistrer. Il faut noter que la sauvegarde est une opération irréversible. De ce fait il est fortement recommandé d'effectuer des copies de vos fichiers d'origine et d'y effectuer vos manipulations. Avant de terminer une séance, vérifier les nécessités de sauvegarde afin de ne pas commettre d'erreurs irréparables sur votre fichier initial.

I.1 Structure et gestion des données

4.1. Caractéristiques des variables et entrée des données

Les variables sont décrites dans le **Data Editor**. Par défaut le Data Editor s'ouvre avec le **Data View**. Pour obtenir la vue des variables, sélectionnez l'option **Variable View** au bas de l'écran. Les variables se trouvent sur les lignes et les colonnes décrivent les caractéristiques des variables.



Les caractéristiques principales dans une vue des variables :

- ✓ **Name** : Le nom des variables. Il débute par une lettre et ne peut se terminer par un point. Il ne contient pas d'espace et ne peut comporter plus de huit (8) caractères
- ✓ **Type** : La nature de variable (numérique, date, etc.).
- ✓ **Width** : Largeur de la donnée. Elle indique le nombre de chiffres qui composent la donnée. Selon le format, la donnée apparaissant dans la cellule peut différer de la valeur entrée. Pour augmenter ou diminuer la largeur, il faut cliquer dans la case correspondante et faire varier la taille.
- ✓ **Decimals** : Nombre de décimaux. Elle indique le nombre de chiffres après la virgule. Il est possible de le faire varier.

-
- ✓ **Labels** : Etiquettes ou description de la variable.
 - ✓ **Values** : Valeurs et description de chacune d'elles.
 - ✓ **Missing** : Valeurs manquantes correspondent à une attribution de certaines valeurs comme manquantes.
 - ✓ **Columns** : nombre de colonnes sur la grille.
 - ✓ **Align** : C'est la disposition des données (données alignées à droite, à gauche, centrées, etc.)
 - ✓ **Measure** : C'est l'échelle de mesure ; variable nominale ou ordinale.

L'entrée des données.

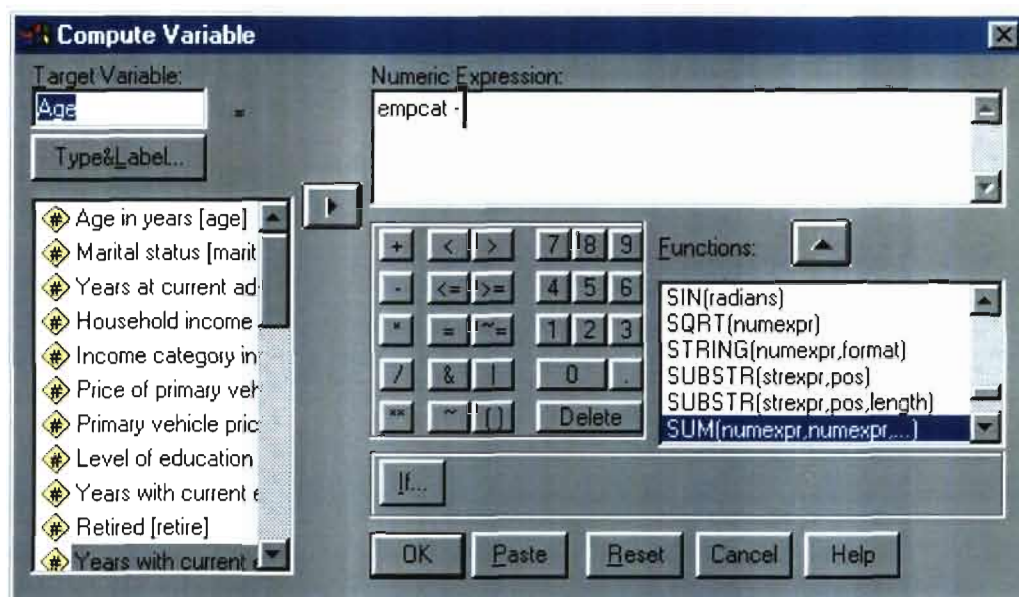
Elle se fait dans le Data Editor dans la vue des données **Data View**.

- ✓ La cellule active est encadrée.
- ✓ Vous pouvez entrer les données dans n'importe quel ordre, par cas ou par variables ou au hasard.
- ✓ Le nom et le numéro de la variable apparaissent dans le coin supérieur gauche de l'éditeur de données.
- ✓ Pour entrer toute donnée non numérique, vous devez d'abord définir le type de la variable.
- ✓ Vous pouvez apporter toutes sortes de modification à vos données dans la grille.

4.1. Transformation de variables

La fonction **COMPUTE** : Elle permet d'attribuer des valeurs à une nouvelle variable en fonction de la transformation d'autres variables.

- ✓ Sélectionnez **COMPUTE** dans le menu **TRANSFORM**. La boîte de dialogue suivante apparaît.



- ✓ Dans la case **TARGET VARIABLE**, introduisez le nom de la nouvelle variable, obtenue par transformation d'une ou de plusieurs variables du tableau de données.
- ✓ Dans la case **NUMERIC EXPRESSION**, tapez la formule adéquate.
- ✓ Sélectionnez la **fonction** adéquate dans la liste et cliquez sur la flèche au-dessus de la liste, à côté de **FUNCTIONS**.
- ✓ Sélectionnez ensuite la (ou les) **variable(s)** du tableau et cliquez sur la flèche à droite de la liste des variables.
- ✓ Cliquez **OK**.

4.2. Suppression de variables

- ✓ Cliquez une fois sur le nom de la variable à effacer. On sélectionne ainsi la colonne Correspondant à la variable à effacer.
- ✓ Sélectionnez **CLEAR** du menu **EDIT**.

4.3. Suppression d'observations ou d'individus

- ✓ Sélectionnez la ligne correspondant à l'individu à effacer en cliquant une fois sur le numéro de la ligne.
- ✓ Sélectionnez **CLEAR** du menu **EDIT** ou appuyez sur la touche **delete** de votre clavier.
- ✓ Pour annuler la dernière opération effectuée, sélectionnez **UNDO** du menu **EDIT**.

4.4. Insérer une variable

- ✓ Positionnez-vous sur la colonne du tableau de données, suivant celle à insérer.
- ✓ Sélectionnez **INSERT VARIABLE** du menu **DATA**.

4.5. Insérer une observation

- ✓ Positionnez-vous sur la ligne du tableau de données, suivant celle à insérer.
- ✓ Sélectionnez **INSERT CASE** du menu **DATA**.

4.6. Trier les observations

- ✓ Dans la rubrique **Data**, sélectionnez l'option **Sort Cases**.
- ✓ Sélectionnez les variables de tri.
- ✓ Choisissez l'option de tri **Ascending** ou **Descending**.
- ✓ Cliquez sur **OK**

4.7. Transposition des cas

- ✓ Sélectionnez l'option **Transpose** sous la rubrique **Data**
- ✓ Sélectionnez les variables à transposer
- ✓ Cliquez sur **OK**

4.8. Fusionner des fichiers de données

Lorsque les données sont distribuées dans plusieurs fichiers, il est nécessaire de les fusionner pour les besoins de l'analyse. Il y'a deux types de fusion :

✓ **L'ajout de cas** : Cette fusion permet d'ajouter dans un fichier principal des cas provenant d'un autre fichier. La procédure est la suivante :

- Sélectionnez dans la rubrique **Data** l'option **Merge Files**, puis l'option **Add Case** puis, ouvrez le fichier contenant les données à ajouter.
- Sélectionnez les variables du fichier de destination.
- Cliquez sur **OK**

✓ **L'ajout de variables** : Les cas doivent être placés dans le même ordre dans les deux fichiers de données. Les variables portant le même nom que ceux du fichier principal seront exclus de la fusion. Il y'a cependant possibilité de les renommer.

- Sélectionnez dans la rubrique **Data** l'option **Merge Files**, puis l'option **Add Variables** puis, ouvrez le fichier contenant les données à ajouter.
- Renommer les variables portant le même nom que ceux du fichier principal afin de les prendre en compte dans la fusion.

4.9. Fragmenter un fichier

- ✓ Sélectionnez l'option **Split File** de la rubrique Data.
- ✓ Sélectionnez les variables de regroupement.
- ✓ Cliquez sur **OK**

4.10. Sélectionner des observations

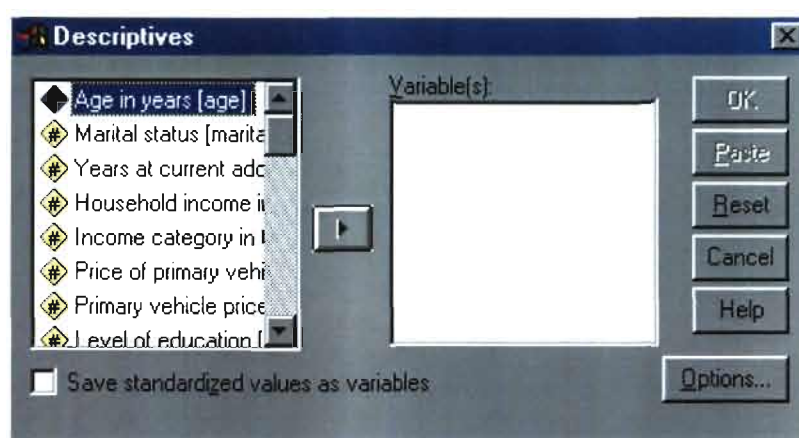
- ✓ Sélectionnez l'option **Select Case** de la rubrique Data.
- ✓ Précisez les conditions de sélection
- ✓ Cliquez sur **OK**

I.1 Réalisation de statistiques descriptives univariées

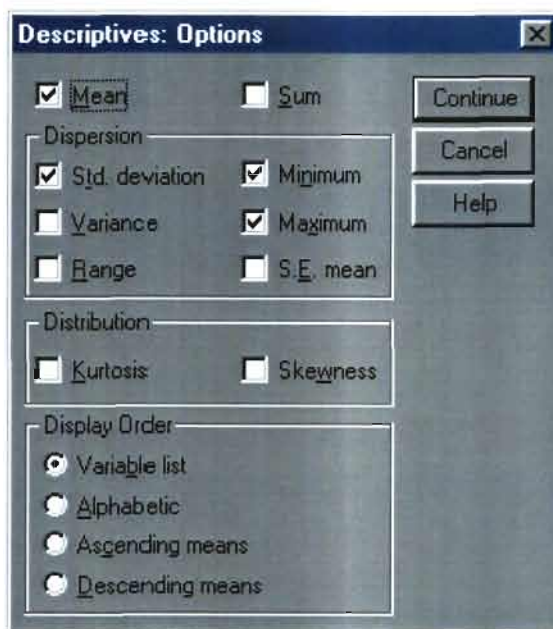
5.1 Variable quantitative

5.1.1. Résumés statistiques

- ✓ Sélectionnez **Descriptive Statistics** du menu **Analyze**, puis cliquez **Descriptives**.
- ✓ Dans la boîte de dialogue **Descriptives**, sélectionnez une variable quantitative et cliquez sur la flèche pour la faire passer du côté **variable(s)**. Vous pouvez sélectionner plusieurs variables en recommençant la même opération ou en appuyant sur la touche **Ctrl** et en cliquant sur les différentes variables qui vous intéressent.



- ✓ Pour choisir les indicateurs statistiques, cliquez **Options** dans la boîte de dialogue **Descriptives** ; pour chaque indicateur, cochez son nom dans la boîte **Descriptives : Options**.

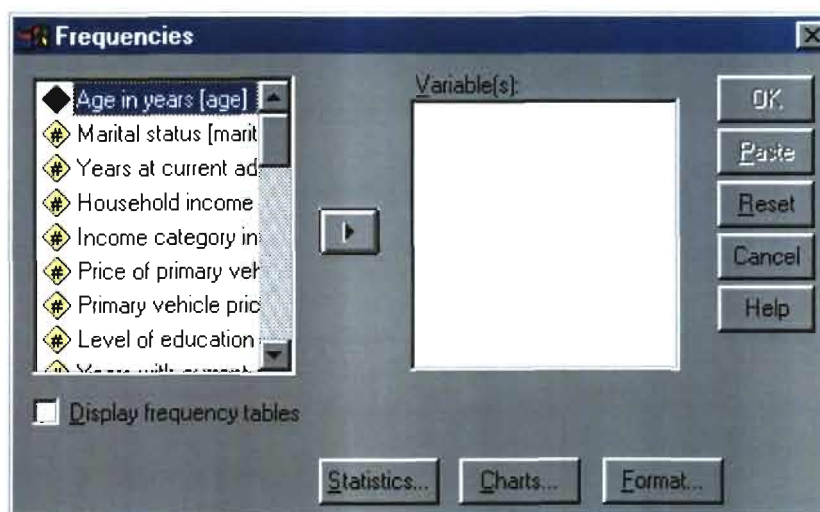


- ✓ Cliquez **Continue**.
- ✓ Cliquez **OK** dans la boîte de dialogue **Descriptives**.

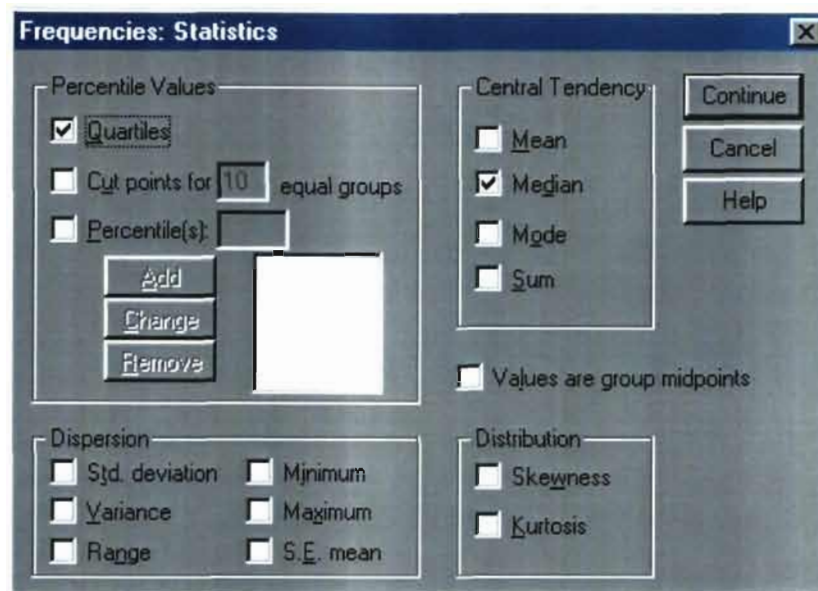
5.1.2. Résumés statistiques et graphiques

Le résumé renvoie un ensemble de statistiques descriptives selon votre choix. La procédure est la suivante.

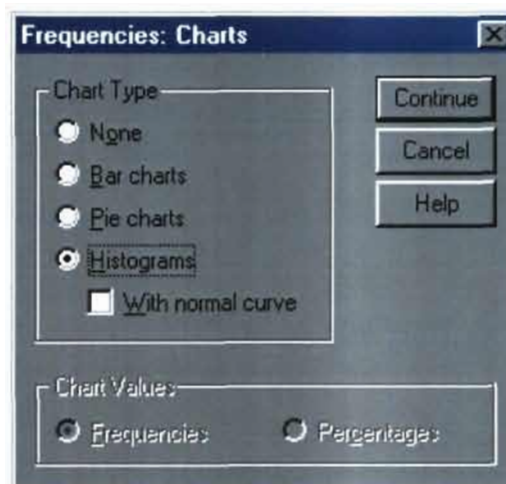
- ✓ Sélectionnez **Descriptive Statistics** du menu **Analyze**, puis cliquez **Frequencies**.
- ✓ Dans la boîte de dialogue **FREQUENCIES**, sélectionnez une variable quantitative et cliquez sur la flèche pour la faire passer du côté **Variable(s)**. N'activez pas **Display Frequency Tables**.



- ✓ Pour choisir les indicateurs statistiques, cliquez **Statistics** dans la boîte de dialogue **FREQUENCIES**; pour chaque indicateur, cochez son nom dans la boîte de dialogue **FREQUENCIES : Statistics**.



- ✓ Cliquez **Continue**.
- ✓ Pour choisir les graphiques, cliquez **Charts** dans la boîte de dialogue **FREQUENCIES**;
- ✓ Cochez le nom du graphique souhaité dans la boîte de dialogue **FREQUENCIES : Charts**.



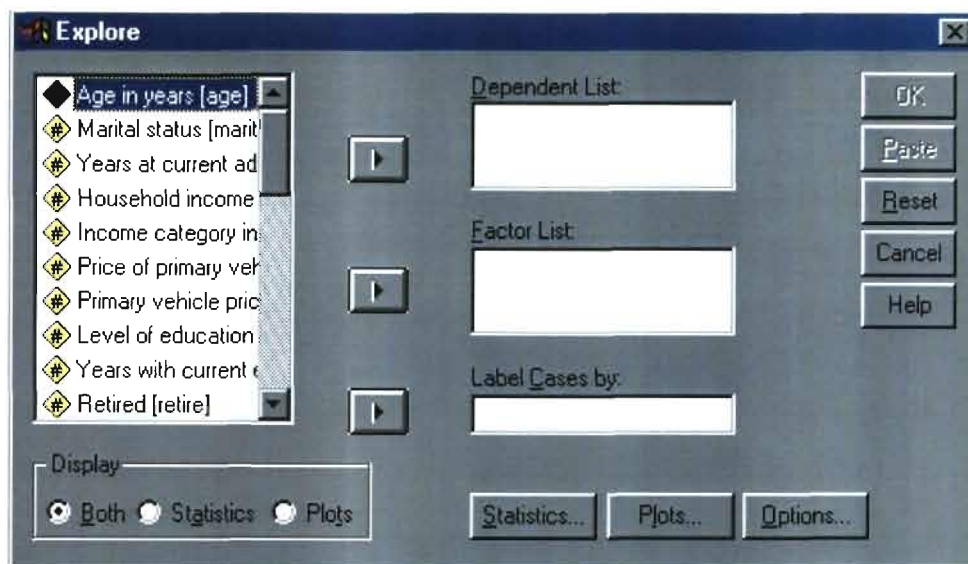
- ✓ Cliquez **Continue**.
- ✓ Cliquez **OK** dans la boîte de dialogue **FREQUENCIES**.

5.1.3. Procédure EXPLORE

La procédure **EXPLORE** permet de réaliser très facilement des analyses descriptives et graphiques sur une ou plusieurs variables. Elle permet aussi l'obtention d'un diagramme en tiges et feuilles.

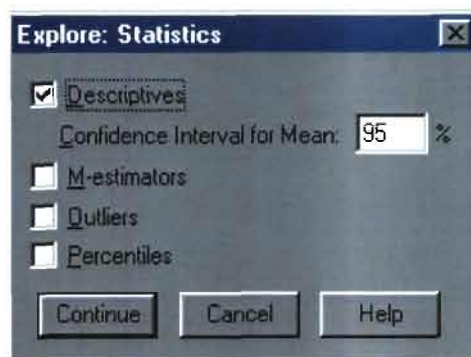
- ✓ Sélectionnez **Descriptive Statistics** du menu **Analyze**, puis cliquez **EXPLORE**.

- ✓ Dans la boîte de dialogue, sélectionnez une ou plusieurs variables quantitatives et cliquez sur la flèche devant **Dependent List**

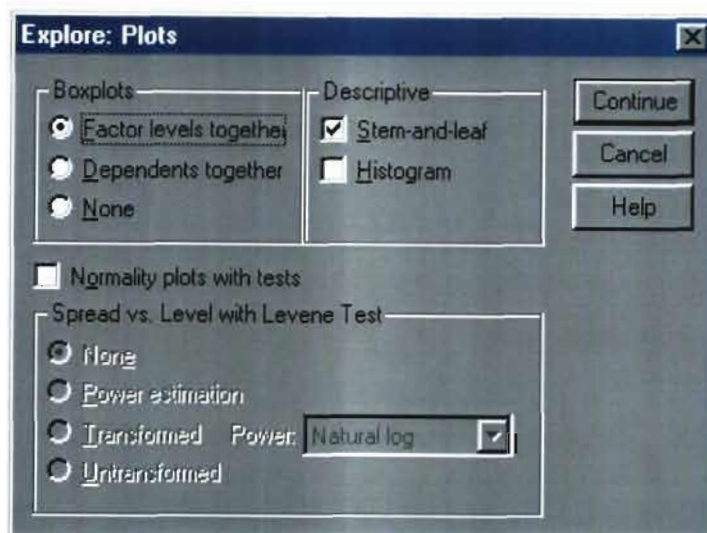


Vous pouvez maintenant choisir le type d'information que vous souhaitez obtenir en cliquant dans la case *display* sur **Statistiques**, sur **Plots** ou sur **Both**, ce dernier vous permettant d'obtenir les statistiques et les graphiques. Selon votre choix, vous pourrez accéder aux options correspondantes en cliquant sur **statistics** ou sur **Plots**.

- ✓ **Statistics** vous permet de choisir les statistiques que vous souhaitez obtenir.



- ✓ **Plots** vous permet de choisir les graphiques que vous souhaitez obtenir.



- ✓ Cliquez sur **Continue** pour chaque groupe d'options.
- ✓ Cliquez **OK** dans la boîte de dialogue **Explore**.

5.2. Variable qualitative : Table de fréquence et graphiques

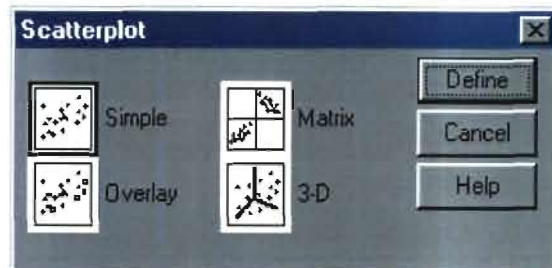
- ✓ Sélectionnez **Descriptive Statistics** du menu **Analyze**, puis cliquez **EXPLORE**.
- ✓ Dans la boîte de dialogue **FREQUENCIES**, sélectionnez une variable qualitative et cliquez sur la flèche pour la faire passer du côté **variable(s)**.
- ✓ Activez **Display frequency Tables**.
- ✓ Pour choisir les graphiques, cliquez **Charts** dans la boîte de dialogue **FREQUENCIES** ;
- ✓ Cliquez sur le nom du graphique souhaité dans la boîte de dialogue **FREQUENCIES : Charts**.
- ✓ Cliquez **Continue**.
- ✓ Cliquez **OK** dans la boîte de dialogue **FREQUENCIES**.

I.1 Réalisation de statistiques bivariées

6.1. Liaison entre variables quantitatives

La procédure **Scatterplot** permet de réaliser des nuages de points de deux variables quantitatives (en 2 dimensions) ou de trois variables (en 3 dimensions).

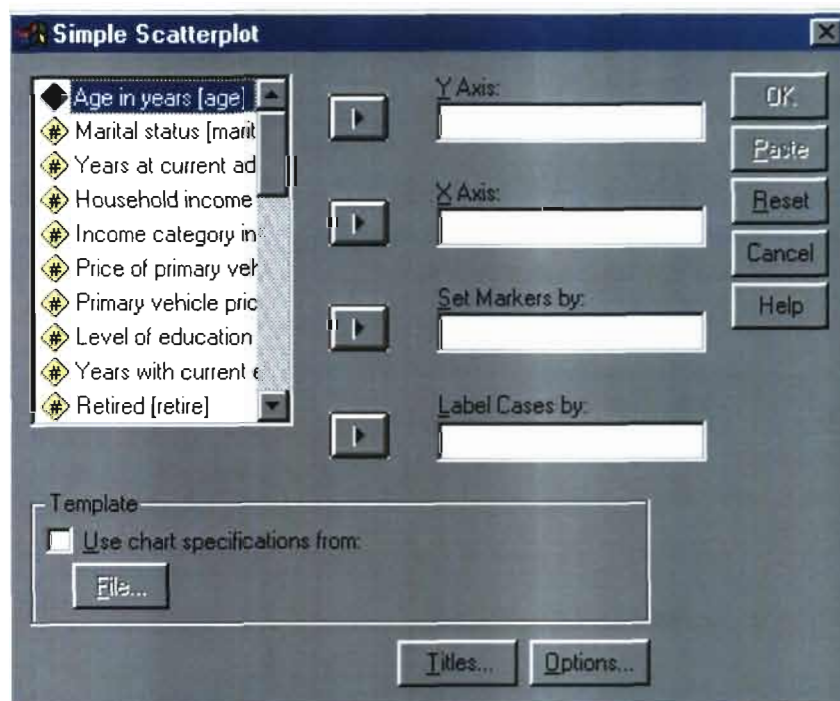
- ✓ Cliquez sur **SCATTER** du menu **GRAPHS**.



6.1.1 Nuages de points en deux dimensions

La procédure **Simple Scatterplot**, permet de réaliser un nuage de points X*Y en deux dimensions avec la possibilité de repérer les points par une variable marqueur qualitative. La variable Y fournit les ordonnées et la variable X les abscisses des points représentés.

- ✓ Dans la boîte de dialogue "**Scatterplot**", sélectionnez l'icône "**Simple**", puis cliquez sur **Define**.
- ✓ Sélectionnez une variable quantitative pour l'axe horizontal X, et cliquez sur la flèche correspondante ; faites de même avec une autre variable quantitative pour l'axe vertical Y ; ces variables doivent être numériques.
- ✓ Cliquez **OK**.



Remarque :

En sélectionnant une variable qualitative et en l'amenant dans la boîte **Set Markers by**, le graphique indique de plus dans quelle catégorie se situe chaque couple de points tracés.

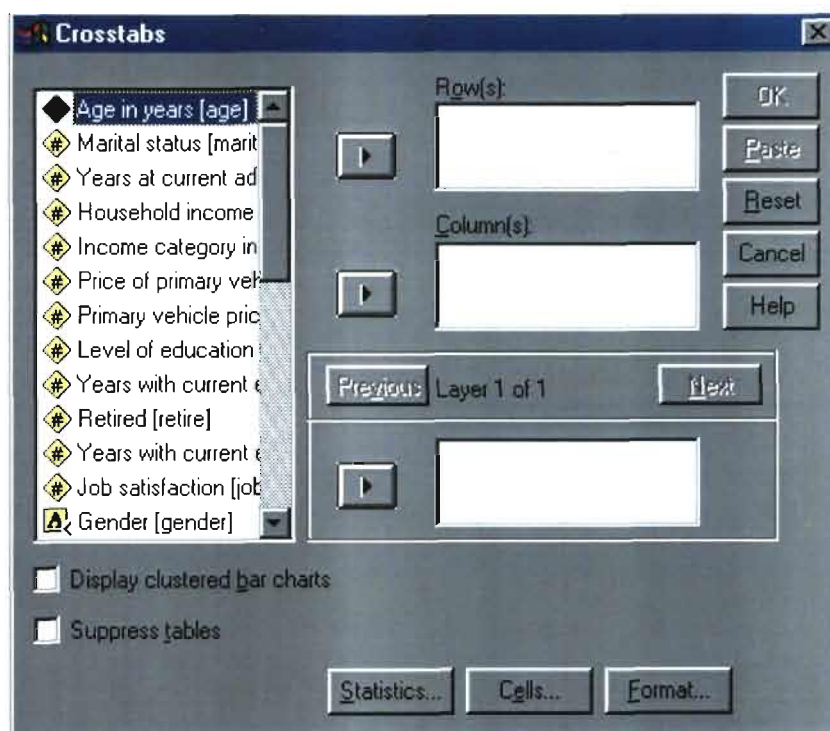
6.1.2. Diagrammes en barres pour variables quantitatives

Confer 6.3. : Etude croisée d'une variable quantitative et d'une variable qualitative.

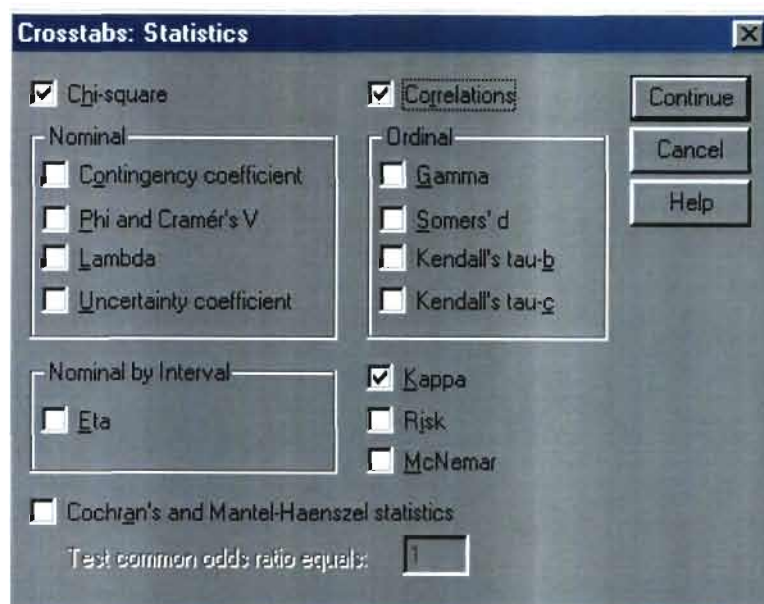
6.2. Liaison entre variables qualitatives**6.2.1. Tables de contingence**

La procédure **CROSSTABS** permet d'obtenir des tables croisées de fréquence de plusieurs variables qualitatives. Les tables à deux entrées (bivariées) concernent le croisement de deux variables qualitatives. On peut aussi créer des tables multivariées (plus de deux variables qualitatives), mais leur lecture devient vite complexe.

- ✓ Sélectionnez **Descriptive Statistics** du menu **Analyze**, cliquez sur **CROSSTABS**.
- ✓ Dans la boîte de dialogue **CROSSTABS**, sélectionnez une variable en ligne puis cliquez sur la flèche correspondante (**Row(s)**) ; faites de même pour une variable en colonne (**Column(s)**)



- ✓ En cliquant sur **Statistics** de la boîte de dialogue **CROSSTABS**, vous pouvez sélectionner une série de statistiques, cliquez sur **Continue**.



✓ Cliquez **OK**.

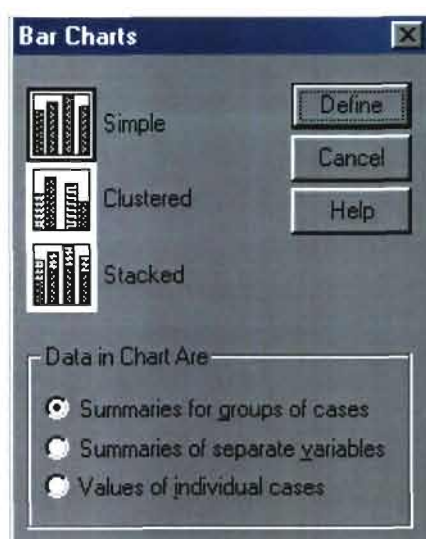
Remarque :

Pour créer des tables de contingences multivariées, et selon la présentation souhaitée de ces tables, il suffit de sélectionner plusieurs variables qualitatives en ligne (**Row(s)**) et plusieurs variables qualitatives en colonnes (**Column(s)**).

6.2.2. Diagrammes en barres

La procédure **Bar Charts** permet de représenter la fréquence observée (ou le pourcentage) pour les différentes modalités d'une variable qualitative (**SIMPLE**) et de décomposer les fréquences en fonction des modalités d'une autre variable qualitative (**Clusters variable** ou **Stacks variable**).

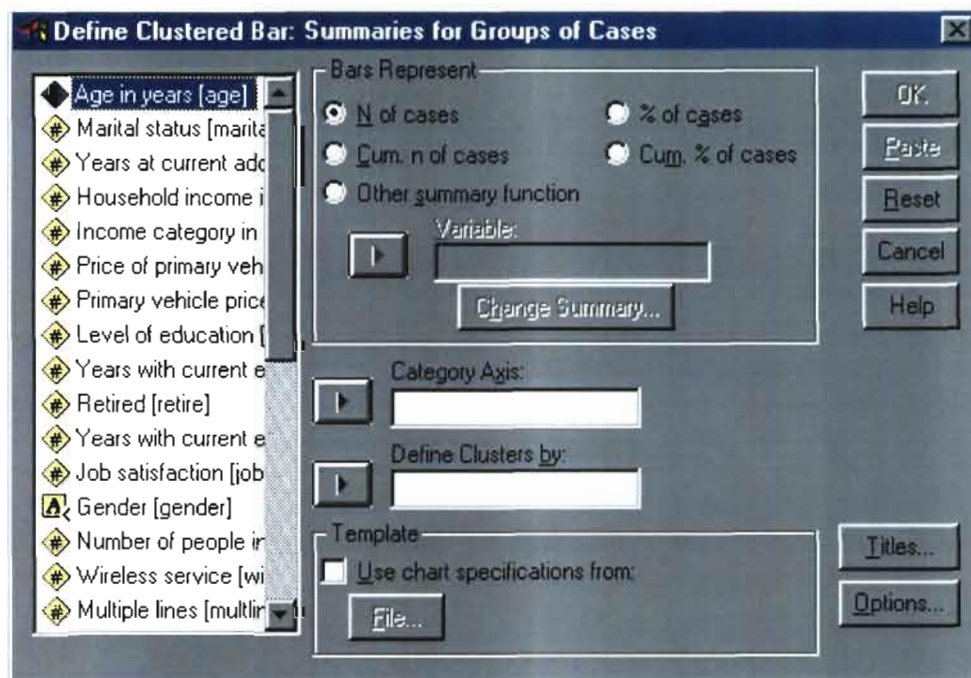
✓ Sélectionnez **BAR** du menu **GRAPHS**.



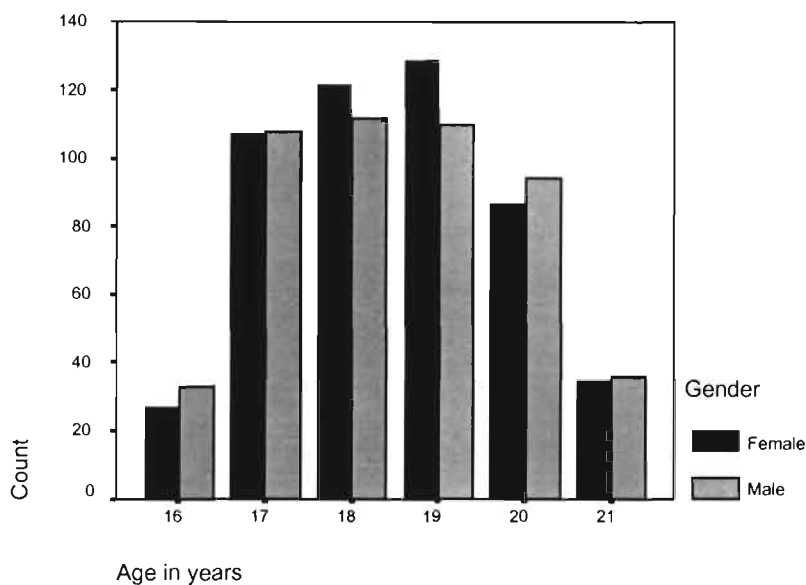
✓ Choisissez **SIMPLE** pour une seule variable

✓ Pour plusieurs variables croisées, sélectionnez l'icône **Clustered**, ou bien l'icône **Stacked**.

- ✓ Pour représenter les diagrammes en barres d'une variable qualitative suivant les modalités d'une autre variable qualitative, sélectionnez l'option **Summaries for groups of cases**, puis cliquez sur **Define**.
- *Si vous avez choisi l'icône Clustered,*
 - ✓ Cette option permet de représenter plusieurs groupes de barres côte à côte suivant les modalités de la variable qualitative spécifiée.
 - ✓ Sélectionnez une variable pour l'axe des catégories, et cliquez sur la flèche en face de la boîte variable **Category Axis**. Il est recommandé de choisir une variable qui ne présente que quelques niveaux différents.
 - ✓ Faites un choix de représentation des données sur l'axe d'échelle (effectifs, effectifs cumulés, pourcentages, pourcentages cumulés, ...) dans la boîte **Bars Represent**.
 - ✓ Sélectionnez une variable de classification et cliquez sur la flèche en face de la boîte **Define Clusters by**.
 - ✓ Cliquez **OK**.

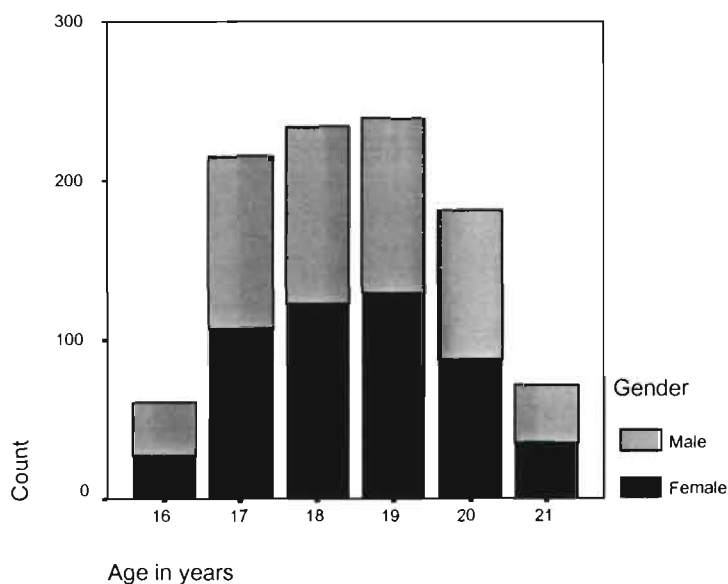


Par exemple, les instructions reprises ci-dessus donneront le graphe suivant. Comme on peut le voir, il y a deux diagrammes en barre, un pour chaque sexe (variable *Category Axis*). En utilisant l'option *Edit*, il vous est possible de modifier le graphique comme par exemple de changer les couleurs.



○ *Si vous avez choisi l'icône Stacked*

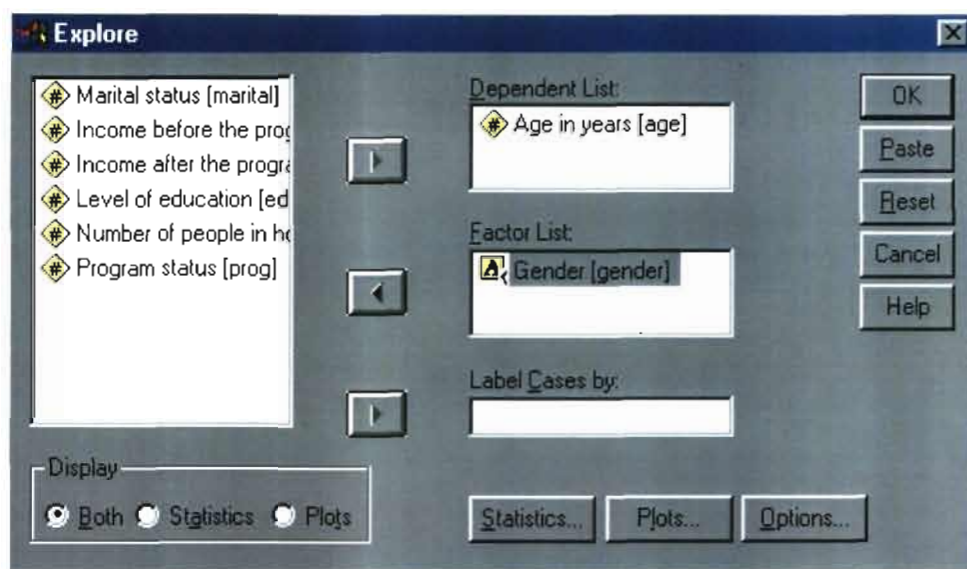
Cette option réalise une superposition des barres qui représente la participation de chaque modalité de la variable évaluée en fonction de la variable qualitative spécifiée. La fenêtre est similaire à celle qui précède. Elle produit un graphique comme celui-ci.



6.3. Etude croisée d'une variable quantitative et d'une variable qualitative.

6.3.1. Résumés statistiques

- ✓ Sélectionnez **Descriptive Statistics** du menu **Analyze**, puis cliquez **EXPLORE**
- ✓ Dans la boîte de dialogue **EXPLORE**, sélectionnez une variable quantitative et cliquez sur la flèche devant **Dependent List**.
- ✓ Ensuite, sélectionnez une variable qualitative et cliquez sur la flèche devant **Factor List**.



- ✓ Pour le reste, reportez-vous à la **procédure Explore**.

6.3.2. Réalisation de graphique

6.3.2.1. Les diagrammes en barres

On peut vouloir représenter la moyenne (ou la médiane) d'une variable quantitative pour différentes modalités d'une ou même de deux variables qualitatives. Pour ce faire, sélectionnez comme précédemment **BAR** du menu **GRAPHS** et dans la boîte de dialogue l'option **Summaries for Group of Cases**, à l'aide du bouton **DEFINE**, sélectionnez **OTHER SUMMARY FUNCTION** et faites entrer la variable quantitative de votre choix.

Vous pouvez modifier la mesure de tendance centrale proposée qui par défaut est la moyenne.

Pour terminer, sélectionnez la ou les variables qualitatives et passez les dans les espaces nommées **CATEGORY AXIS** (pour la 1ère variable) et **DEFINE CLUSTERS BY** (pour la seconde).

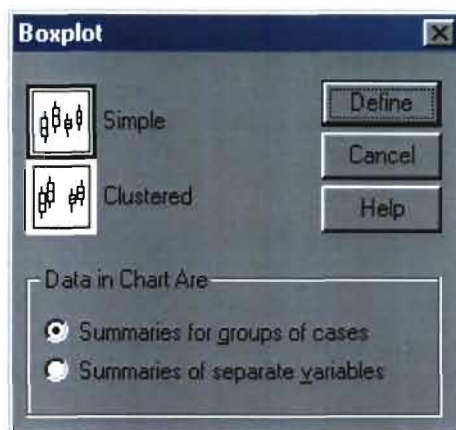
Si vous désirez représenter la valeur des moyennes de plusieurs variables quantitatives sur le même graphique, choisissez l'option **Summaries of separate variables** dans la boîte de dialogue **BARS**.

Sélectionnez les variables quantitatives de votre choix et ajoutez-les sous l'intitulé **BARS REPRESENT**. Ensuite, sélectionnez la variable qualitative sous **CATEGORY AXIS**.

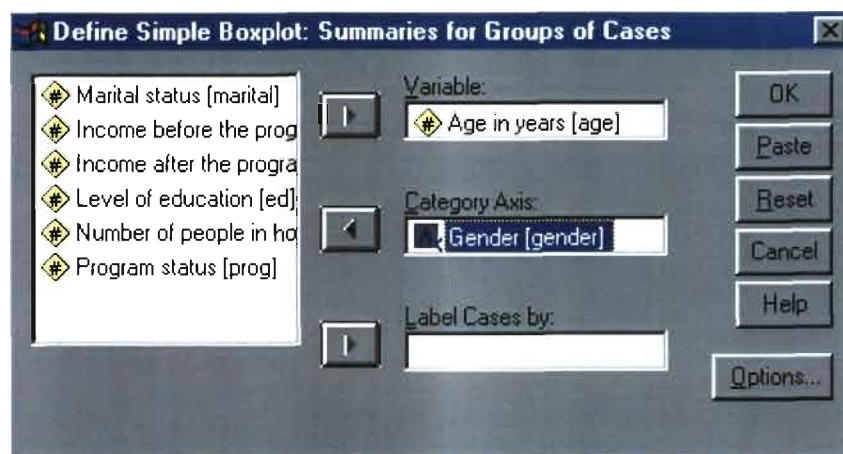
Pour représenter les diagrammes en barres des valeurs d'une ou de plusieurs variables quantitatives pour chaque individu, sélectionnez, dans la boîte de dialogue **Bar Charts**, l'option **Values of individual cases**.

6.3.2.1. Les Boxplots croisés avec une variable qualitative

- ✓ Choisissez **BOXPLOT** du menu **GRAPHS**.
- ✓ Sélectionnez l'icône **Simple**, puis l'option **Summaries for groups of cases**.



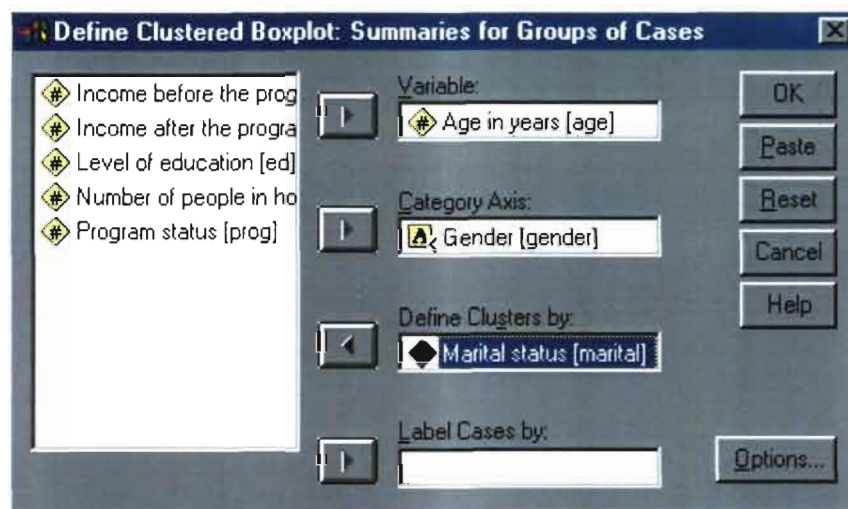
- ✓ Cliquez **DEFINE**.



- ✓ Sélectionnez une variable quantitative, et amenez-la dans la boîte **Variable** en cliquant sur la flèche correspondante.
- ✓ Sélectionnez une variable qualitative, et amenez-la dans la boîte **Category Axis** en cliquant sur la flèche correspondante.
- ✓ Cliquez **OK**

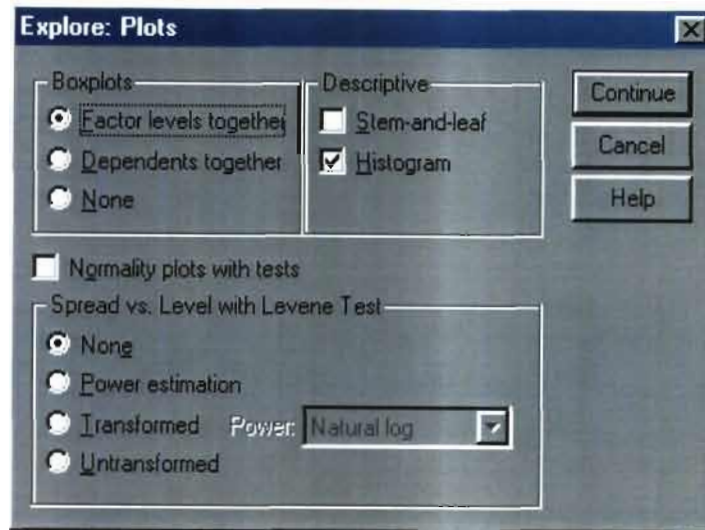
6.3.2.2. Les Boxplots croisés avec deux qualitatives

- ✓ Choisir **BOXPLOT** du menu **GRAPHS**.
- ✓ Sélectionnez **Clustered**, puis l'option **Summaries for groups of cases**,
- ✓ Sélectionnez une variable quantitative, et amenez-la dans la boîte **Variable** en cliquant sur la flèche correspondante.
- ✓ Sélectionnez une variable qualitative, et amenez-la dans la boîte **Category Axis** en cliquant sur la flèche correspondante.
- ✓ Sélectionnez une deuxième variable qualitative, et amenez-la dans la boîte **Define Clusters by** en cliquant sur la flèche correspondante.
- ✓ Cliquez **OK**.



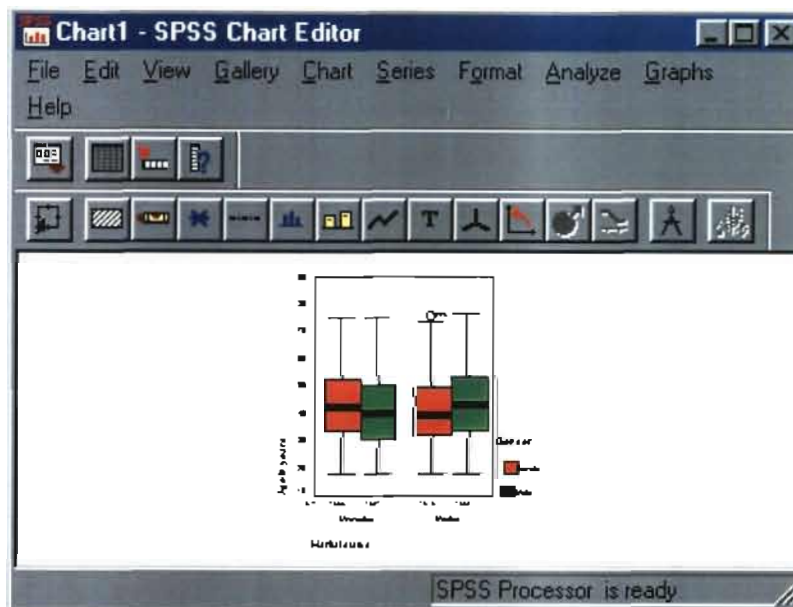
6.3.2.3. Faire un histogramme sur un sous-groupe de données quantitatives

Pour réaliser ce genre de graphique, vous pouvez utiliser la procédure **EXPLORE** (voir plus haut) en ne demandant que les histogrammes dans l'option **Plot**. Vous pouvez aussi bien sûr demander les Boxplots si vous le souhaitez.



6.3.3. Mise en forme des graphes

Pour modifier un élément du graphique, il faut le sélectionner. Pour cela, il suffit de double-cliquer dessus ou tout près de lui. Une fenêtre **Chart Editor** s'ouvre avec le graphique sélectionné.



Cette fenêtre vous permet de changer de nombreuses caractéristiques de votre graphique. Pour cela vous devez cliquer sur l'élément que vous voulez changer :

- ✓ Pour la taille des marques, dans le menu **Format**, sélectionnez **Marker**, choisissez la taille dans la fenêtre **Size**, cliquez sur **Apply** ;
- ✓ Pour leur forme dans le menu **Format**, sélectionnez **Marker**, choisissez la forme dans la fenêtre **Style**, cliquez sur **Apply**
- ✓ Pour la couleur dans le menu **Format**, sélectionnez **Color**, choisissez la couleur, cliquez sur **Apply**.

- ✓ Pour la mise en forme de la légende par exemple, double cliquez sur celui et appliquez le format voulu. Vous pouvez la centrer sélectionner **Center** dans **Title Justification** mais aussi changer beaucoup d'autres paramètres.

I.1 Manipulation de commandes et syntaxe.

La plupart des commandes sont exécutées en passant par le menu et les boîtes de dialogue. Il est aussi possible d'exécuter ces commandes à partir de la page de syntaxe qui peut être conservée et modifiée à votre gré. Même en exécutant les commandes à partir d'une boîte de dialogue, avant de les exécuter, elles peuvent être ajoutées à la page de syntaxe.

7.1. Ouverture d'une page de syntaxe.

A partir d'une boîte de dialogue, si vous appuyer sur le bouton **Paste**, le progiciel ouvrira une page de syntaxe à condition qu'une page de syntaxe ne soit pas déjà ouverte. Par le menu vous pouvez ouvrir plusieurs pages de syntaxe en faisant à chaque fois **File>Open>Syntax**.

7.2. Règles d'écriture des commandes.

Les commandes doivent être écrites selon des règles particulières :

- ✓ Chaque commande débute sur une nouvelle ligne et se termine par un point.
- ✓ La plupart des commandes sont séparés par une barre oblique(/).
- ✓ Le nom de chaque variable doit être écrit au complet.
- ✓ Un texte inséré entre guillemet doit être écrit sur une même ligne.
- ✓ Chaque ligne de commande ne peut contenir plus de 80 caractères.
- ✓ La ligne de commentaire commence par le symbole (*) et se termine par un point (.)
- ✓ Un point(.) au lieu d'une virgule doit être utilisé pour indiquer le décimal.
- ✓ Le logiciel ne différencie pas les lettres minuscules des lettres MAJUSCULES sauf pour les noms de fichiers.
- ✓ Il est fortement recommandé d'éviter d'insérer un point dans le nom d'une variable.
- ✓ Il ne faut jamais utiliser la touche "*Tabulation*" comme remplaçante de plusieurs espaces blancs. SPSS interprète incorrectement la touche "*Tabulation*". Il ne faut donc jamais l'utiliser. Pour insérer plusieurs espaces blancs dans une instruction, on tape plusieurs fois la touche "*espace blanc*".

7.3. Exécution d'une commande.

L'exécution des commandes se fait de la façon suivante :

- ✓ **Run >All** ou le raccourci suivant qui figure sur la barre d'outil ► permet d'exécuter toutes les commandes saisies.
- ✓ **Run Selection:** Exécute la commande sélectionnée.

- ✓ **Run Current:** Exécute la commande sur laquelle le curseur est positionné.
- ✓ **Run To End:** Exécute toutes les commandes situées entre la position actuelle du curseur et la fin du fichier.

7.4. Exemple de commandes.

| Syntaxe | Rôle |
|--|---|
| GET FILE = <i>nom_du_fichier</i> | Accéder à un fichier SPSS |
| RECODE <i>variables(anciennes valeurs = nouvelle valeur)</i> (<i>anciennes valeurs = nouvelle valeur</i>) | Transformer une suite de valeurs ou un intervalle en une nouvelle valeur. |
| MATCH FILES FILE = <i>nom_du_1^{er}_fichier</i> /FILE = <i>nom_du_2eme_fichier</i> | Fusionner des fichiers de données |
| COUNT <i>variable=liste de variables(valeurs admissibles)</i> | ``Compter" les occurrences parmi les valeurs admissibles associées aux variables mentionnées |
| COMPUTE <i>variable=expression arithmétique</i> | Créer ou modifier une variable à l'aide d'opérations arithmétiques |
| SORT CASES BY <i>variables</i> | Trier le fichier de données selon les variables mentionnées |
| SPLIT FILE BY <i>variable</i> | Permettre d'exécuter les procédures demandées pour chaque sous-groupe de données déterminé par les valeurs de la variable désignée. |
| WEIGHT BY <i>variable</i> | Permettre d'associer à une donnée un poids (pondération) proportionnel à la valeur de la variable désignée. |
| EXECUTE | Permettre à certaines instructions d'être exécutées, c'est à dire d'être équivalentes à des procédures |

7.5. Placer des commandes sur la page de syntaxe.

Dans la plupart des boîtes de dialogues, on trouve le bouton **Paste**. En cliquant dessus, les opérations exécutées s'inscriront sur la page de syntaxe active.

Avant d'exécuter les analyses, vous pouvez exiger que la syntaxe s'inscrive sur la page des résultats en même temps que les résultats statistiques eux-mêmes. Pour y parvenir,

faites : **Edit>Options>Viewer**. Cochez l'option **Displays commands in log**. Une fois que vos commandes seront exécutées vous aurez la syntaxe dans le Output File. Vous pouvez l'exporter dans une page de syntaxe en faisant **Copy Paste** après sélection.

I.1 Exemples de programme SPSS

8.1. Calcul d'un ratio : Taux d'accès à l'eau potable.

Nous allons effectuer le calcul du taux d'accès à l'eau potable sur la base de données de l'Enquête Prioritaire de 1998 (EPII). Le fichier Ep2fi6.sav intitulé «logement et confort » sera utilisé. Dans ce fichier nous avons la variable s6q15 qui renvoie le mode d'approvisionnement en eau du ménage.

Pour obtenir le taux d'accès à l'eau potable des ménages, il nous faut déterminer dans un premier temps le nombre de ménage ayant accès à l'eau potable.

La variable s6q15 à plusieurs modalités :

- 1 Rivière, cours d'eau
- 2 Puits ordinaire
- 3 Puits busés
- 4 Forage
- 5 Fontaine publique
- 6 Robinet intérieur propre
- 7 Robinet intérieur partagé
- 8 Autres

Nous allons considérer comme ayant accès à l'eau potable les ménages n'utilisant pas les rivières, les cours d'eau ou les puits ordinaires comme mode d'approvisionnement.

Pour le calcul, nous utiliserons la fonction **IF** pour la sélection, et **COMPUTE** pour le calcul du ratio. Nous utiliserons aussi **AGGREGATE** et **MATCH FILES** pour les calculs intermédiaires.

***Syntaxe de calcul du taux d'accès à l'eau potable.**

*Ouverture du fichier renfermant le mode d approvisionnement en eau.
GET FILE = ep2fi6.sav.

*Sélection des individus ayants accès à l'eau potable.
IF (s6q15<>1) **AND** (s6q15<>2).

*Sauvegarde du fichier de travaille.
SAVE OUTFILE= Eaupotable.sav.

*Création d'une variable d'agrégation.
COMPUTE indic = 1.
VARIABLE LABELS indic 'variable d'agrégation.

*Agrégation pour calculer la taille de l'échantillon.
AGGREGATE OUTFILE= ficha.sav
/BREAK = indic
/taille 'taille totale de léchantillon' = **SUM** (indic).

*Fusion du résultat au fichier de travail.
MATCH FILES /FILE = */**TABLE** = ficha.sav **/BY** indic.

*Agrégation pour le calcul du nombre des ménages ayants accès à l'eau potable.
AGGREGATE OUTFILE= fichb.sav
/BREAK = indic
/Nbmen 'totale des ménages ayant de l eau Potable' = **SUM** (existe).

*Fusion du résultat au fichier de travail.
MATCH FILES /FILE = */**TABLE** = fichb.sav **/BY** indic.

*Calcul du ratio (Nombre des ménages ayant accès à l'eau potable/ Effectif total)
COMPUTE txmen= 100*(Nbmen/taille).
VARIABLE LABELS txmen 'Taux des ménages ayant de leau potable'.

EXECUTE.

8.2. Calcul de l'indice de GINI et tracé de la courbe de Lorentz

L'indice de GINI permet de mesurer l'inégalité des revenus. Il est déduit de la courbe de LORENTZ. Cette courbe met en rapport la distribution des effectifs cumulés avec les masses de revenus cumulés.

$$\text{Indice de GINI} = 1 - \sum_{i=1}^N f_i (G_{ei} + G_{e(i-1)}) \text{ avec } G_{e0} = 0$$

f_i : Fréquences des l'effectifs cumulés
 G_{ei} : Masses de revenus cumulés.

Pour le calcul des fréquences, il est nécessaire d'avoir l'effectif total des individus de la base. Nous avons créé une variable appelée indic, constante égale à 1 qui nous a servit de clé d'agrégation dans le calcul de la taille de l'échantillon. Nous avons aussi fait usage de la commande **MATCH FILES / FILE = */TABLE** afin d'affecter la valeur de la taille à toutes les classes.

On a ensuite calculé la fréquence relative de chaque classe en divisant l'effectif de la classe par la taille ($f_i = \text{effect}(i)/\text{taille}$). Après avoir calculé les fréquences relatives, nous avons calculé les fréquences cumulées en utilisant la fonction **CSUB**. Le calcul des masses cumulées s'est fait de façon similaire.

Pour aboutir finalement au calcul de l'indice nous avons utilisé d'autres commandes telle la boucle **DO IF**, **MATCH FILES** et la fonction **LAG** qui nous a permis de repérer G_{ei} et $G_{e(i-1)}$.

Pour la représentation graphique, nous aurons en abscisse les masses relatives cumulées et en ordonnée les fréquences relatives cumulées. La commande **PLOT** permet d'afficher la courbe ainsi que la première bissectrice.

* *Calcul de l'indice de GINI et représentation de la courbe de LORENTZ* *.

* *Ouverture du fichier contenant le revenu des ménages.*

GET FILE = ep2fi11.sav.

* *Calcul de l'effectif des observations par classe de salaire et création du fichier de travail.*

AGGREGATE

/OUTFILE= basedetravail.sav

/BREAK=s94c2

/effect 'effectif par classe'= N.

GET FILE basedetravail.sav.

* *Calcul de la taille l'échantillon et fusion de fichiers par l'intermédiaire du fichier fich1.sav.*

COMPUTE indic=1.

VARIABLE LABELS indic 'variable daggregation'.

AGGREGATE OUTFILE =fich1.sav

/BREAK=indic

/taille 'Talle de l échantillon' = **SUM**(effect).

MATCH FILES /FILE = * **/TABLE** =fich1.sav **/BY** indic.

* *Calcul des fréquences relatives et cumulées des classes.*

COMPUTE freq = effect/taille.

VARIABLE LABELS freq 'Fréquences relatives'.

CREATE freqcum =**CSUB** (freq).

VARIABLE LABELS freqcum 'Fréquence cumulée'.

* *calcul de la masse et de la somme cumulée des masses par l'intermédiaire de fich2.sav.*

COMPUTE masse= s94c2*effect.

VARIABLE LABELS masse 'Masse par classe'.

AGGREGATE OUTFILE = fich2.sav

/BREAK = indic

/smasse 'somme cumulée des masses' = **SUM**(masse).

MATCH FILES /FILE = * **/TABLE** =fich2.sav **/BY** indic.

* *Calcul des fréquences relatives et cumulées des masses.*

COMPUTE freqmas =masse/smasse.

VARIABLE LABELS freqmas 'fréquence des masses'.

CREATE fmascum =**CSUB**(freqmas).

VARIABLE LABELS fmascum 'Fréquence cumulée des masses'.

* *Calcul du produit des sommes cumulées des masses par les fréquences.*

DO IF (\$casenum=1).

COMPUTE somcum=fmascum.

ELSE.

COMPUTE somcum=fmascum + **LAG**(fmascum).

END IF.

COMPUTE prod=somcum*freq.

**Calcul de la somme des $f_i(G e_i + G e_{i-1})$.*

```
AGGREGATE OUTFILE=fich3.sav
  /BREAK=indic
  /somprod=SUM(prod).
```

```
MATCH FILES /FILE = * /TABLE =fich3.sav /BY indic.
```

```
COMPUTE GINI=1-somprod.
VARIABLE LABELS GINI 'indice de gini'.
```

**Courbe de LORENTZ.*

PLOT

```
/HSIZE=150
/VSIZE=150
/MISSING=LISTWISE
/FORMAT=OVERLAY
/TITLE="courbe de lorentz"
/HORIZONTAL="Masses cumulées" MIN(0) MAX(1)
/VERTICAL="Fréquences cumulées" MIN(0) MAX(1)
/PLOT=fmascum fmascum WITH freqcum fmascum (PAIR).
```

EXECUTE.

I.1 Bibliographie

9.1. Les livres

Livre : Analyse de données avec SPSS pour Windows versions 10.0 et 11.0

Livre : Analyse de données avec SPSS pour Windows versions 10.0 et 11.0 - Applications en gestion, assurance qualité et en sciences humaines - SMG

Auteur: F. Ouellet, G. Baillargeon

Livre : Etudes Marketing avec SPSS

Livre : Etudes Marketing avec SPSS - Pearson Education

Auteur: N. Malhotra, J. Décaudin, A. Bouguerra

SPSS Programming and Data Management.

Livre: Gestion des bases de données

Auteur: Raynal Levesque

Livre : Traitement des données avec SPSS pour Windows - Edition 8.0/9.0

Livre : Traitement des données avec SPSS pour Windows - Edition 8.0/9.0 - SMG

Auteur: F.Ouellet G.Baillargeon

9.2. Les sites Internet

Pour vous perfectionner dans l'exploitation du logiciel SPSS, il est très nécessaire de visiter des sites Internet et de participer aux foires à questions.

Nous vous recommandons particulièrement les sites suivant :

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

<http://pages.infinet.net/rlevesqu/SampleSyntax.htm>

www.spsstools.net

<http://www.sosig.ac.uk/statistics/>

SOMMAIRE

| | | |
|-------------|---|------------|
| I. | La Présentation générale du logiciel STATA | 119 |
| 1.2 | Installation et spécification | 119 |
| 1.3 | Représentation schématique du logiciel Stata | 121 |
| II. | Le vocabulaire de base | 122 |
| 2.1 | Le démarrage et l'arrêt | 122 |
| 2.2 | Les quatre fenêtres | 122 |
| 2.3 | La barre de menu | 124 |
| 2.4 | La barre d'outils | 126 |
| 2.5 | Les différents types de fichiers | 127 |
| III. | La programmation avec le logiciel STATA | 130 |
| 3.1 | Syntaxe générale des commandes STATA | 130 |
| 3.2 | La gestion et la manipulation des données | 132 |
| 3.3 | La création et l'utilisation des variables | 135 |
| 3.4 | La gestion des fichiers | 135 |
| 3.5 | La création et l'utilisation des fichiers 'log' et 'do' | 144 |
| 3.6 | Travailler en ligne avec STATA | 145 |
| 3.7 | Ecrire des programmes avec STATA | 148 |
| 3.8 | Quelques cas pratiques | 150 |
| 3.9 | La réalisation de graphiques | 152 |
| IV. | Bibliographie | 154 |
| 4.1 | les livres et articles | 154 |
| 4.2 | les sites Internet | 154 |

I. La Présentation générale du logiciel STATA

1.1 Identification

Ce guide d'utilisation destiné aux utilisateurs moyens de STATA présente quelques fonctionnalités du logiciel. STATA est un logiciel d'analyses statistiques développé et distribué par STATA CORPORATION 4905 Lakeway Drive, College Station, Texas 77845, United States. Tel 1-979-696-4600. Fax 1-979-696-4601. Email : stata@stata.com Corporate website : <http://www.stata.com>. Ce présent guide est réalisé avec la version 7.00 du dit logiciel qui comporte un langage de programmation qui peut s'apprendre assez rapidement, de nombreuses fonctions de calculs statistiques, de modélisations et de réalisation de graphiques complexes propre aux publications scientifiques. STATA réalise ces opérations avec le maximum d'optimisation ce qui le rend très rapide (d'où son caractère interactif). Il faut remarquer que les applications statistiques couvertes par le logiciel STATA concernent principalement l'économétrie, les sciences sociales et la biostatistique.

1.2 Installation et spécification

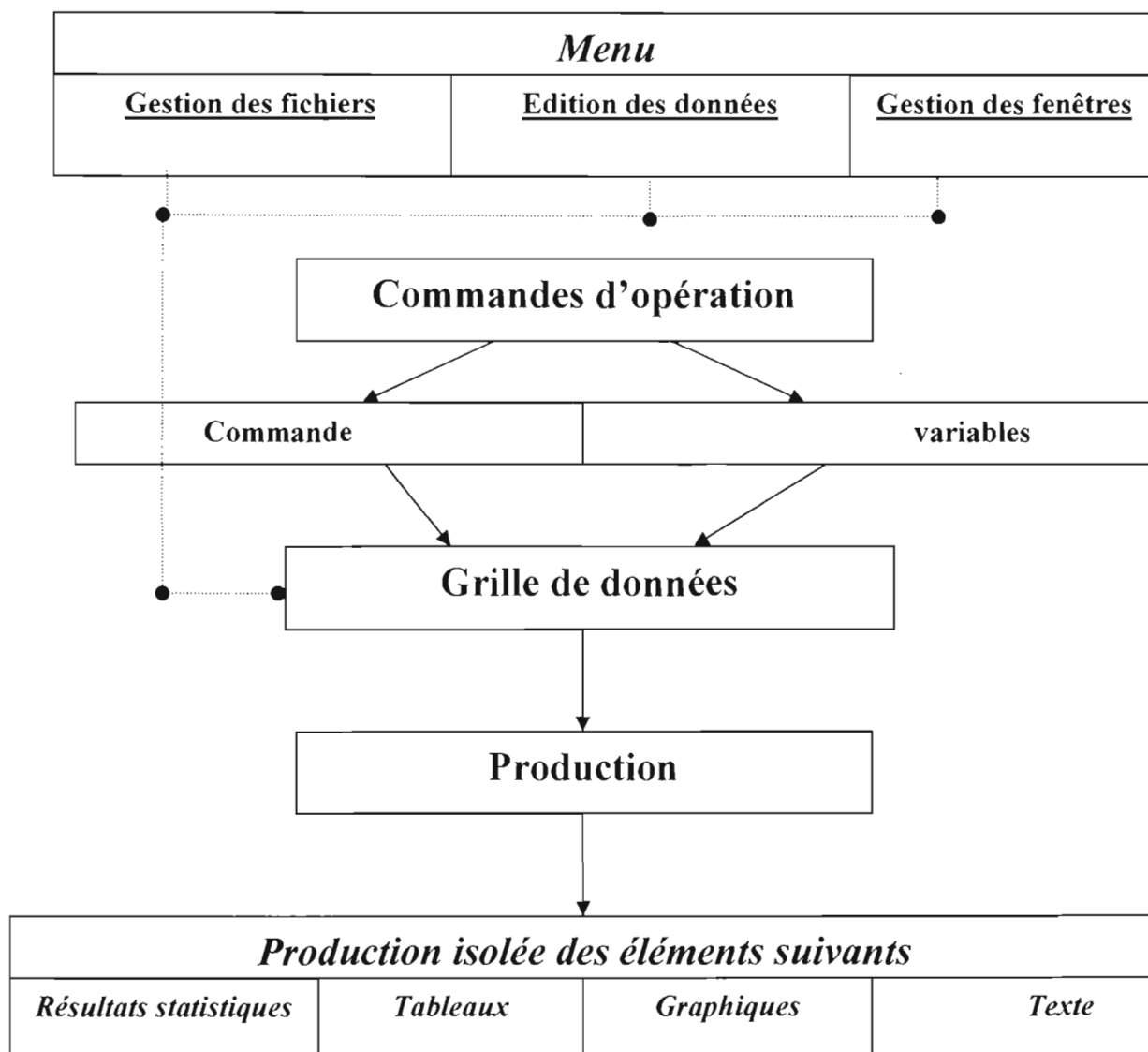
STATA existe sous deux formes le 'Intercooled STATA' et le 'Small STATA' ; Le 'Intercooled STATA' contient toutes les fonctionnalités du logiciel et à un champ d'application encore plus vaste alors que le 'Small STATA' contient certaines restrictions tant au niveau des fonctionnalités que de la mémoire adressable. Ce logiciel existe pour toutes les plates formes couramment rencontrées telles que Windows 9x/NT/2000, Macintosh (MachOs), et plusieurs distributions du système Unix (Digital Unix, Hp-UX, AIX, Linux, Solaris) . Les fichiers sont portables d'une plate forme à l'autre sans conversion, et l'interface est pratiquement identique sur les trois plates formes.

STATA peut être installé soit sur un ordinateur individuel ou portable, soit sur un serveur à l'aide d'un CD-ROM d'installation (double cliquer sur le fichier setup.exe dans le répertoire stata7) ou directement au travers de l'Internet (vous le trouverez sur le site <http://www.stata.com>). Au cours de l'installation, il nous sera demandé la clé qui est constituée d'un numéro de série, d'un code et d'un numéro autorisation. Installé, STATA 7 occupe environ 10 Mb d'espace disque.

STATA peut être démarré suivant deux modes : soit le mode graphique (à l'aide de la commande *xstata*) , soit le mode texte appelé stata console (à l'aide de la commande *stata*).

Si vous êtes sur un système d'exploitation Windows, l'icône STATA sera par défaut dans le menu programme de Démarrer.

1.3 Représentation schématique du logiciel Stata



Le menu est constitué de façon générale de commandes de gestions de fichiers, d'édition des données, et de gestion des fenêtres. Le traitement avec stata se faisant par des lignes de commandes), nous avons donc au deuxième niveau les commandes d'opération qui eux au cours de leur exécution vont agir sur les données pour produire des résultats qui peuvent être sous forme de résultats statistiques, tableaux, graphiques, ou textes.

II. Le vocabulaire de base

2.1 Le démarrage et l'arrêt

Pour le démarrage du logiciel Stata déjà installé, cliquer sur **Intercooled STATA**, exécutable se trouvant dans le dossier Stata du menu programme de **DEMARRER**.

Pour arrêter allez au niveau du menu file et cliquer sur exit ou taper la touche ctrl+F4 ou taper au niveau de la fenêtre de commande **exit, clear**. Après avoir exécuté une de ces commandes STATA se ferme avec suppression de toutes les données en mémoire.

2.2 Les quatre fenêtres

Stata dispose de quatre fenêtres qui sont : **Command, Results, Review, Variables**.

Fenêtre Command



Saisir la commande dans cette fenêtre et taper sur entrée.

Fenêtre Results

```

Stata Results
-----
Statistics/Data Analysis 7.0
Copyright 1989-2004
Stata Corporation
4905 Ashmead Drive
College Station, Texas 77845 USA
800-537-9000 http://www.stata.com
379-696-4200 http://www.stata.com
379-696-4200 (fax)

Single-user Stata for Windows (64-bit) license
Serial number: 197049693
Licensee: JSA
License: 48222F

Notes
-----
1. (maxsize option) 0.70 GB allocated to data

. use "C:\Enquête Prioritaire 98\data\stata\calculind\tauxalph98\fic1.dta", clear
. tabulate
. list realized
. describe

Contains data from C:\Enquête Prioritaire 98\data\stata\calculind\tauxalph98\fic1.dta
   obs:      50,721
   vars:      2
   size:     1,014,420 (0.00% of memory free)
   on:      6 Sep 2004 12:11

+-----+-----+-----+-----+-----+
| variable name | storage | display | value | variable label |
|-----|-----|-----|-----|-----|
| g5            | double | 00.00   | 0000   | Freqente en 97/98 |
| s9q2         | double | 01.00   | 0000   | Freqente en 97/98 |
+-----+-----+-----+-----+
Sorted by:  g5

```

Le résultat de l'exécution de la commande est affiché dans la fenêtre

Fenêtre Review

```

Review
-----
use "C:\Enquête Prioritaire 98\data\stata\calculind\tauxalph98\fic1.dta", clear
tabulate
describe
dir
ls
type
remove

```

Chaque commande exécutée est enregistrée dans cette fenêtre. Cliquer sur l'une des commandes pour l'afficher dans la fenêtre d'édition des commandes ; Double-cliquer sur une des commandes pour l'exécuter directement.

Fenêtre **Variables**



Cette fenêtre contient la liste des variables stockées en mémoire. Cliquer sur une variable pour la mettre dans la fenêtre de saisie des commandes.

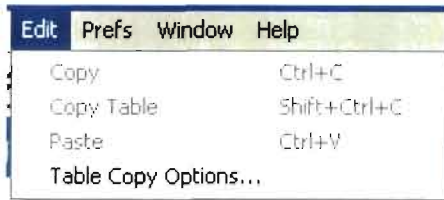
Ces 4 fenêtres sont redimensionnables indépendamment et pour sauvegarder les dimensions, cliquer sur le menu **Prefs, Save Windowing Preferences**.

2.3 La barre de menu

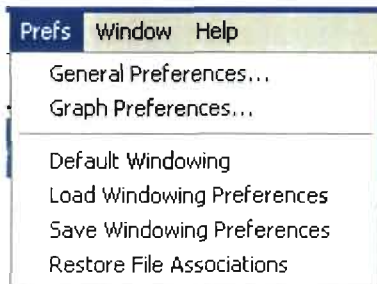
La barre de menu est constituée de cinq éléments qui sont : **File, Edit, Prefs, Windows, help**.



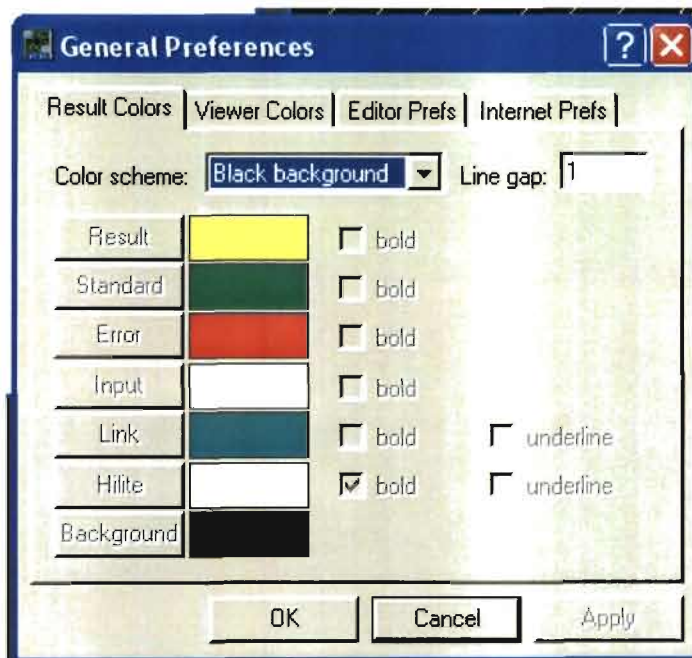
FILE : Le menu File présente les différentes opérations externes que l'on peut effectuer sur un fichier telles que ouvrir un fichier existant (*Open*), la sauvegarde (*save, save as, filename, save graph*), la création de fichier **Do** (type de macro stata) (*Do*), et l'impression de fichier (*Print Graph, Print Result*).



EDIT : Le menu Edit contient les fonctions qui permettent de copier et de coller.

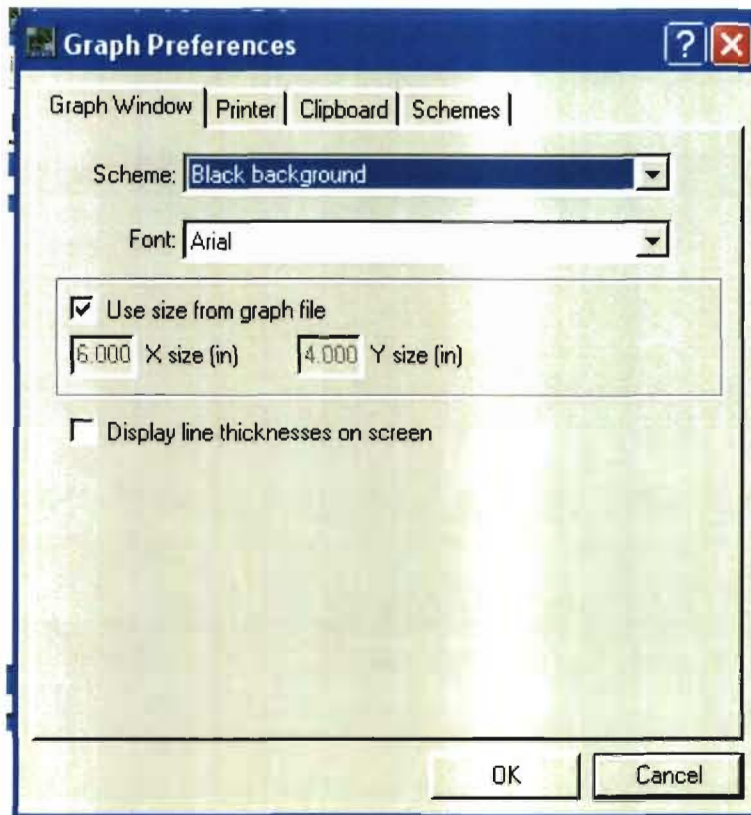


PREFS : Le menu Prefs contient les choix (préférences) de l'utilisateur en ce qui concerne les couleurs et les emplacements des fenêtres par défaut.



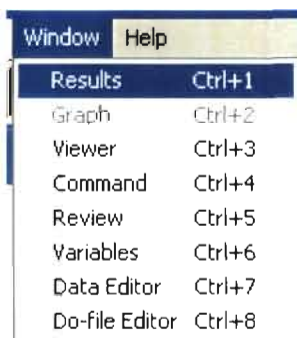
General Preferences :

Choix de la couleur de l'arrière plan de la fenêtre **Results**, personnalisation de la connexion à l'Internet,...



Graph Preferences

Choix de l'arrière plan de la fenêtre graphique,...



WINDOWS: Le menu Windows contient les fonctions d'affichage des fenêtres de saisie de commande (*command*), de résultats (*results*), de variables (*variables*), de saisie de données (*data editor*),...



HELP: Le menu help contient des fonctions de recherche de commandes ou de mot clé et d'affichage de l'aide correspondant à l'élément recherché ;recherche effectuer soit en local ou en utilisant les ressources de l'Internet ; il nous donne aussi les informations sur STATA.

2.4 La barre d'outils

| Symbole | Désignation |
|---|--|
|  | Open : Ouverture d'un fichier de donnée stata (<i>use</i>) |
|  | Save : Enregistrement des données en mémoire sur le disque dur (<i>save</i>) |
|  | Print : Impression d'un fichier log ou graph |
|  | Begin log : Ouvrir, suspendre ou fermer un fichier log (<i>log</i>) |
|  | Start viewer : Affichage de la fenêtre log à l'écran (si elle est ouverte) |
|  | Bring results windows to front : Affichage de la fenêtre résultat |
|  | Do-file editor : editeur de fichier do file |
|  | Data editor : éditeur de données (<i>edit</i>) |
|  | Browser the data in memory : navigateur (<i>browse</i>) |
|  | Bring graph. Windows to front : Affichage de la fenêtre du graph. (si cela a été crée) |
|  | Clear-more-condition : Faire défiler une autre page quand 'more' est exécutée |
|  | Break : arrêt d'une commande en exécution ou un do file |

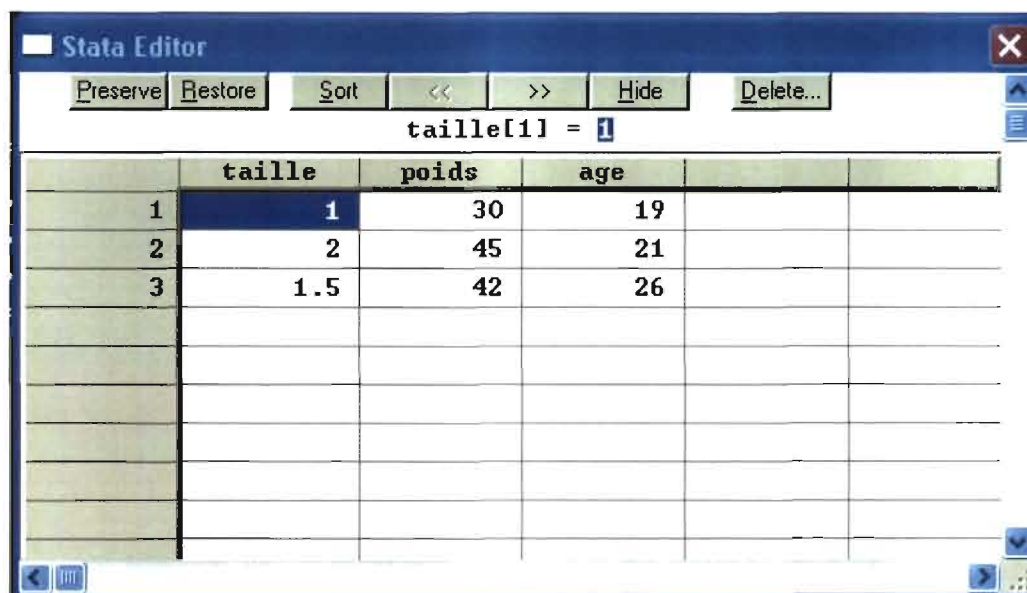
2.5 Les différents types de fichiers

Au cours de l'utilisation du logiciel STATA, nous pouvons être amenés à manipuler plusieurs types de fichiers nous avons entre autres :

- Les fichiers *.dta*

Ce sont des fichiers de données propres à Stata. La lecture de ces fichiers se fait à l'aide de la commande *use* et la sauvegarde des données du fichier avec la commande *save*.

Ex :



The screenshot shows the Stata Editor window with a data table. The table has three columns: 'taille', 'poids', and 'age'. The first three rows contain data, and the rest are empty. The first row is highlighted in blue.

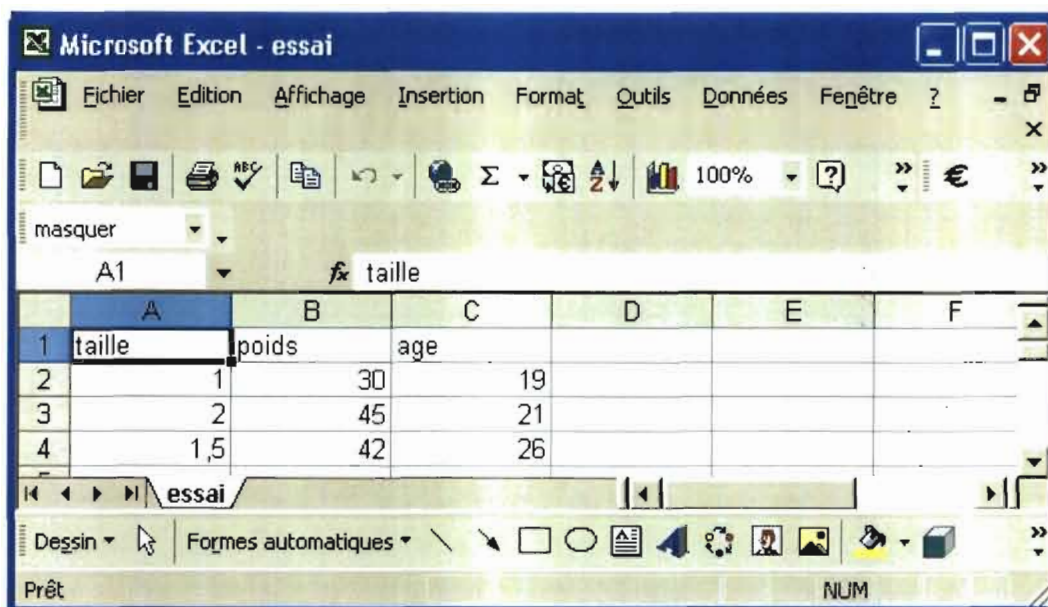
| | taille | poids | age |
|---|--------|-------|-----|
| 1 | 1 | 30 | 19 |
| 2 | 2 | 45 | 21 |
| 3 | 1.5 | 42 | 26 |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

- Les fichiers *.csv*

Ce sont des fichiers externes à STATA ; mais grâce à cette extension, STATA est à mesure de lire le contenu de ces fichiers de données. Ces fichiers sont généralement des fichiers Excel, Access, Dbase...

La lecture de ces fichiers se fait à l'aide de la commande *insheet*.

Ex.



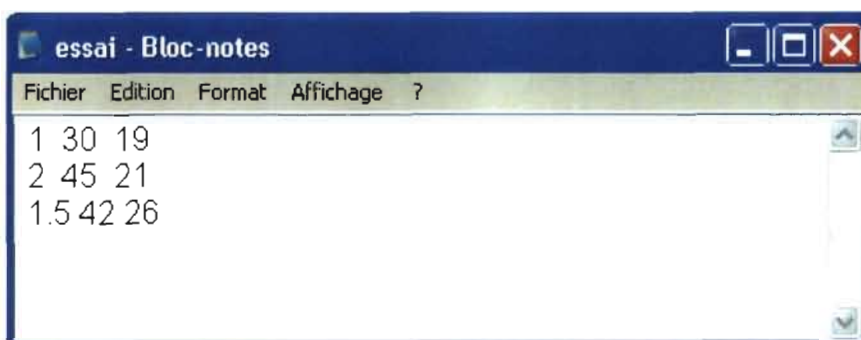
The screenshot shows a Microsoft Excel window titled "Microsoft Excel - essai". The spreadsheet has the following data:

| | A | B | C | D | E | F |
|---|--------|-------|-----|---|---|---|
| 1 | taille | poids | age | | | |
| 2 | 1 | 30 | 19 | | | |
| 3 | 2 | 45 | 21 | | | |
| 4 | 1,5 | 42 | 26 | | | |

- Les fichiers *.raw*

Ce sont des fichiers externes à STATA. Au format ascii, ils ne contiennent que les données manipulées. La lecture se fait grâce à la commande *infile*.

Ex :



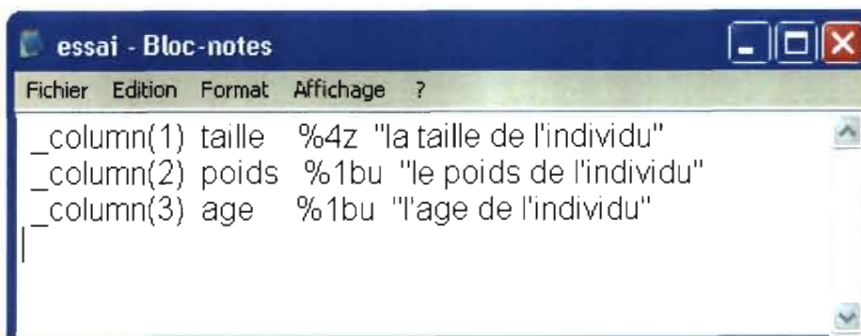
The screenshot shows a Notepad window titled "essai - Bloc-notes" containing the following text:

```
1 30 19
2 45 21
1.5 42 26
```

- les fichiers *.dct*

Ce sont des fichiers « dictionnaires » ; Appartenant à STATA, il facilite au système la lecture de fichier généralement les fichiers *.raw*. Ils donnent les informations sur les variables utilisées dans les fichiers *.raw* à savoir désignation de la variable, le format, la taille, et le label.

Ex :



The screenshot shows a Notepad window titled "essai - Bloc-notes" containing the following text:

```
_column(1) taille %4z "la taille de l'individu"
_column(2) poids %1bu "le poids de l'individu"
_column(3) age %1bu "l'age de l'individu"
```

- les fichiers **.smcl**

Ce sont des fichiers utilisés pour la réalisation de session de travail. (confère niveau 3.5)

- les fichiers **.do**

Ce sont des fichiers contenant un ensemble de commandes spécifiques pour la réalisation d'une tâche précise. (confère niveau 3.5)

III. La programmation avec le logiciel STATA

3.1 Syntaxe générale des commandes STATA

La manipulation des données se fait à l'aide de directives dont le format le plus rencontré est le suivant :

```
[by variable list :] command [variable list] [if condition] [in range]
      [using file name] [[weights]], [options]
```

Les crochets signifient que l'argument de la commande peut être omis ; mais les crochets autour de *weights* sont requis quand bien même la valeur est spécifiée.

Le premier *variable list* : représente la liste des variables groupées.

Le second *variable list* : représente la liste des variables à laquelle sera appliquée la commande

Range et *Condition* : permettent d'effectuer une sélection appropriée dans une liste de variable donnée.

Using file name : utilisé pour effectuer les opérations sur les fichiers (chargement en mémoire, importation et exportation) .

Weights : spécifie le poids qui sera appliqué au niveau des observations.

Options : spécifie ceux dont stata pourrait avoir besoin pour exécuter le programme.

3.2 La gestion et la manipulation des données

Lecture du fichier de données

Dans cette section, nous détaillerons la manipulation de trois commandes principaux, à savoir *use* (utilisée pour la lecture de fichier au format stata), *insheet* (utilisée pour la lecture de fichier au format CSV provenant généralement d'application différente telle que Excel), et *infile* (utilisée pour la lecture des fichiers 'plats' tel que le block note ; les données étant enregistrées au format .raw) .

Use

Lecture des fichiers au format stata

```
Use [filename]
```

Filename est un fichier stata ; c'est-à-dire enregistré sous la forme *filename.dta*

.dta est l'extension des fichiers stata.

Le fichier stata peut être lu aussi depuis un répertoire.

Ex : `use \home\data\filename.dta`

Insheet

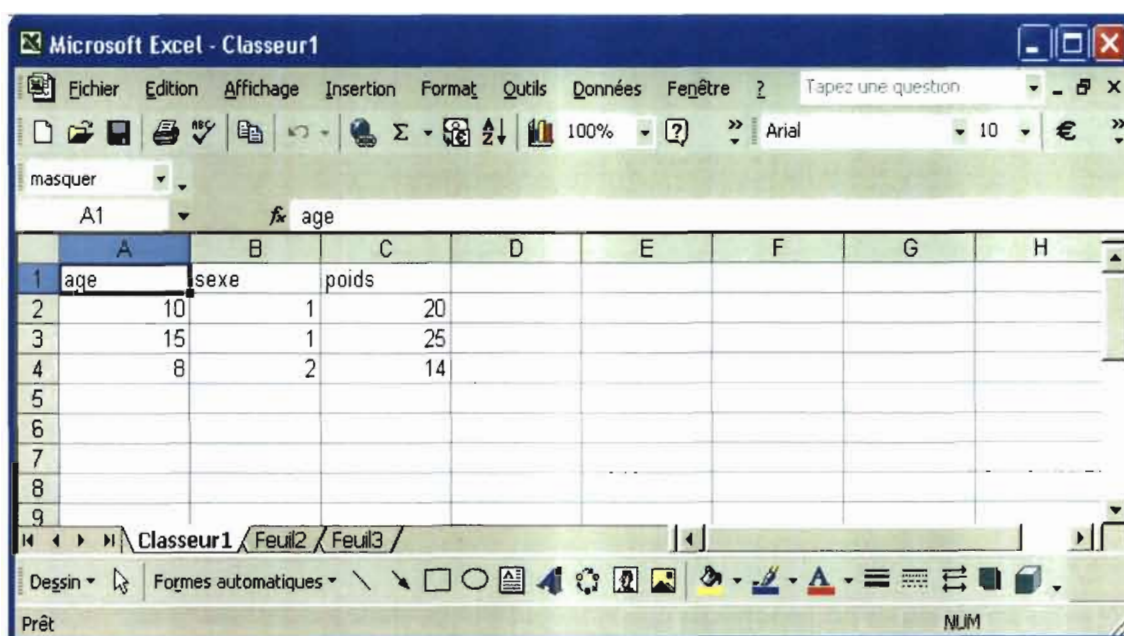
Lecture des fichiers au format .csv

Insheet using `[filename]`

Vous pouvez utiliser une feuille de calcul ou tout autre logiciel vous permettant d'enregistrer vos données sous format .csv

Insheet using `filename.csv`

Exemple avec Excel



Infile

Lecture des fichiers téléchargés via l'Internet ; généralement au format ASCII.

Infile `[variable list]` using `[filename]`

Vous pourrez enregistrer votre fichier sous format .raw

Infile `var1 var2 var3` using `mydata.raw`

Ex :

1 55 4.5

2 23 3.2
3 34 3.4
4 52 7.1

Sauvegarde du fichier de données

Pour enregistrer les fichiers au format stata, on utilisera la commande *save*

Save

Permet la sauvegarde des fichiers de données

Save [filename],[option]

Le fichier de sauvegarde *filename* sera automatiquement au format stata (.dta)

Save dispose d'une option qui est régulièrement employée à savoir *replace*.

Ex :

save mydata, replace

Après avoir sauvegardé le fichier encours sous le nom mydata, celui-ci est automatiquement remplacé par le nouveau fichier mydata.

Gestion de l'espace mémoire réservé aux données

L'espace mémoire réservé aux données stata étant limité, alors nous serons amener soit à compresser nos données ou à augmenter l'espace.

desc using [filename]

Donne les caractéristiques du fichier *filename* (nombre d'observations, de variables, et la taille).

desc using mydata.dta

Ex :

Contains data

Obs: 1,248

Vars:104

Size : 11,576,000

set memory *taille*

Effectue une allocation mémoire de valeur *taille* à la session courante.

Set memory 12m

compress

Cette commande effectue la compression des données en mémoire.

compress

Description des données

Concernant la description des données en mémoire les deux commandes couramment employées sont : **describe** et **list**.

describe

Cette commande donne les informations concernant les données en mémoire.

Ex. de description du fichier EP2FI11 de l'enquête prioritaire 98

Describe

```

Stata Results
-----
Statistics/Data Analysis 7.0 Copyright 1994-2001
State Corporation
3702 Lakeside Drive
College Station, Texas 77946 USA
(505) 843-7171 http://www.stata.com
(505) 843-7000 stata@stata.com
(505) 843-7001 (fax)

Single-user, single-seat Windows desktop license
Serial number 117042004
Licensee's full name
Licensee's full name

Notice:
1. License option 0.7% RW allocated to data

. use "C:\DATA\indices98\EP2FI11.dta"
no room to add more observations
r(9011)

. set memory 40000
40000

. use "C:\DATA\indices98\EP2FI11.dta", clear

. describe

Contains data from C:\DATA\indices98\EP2FI11.dta
      obs      296,730
      vars         6
      size    15,429,960 (62.2% of memory free)

-----
variable name      storage      display      value      variable label
                    type      format
-----
95                 double      16.00              N?ZD
98                 double      22.00              N?N?nage dans la ZD
94c0               double      22.00              1940
94c1               double      22.00              1941
94c2               double      22.00              1942
94c3               double      22.00              1943
-----
Sorted by

```

Affiche le nombre d'observations, de variables, la taille, la liste des variables, leur type, leur format et un commentaire concernant chacune d'elles.

list

Affiche à l'écran les données.

Ex. sur le fichier EP2FI11 de l'enquête prioritaire 98

List

Cette commande est utilisée pour renommer les variables.

Ex : **rename** *old_name new_name*

drop

Cette commande est utilisée pour suppression des variables.

Ex : **drop** *var1 var2*

keep

Cette commande est utilisée pour la conservation, l'enregistrement du contenu des variables.

Ex : **keep** *var3 var4*

Variable particulière

-Variable factice, variable n'ayant pas d'identification et dont la valeur ne peut être utilisée par le programmeur.

Ex : `tabulate taille, generate(variable_factice)`

La variable *variable_factice* sera créée et contiendra chaque résultats de la tabulation.

-Variable caractère, la variable qui sera créée sera du type caractère. Le nom de la variable sera précédé de la taille de la dite variable.

`generate str5 nom_variable = 'john'`

str signifie string et 5 représente le nombre de caractères maximum

-Variable date : stata stocke les dates au format mois, jour, année de type numérique grâce à la commande `mdy()`.

Ex : `generate mydate = mdy(mois,jour,année)`

3.4. La gestion des fichiers

Concernant la gestion des fichiers, nous verrons quatre (4) commandes à savoir **sort**, **append**, **merge**, **collapse**.

Sort

Cette commande permet d'effectuer le tri des données des différentes observations d'un fichier dans un ordre spécifique.

Ex : **sort** *nom_variable*

tri les données du fichier par rapport à *nom_variable*.

Append

Cette commande permet d'effectuer la jointure de deux fichiers ayant les mêmes variables ; autrement dit, elle permet d'augmenter les données du fichier principal.

Ex : *use fichier_primaire*

append using *fichier_secondaire*

Merge

Comme la précédente, elle permet la jointure de deux fichiers ; mais la particularité de merge est qu'elle joint en une seule prise des fichiers ayant des variables différentes par rapport à une variable identifiant existant dans les deux fichiers.

Ex : *use fichier_primaire*

sort id

merge id using *fichier_secondaire*

Collapse

Cette commande est utilisée pour la création des fichiers de données contenant des variables calculées comme la moyenne, la somme, etc.

Ex : *collapse* (mean) var1, var2

3.5 La création et l'utilisation des fichiers 'log' et 'do'

Fichiers 'log'

Les fichiers log d'extension *.smcl* contiennent les enregistrements des commandes saisies et des résultats obtenus durant la session stata.

Pour la création d'un fichier log on utilise :

Ex : *Log using filename*

filename représente un nom donné au fichier log

Pour ouvrir une session log, on utilise la même commande :

Ex : *Log using filename*

filename représente le nom du fichier log

Pour la fermeture d'une session log, on utilise la commande :

Ex : *Log close*

Sans paramètre, c'est la fermeture de la session log en cours.

Fichier 'do'

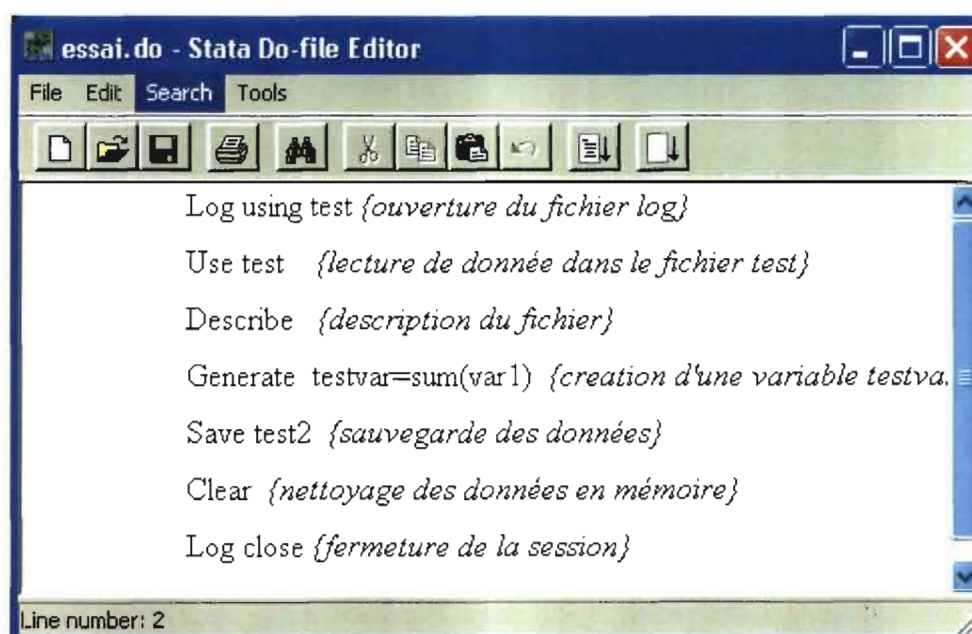
Les fichiers do d'extension *.do* contiennent un ensemble de commandes qui peuvent être saisies sur n'importe quel éditeur à condition d'être enregistrées au format *.do* ; Ceux-ci pourront être exécutés plus tard.

Pour exécuter un fichier do, faire :

Ex : **do** mydata

Exemple de fichier do avec un fichier log

Nom du fichier *essai.do*



3.6 Travailler en ligne avec STATA

Cette section contient les commandes classiques utilisées assez souvent pour les traitements.

who

Cette commande affiche l'identifiant des utilisateurs actuellement connectés sur stata.

Ex : **who**

pwd

Cette commande affiche le répertoire de travail.

Ex : **pwd**

ls

Cette commande affiche le contenu du répertoire de travail.

Ex : *ls*

mkdir

Cette commande permet la création de répertoire.

Ex : *mkdir nom_repertoire*

cd

Cette commande permet de changer de répertoire de travail.

Ex : *mkdir essai*

cd essai

summarize

Cette commande donne les renseignements concernant les variables en affichant dans un tableau le nombre d'observations, la moyenne, etc.

summarize list_variable

Il faut remarquer que Summarize sans *list_variable* affiche les informations concernant toutes les variables du fichier.

Ex : avec le fichier EP2FI11 de l'enquête prioritaire 98

Summarize



```

Stata Results
-----
Stata 7.0
Copyright 1994-2003
Stata Corporation
4905 Lakeshore Drive
College Station, Texas 77940 USA
300-STATA-PC (HTTP://WWW.STATA.COM)
300-226-4600 (STATA@STATA.COM)
300-226-4601 (Fax)

Single-user State for Windows perpetual license:
Serial number: 13704653
Licensee for: an244

Active:
) 1-pass option? 0.96 MB allocated to data

. use "C:\DATA\indices98\EP2FI11.dta"
no more to add more observations
(1901)

. set memory 40000
(40000)

. use "C:\DATA\indices98\EP2FI11.dta", clear

. summarize

+-----+-----+-----+-----+-----+
| Variable | Obs | Mean | Std. Dev. | Missing | Total |
+-----+-----+-----+-----+-----+
| v94-0 | 296730 | 212.9289 | 122.6829 | 1 | 425 |
| v94-1 | 296730 | 14.82826 | 13.68659 | 0 | 93 |
| v94-2 | 296730 | 18 | 10.09952 | 1 | 303 |
| v94-3 | 296730 | 1.91684 | 3.09447 | 1 | 9 |
| v94-4 | 174694 | 3237080 | 5.00e+08 | 0 | 9 |
| v94-5 | 247410 | 2.22868 | 1.670604 | 1 | 9 |
+-----+-----+-----+-----+-----+

```

tabulate

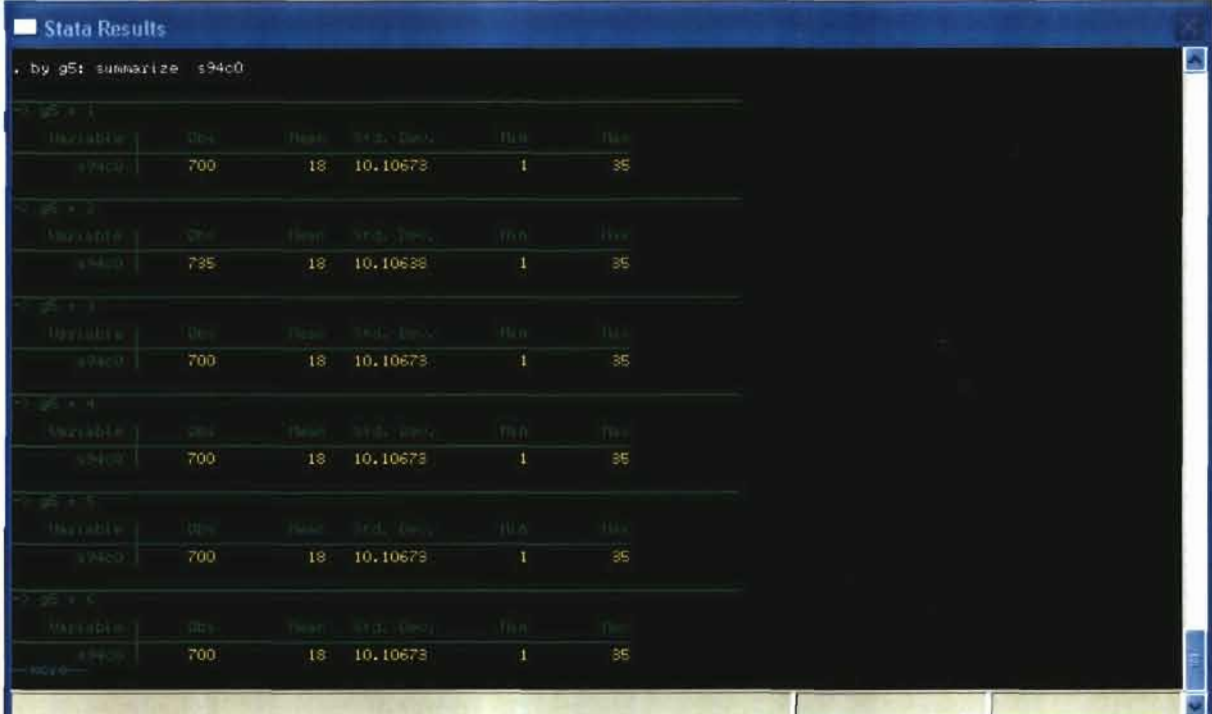
Cette commande prenant en argument une variable affiche la table des fréquences de la dite variable. Au cas où elle a deux arguments, nous aurons un tableau croisé.

tabulate list_variable

Ex : avec le fichier EP2FI11 de l'enquête prioritaire 98

tabulate g5

| g5 | Pct | Cum |
|-----|------|-------|
| 700 | 0.24 | 0.24 |
| 700 | 0.24 | 0.48 |
| 700 | 0.24 | 0.72 |
| 700 | 0.24 | 0.96 |
| 700 | 0.24 | 1.20 |
| 700 | 0.24 | 1.44 |
| 700 | 0.24 | 1.68 |
| 700 | 0.24 | 1.92 |
| 700 | 0.24 | 2.16 |
| 700 | 0.24 | 2.40 |
| 700 | 0.24 | 2.64 |
| 700 | 0.24 | 2.88 |
| 700 | 0.24 | 3.12 |
| 700 | 0.24 | 3.36 |
| 700 | 0.24 | 3.60 |
| 700 | 0.24 | 3.84 |
| 700 | 0.24 | 4.08 |
| 700 | 0.24 | 4.32 |
| 700 | 0.24 | 4.56 |
| 700 | 0.24 | 4.80 |
| 700 | 0.24 | 5.04 |
| 700 | 0.24 | 5.28 |
| 700 | 0.24 | 5.52 |
| 700 | 0.24 | 5.76 |
| 700 | 0.24 | 6.00 |
| 700 | 0.24 | 6.24 |
| 700 | 0.24 | 6.48 |
| 700 | 0.24 | 6.72 |
| 700 | 0.24 | 6.96 |
| 700 | 0.24 | 7.20 |
| 700 | 0.24 | 7.44 |
| 700 | 0.24 | 7.68 |
| 700 | 0.24 | 7.92 |
| 700 | 0.24 | 8.16 |
| 700 | 0.24 | 8.40 |
| 700 | 0.24 | 8.64 |
| 700 | 0.24 | 8.88 |
| 700 | 0.24 | 9.12 |
| 700 | 0.24 | 9.36 |
| 700 | 0.24 | 9.60 |
| 700 | 0.24 | 9.84 |
| 700 | 0.24 | 10.08 |
| 700 | 0.24 | 10.32 |
| 700 | 0.24 | 10.56 |
| 700 | 0.24 | 10.80 |
| 700 | 0.24 | 11.04 |
| 700 | 0.24 | 11.28 |
| 700 | 0.24 | 11.52 |
| 700 | 0.24 | 11.76 |
| 700 | 0.24 | 12.00 |
| 700 | 0.24 | 12.24 |
| 700 | 0.24 | 12.48 |
| 700 | 0.24 | 12.72 |
| 700 | 0.24 | 12.96 |
| 700 | 0.24 | 13.20 |
| 700 | 0.24 | 13.44 |
| 700 | 0.24 | 13.68 |
| 700 | 0.24 | 13.92 |
| 700 | 0.24 | 14.16 |
| 700 | 0.24 | 14.40 |
| 700 | 0.24 | 14.64 |
| 700 | 0.24 | 14.88 |
| 700 | 0.24 | 15.12 |
| 700 | 0.24 | 15.36 |
| 700 | 0.24 | 15.60 |
| 700 | 0.24 | 15.84 |
| 700 | 0.24 | 16.08 |
| 700 | 0.24 | 16.32 |
| 700 | 0.24 | 16.56 |
| 700 | 0.24 | 16.80 |
| 700 | 0.24 | 17.04 |
| 700 | 0.24 | 17.28 |
| 700 | 0.24 | 17.52 |
| 700 | 0.24 | 17.76 |
| 700 | 0.24 | 18.00 |
| 700 | 0.24 | 18.24 |
| 700 | 0.24 | 18.48 |
| 700 | 0.24 | 18.72 |
| 700 | 0.24 | 18.96 |
| 700 | 0.24 | 19.20 |
| 700 | 0.24 | 19.44 |
| 700 | 0.24 | 19.68 |
| 700 | 0.24 | 19.92 |
| 700 | 0.24 | 20.16 |
| 700 | 0.24 | 20.40 |
| 700 | 0.24 | 20.64 |
| 700 | 0.24 | 20.88 |
| 700 | 0.24 | 21.12 |
| 700 | 0.24 | 21.36 |
| 700 | 0.24 | 21.60 |
| 700 | 0.24 | 21.84 |
| 700 | 0.24 | 22.08 |
| 700 | 0.24 | 22.32 |
| 700 | 0.24 | 22.56 |
| 700 | 0.24 | 22.80 |
| 700 | 0.24 | 23.04 |
| 700 | 0.24 | 23.28 |
| 700 | 0.24 | 23.52 |
| 700 | 0.24 | 23.76 |
| 700 | 0.24 | 24.00 |
| 700 | 0.24 | 24.24 |
| 700 | 0.24 | 24.48 |
| 700 | 0.24 | 24.72 |
| 700 | 0.24 | 24.96 |
| 700 | 0.24 | 25.20 |
| 700 | 0.24 | 25.44 |
| 700 | 0.24 | 25.68 |
| 700 | 0.24 | 25.92 |
| 700 | 0.24 | 26.16 |
| 700 | 0.24 | 26.40 |
| 700 | 0.24 | 26.64 |
| 700 | 0.24 | 26.88 |
| 700 | 0.24 | 27.12 |
| 700 | 0.24 | 27.36 |
| 700 | 0.24 | 27.60 |
| 700 | 0.24 | 27.84 |
| 700 | 0.24 | 28.08 |
| 700 | 0.24 | 28.32 |
| 700 | 0.24 | 28.56 |
| 700 | 0.24 | 28.80 |
| 700 | 0.24 | 29.04 |
| 700 | 0.24 | 29.28 |
| 700 | 0.24 | 29.52 |
| 700 | 0.24 | 29.76 |
| 700 | 0.24 | 30.00 |
| 700 | 0.24 | 30.24 |
| 700 | 0.24 | 30.48 |
| 700 | 0.24 | 30.72 |
| 700 | 0.24 | 30.96 |
| 700 | 0.24 | 31.20 |
| 700 | 0.24 | 31.44 |
| 700 | 0.24 | 31.68 |
| 700 | 0.24 | 31.92 |
| 700 | 0.24 | 32.16 |
| 700 | 0.24 | 32.40 |
| 700 | 0.24 | 32.64 |
| 700 | 0.24 | 32.88 |
| 700 | 0.24 | 33.12 |
| 700 | 0.24 | 33.36 |
| 700 | 0.24 | 33.60 |
| 700 | 0.24 | 33.84 |
| 700 | 0.24 | 34.08 |
| 700 | 0.24 | 34.32 |
| 700 | 0.24 | 34.56 |
| 700 | 0.24 | 34.80 |
| 700 | 0.24 | 35.04 |
| 700 | 0.24 | 35.28 |
| 700 | 0.24 | 35.52 |
| 700 | 0.24 | 35.76 |
| 700 | 0.24 | 36.00 |
| 700 | 0.24 | 36.24 |
| 700 | 0.24 | 36.48 |
| 700 | 0.24 | 36.72 |
| 700 | 0.24 | 36.96 |
| 700 | 0.24 | 37.20 |
| 700 | 0.24 | 37.44 |
| 700 | 0.24 | 37.68 |
| 700 | 0.24 | 37.92 |
| 700 | 0.24 | 38.16 |
| 700 | 0.24 | 38.40 |
| 700 | 0.24 | 38.64 |
| 700 | 0.24 | 38.88 |
| 700 | 0.24 | 39.12 |
| 700 | 0.24 | 39.36 |
| 700 | 0.24 | 39.60 |
| 700 | 0.24 | 39.84 |
| 700 | 0.24 | 40.08 |
| 700 | 0.24 | 40.32 |
| 700 | 0.24 | 40.56 |
| 700 | 0.24 | 40.80 |
| 700 | 0.24 | 41.04 |
| 700 | 0.24 | 41.28 |
| 700 | 0.24 | 41.52 |
| 700 | 0.24 | 41.76 |
| 700 | 0.24 | 42.00 |
| 700 | 0.24 | 42.24 |
| 700 | 0.24 | 42.48 |
| 700 | 0.24 | 42.72 |
| 700 | 0.24 | 42.96 |
| 700 | 0.24 | 43.20 |
| 700 | 0.24 | 43.44 |
| 700 | 0.24 | 43.68 |
| 700 | 0.24 | 43.92 |
| 700 | 0.24 | 44.16 |
| 700 | 0.24 | 44.40 |
| 700 | 0.24 | 44.64 |
| 700 | 0.24 | 44.88 |
| 700 | 0.24 | 45.12 |
| 700 | 0.24 | 45.36 |
| 700 | 0.24 | 45.60 |
| 700 | 0.24 | 45.84 |
| 700 | 0.24 | 46.08 |
| 700 | 0.24 | 46.32 |
| 700 | 0.24 | 46.56 |
| 700 | 0.24 | 46.80 |
| 700 | 0.24 | 47.04 |
| 700 | 0.24 | 47.28 |
| 700 | 0.24 | 47.52 |
| 700 | 0.24 | 47.76 |
| 700 | 0.24 | 48.00 |
| 700 | 0.24 | 48.24 |
| 700 | 0.24 | 48.48 |
| 700 | 0.24 | 48.72 |
| 700 | 0.24 | 48.96 |
| 700 | 0.24 | 49.20 |
| 700 | 0.24 | 49.44 |
| 700 | 0.24 | 49.68 |
| 700 | 0.24 | 49.92 |
| 700 | 0.24 | 50.16 |
| 700 | 0.24 | 50.40 |
| 700 | 0.24 | 50.64 |
| 700 | 0.24 | 50.88 |
| 700 | 0.24 | 51.12 |
| 700 | 0.24 | 51.36 |
| 700 | 0.24 | 51.60 |
| 700 | 0.24 | 51.84 |
| 700 | 0.24 | 52.08 |
| 700 | 0.24 | 52.32 |
| 700 | 0.24 | 52.56 |
| 700 | 0.24 | 52.80 |
| 700 | 0.24 | 53.04 |
| 700 | 0.24 | 53.28 |
| 700 | 0.24 | 53.52 |
| 700 | 0.24 | 53.76 |
| 700 | 0.24 | 54.00 |
| 700 | 0.24 | 54.24 |
| 700 | 0.24 | 54.48 |
| 700 | 0.24 | 54.72 |
| 700 | 0.24 | 54.96 |
| 700 | 0.24 | 55.20 |
| 700 | 0.24 | 55.44 |
| 700 | 0.24 | 55.68 |
| 700 | 0.24 | 55.92 |
| 700 | 0.24 | 56.16 |
| 700 | 0.24 | 56.40 |
| 700 | 0.24 | 56.64 |
| 700 | 0.24 | 56.88 |
| 700 | 0.24 | 57.12 |
| 700 | 0.24 | 57.36 |
| 700 | 0.24 | 57.60 |
| 700 | 0.24 | 57.84 |
| 700 | 0.24 | 58.08 |
| 700 | 0.24 | 58.32 |
| 700 | 0.24 | 58.56 |
| 700 | 0.24 | 58.80 |
| 700 | 0.24 | 59.04 |
| 700 | 0.24 | 59.28 |
| 700 | 0.24 | 59.52 |
| 700 | 0.24 | 59.76 |
| 700 | 0.24 | 60.00 |
| 700 | 0.24 | 60.24 |
| 700 | 0.24 | 60.48 |
| 700 | 0.24 | 60.72 |
| 700 | 0.24 | 60.96 |
| 700 | 0.24 | 61.20 |
| 700 | 0.24 | 61.44 |
| 700 | 0.24 | 61.68 |
| 700 | 0.24 | 61.92 |
| 700 | 0.24 | 62.16 |
| 700 | 0.24 | 62.40 |
| 700 | 0.24 | 62.64 |
| 700 | 0.24 | 62.88 |
| 700 | 0.24 | 63.12 |
| 700 | 0.24 | 63.36 |
| 700 | 0.24 | 63.60 |
| 700 | 0.24 | 63.84 |
| 700 | 0.24 | 64.08 |
| 700 | 0.24 | 64.32 |
| 700 | 0.24 | 64.56 |
| 700 | 0.24 | 64.80 |
| 700 | 0.24 | 65.04 |
| 700 | 0.24 | 65.28 |
| 700 | 0.24 | 65.52 |
| 700 | 0.24 | 65.76 |
| 700 | 0.24 | 66.00 |
| 700 | 0.24 | 66.24 |
| 700 | 0.24 | 66.48 |
| 700 | 0.24 | 66.72 |
| 700 | 0.24 | 66.96 |
| 700 | 0.24 | 67.20 |
| 700 | 0.24 | 67.44 |
| 700 | 0.24 | 67.68 |
| 700 | 0.24 | 67.92 |
| 700 | 0.24 | 68.16 |
| 700 | 0.24 | 68.40 |
| 700 | 0.24 | 68.64 |
| 700 | 0.24 | 68.88 |
| 700 | 0.24 | 69.12 |
| 700 | 0.24 | 69.36 |
| 700 | 0.24 | 69.60 |
| 700 | 0.24 | 69.84 |
| 700 | 0.24 | 70.08 |
| 700 | 0.24 | 70.32 |
| 700 | 0.24 | 70.56 |
| 700 | 0.24 | 70.80 |
| 700 | 0.24 | 71.04 |
| 700 | 0.24 | 71.28 |
| 700 | 0.24 | 71.52 |
| 700 | 0.24 | 71.76 |
| 700 | 0.24 | 72.00 |
| 700 | 0.24 | 72.24 |
| 700 | 0.24 | 72.48 |
| 700 | 0.24 | 72.72 |
| 700 | 0.24 | 72.96 |
| 700 | 0.24 | 73.20 |
| 700 | 0.24 | 73.44 |
| 700 | 0.24 | 73.68 |
| 700 | 0.24 | 73.92 |
| 700 | 0.24 | 74.16 |
| 700 | 0.24 | 74.40 |
| 700 | 0.24 | 74.64 |
| 700 | 0.24 | 74.88 |
| 700 | 0.24 | 75.12 |
| 700 | 0.24 | 75.36 |
| 700 | 0.24 | 75.60 |
| 700 | 0.24 | 75.84 |
| 700 | 0.24 | 76.08 |
| 700 | 0.24 | 76.32 |
| 700 | 0.24 | 76.56 |
| 700 | 0.24 | 76.80 |



```

Stata Results
. by g5: summarize s94c0

+-----+
|> g5 = 1|
+-----+
| Variable | Obs | Mean | Std. Dev. | Min | Max |
+-----+
| s94c0    | 700 | 18    | 10.10673  | 1   | 35  |
+-----+
|> g5 = 2|
+-----+
| Variable | Obs | Mean | Std. Dev. | Min | Max |
+-----+
| s94c0    | 735 | 18    | 10.10688  | 1   | 35  |
+-----+
|> g5 = 3|
+-----+
| Variable | Obs | Mean | Std. Dev. | Min | Max |
+-----+
| s94c0    | 700 | 18    | 10.10673  | 1   | 35  |
+-----+
|> g5 = 4|
+-----+
| Variable | Obs | Mean | Std. Dev. | Min | Max |
+-----+
| s94c0    | 700 | 18    | 10.10673  | 1   | 35  |
+-----+
|> g5 = 5|
+-----+
| Variable | Obs | Mean | Std. Dev. | Min | Max |
+-----+
| s94c0    | 700 | 18    | 10.10673  | 1   | 35  |
+-----+
|> g5 = 6|
+-----+
| Variable | Obs | Mean | Std. Dev. | Min | Max |
+-----+
| s94c0    | 700 | 18    | 10.10673  | 1   | 35  |
+-----+

```

3.7 Ecrire des programmes avec STATA

La création et la manipulation de macro

Les macros peuvent être vu comme étant des spécifications permettant de paramétrer des instructions.

Stata dispose de cinq commandes pour la définition de macro locales, à savoir : **local**, **global**, **tempvar**, **tempname**, **tempfile**.

Local

Cette commande correspond à la définition d'une macro local ;c'est à dire qui ne peut être utilisé qu'à l'intérieur d'une procédure ou fonction.

La syntaxe est la suivant : **local name [=exp | :ext_fct | [''[string]''[']]**.

name est le nom de la macro local

exp est une expression numérique

ext_fct représente le type, le format ou le label de la macro définie

string représente une valeur initiale de la macro

Ex : **local** nbre=1

global

Cette commande correspond à la définition d'une macro global ;c'est à dire qui ne peut être utilisé dans tous le programme.

La syntaxe est la suivant : **global** name [=exp | :ext_fct | ['][string][']].

name est le nom de la macro local

exp est une expression numérique

ext_fct représente le type, le format ou le label de la macro définie

string représente une valeur initiale de la macro

Ex : **global** compteur=2

Tempvar

Cette commande définit des variables correspondant à des noms de macro local ;Elles sont utilisées de façon temporaire dans le fichier de données

La syntaxe est la suivante : **tempvar** localname

Ex : **tempvar** nbre

Tempname

Cette commande définit des noms correspondant à des noms de macro local ;ils sont utilisés comme des scalaires temporaires ou des matrices. Ils sont automatiquement détruits après l'exécution du programme ou du do-files.

La syntaxe est la suivante : **tempname** localname

Ex : **tempvar** nbre

Tempfile

Cette commande définit des fichiers de données. Le nom de ces fichiers correspond aux noms de macro local et ces fichiers contiennent des données correspondant à la macro local identifié; ils sont utilisés de façon temporaire.

La syntaxe est la suivante : **tempfile** localname

Ex : **tempfile** essai

Notions de branchement conditionnel et de boucle

Commande If

Le 'if' commande de programmation en stata permet d'évaluer une expression. Si le résultat est vrai c'est à dire différent de 0, alors les commandes suivant le **if** sont exécutées ; Sinon c'est ceux suivant le 'else' qui sont exécutées.

```
Syntaxe : if expression {
                Commands
            }
            else command
```

Commande forvalues

La commande *forvalues* répète la macro local définie suivant la condition et exécute les instructions entre parenthèses.

```
Syntaxe : forvalues localname=range {
                Commands_referant_localname
            }
```

avec range étant un intervalle de nombres entiers. La boucle est effectuée au moins 0 fois.

Ex : **forvalues** i=1/100 {generate x'i'=uniform()}

Création de 100 variables ayant des contenus variés type x1, x2, ..., x100

3.8. Quelques cas pratiques

Exemples pris sur les Enquêtes Prioritaires 94 & 98

- **Calcul du taux d'alphabétisation (homme, femme, général)**

```
clear
drop _all /*suppression de toutes les variables existantes en mémoire*/
use EP2FI1 /*utilisation du fichier EP2FI1*/
sort g5 /*tri du fichier par rapport à g5*/
save EP2FI1, replace /*sauvegarde des données*/
use EP2FI4/*suppression des données de EP2FI1 de la mémoire et usage de EP2FI4*/
sort g5 /*tri par rapport à g5*/
save EP2FI4, replace /*sauvegarde des données*/
/*-----*/
use EP2FI1
merge g5 using EP2FI4 /*jointure des fichiers par g5*/
save fichjoint
gen Ad = 1 if s1q10 > 15 /*determination des adultes*/
gen AA = 1 if s1q10 > 15 & s4q2~=0 /*determination des adultes alphabétisés*/
gen HA = 1 if s1q6==1 & s4q2~=0 & Ad==1 /*determination des adultes hommes alph*/
gen FA = 1 if s1q6==2 & s4q2~=0 & Ad==1 /*determination des adultes femesm alph*/
gen NH=1 if s1q6==1 & Ad==1 /*nombre homme adultes*/
gen NF=1 if s1q6==2 & Ad==1 /*nbre femme adultes*/
/*-----*/
gen NNH=sum(NH) /*somme des hom*/
```

```

gen NNF=sum(NF) /*somme des femmes*/
gen NHA=sum(HA) /*somme des hommes alphabétisés*/
gen NFA=sum(FA)/*somme des femmes alphabétisées*/
gen NA=sum(Ad)/*somme des adultes*/
gen NAA=sum(AA)/*somme des adultes alphabétisés*/
/*-----*/
gen TauxAlph=(NAA[_N]/NA[_N])/*taux adultes alphabétisés*/
gen TauxAlphHom=(NHA[_N]/NNH[_N])/*taux hommes alphabétisés*/
gen TauxAlphFem=(NFA[_N]/NNF[_N])/*taux femmes alphabétisés*/
/*-----*/
summarize /* Résumé des différents calculs */
/*-----*/
exit

```

○ Détermination de l'indice de Gini

```

clear
drop _all
use EP2FI11
display "                TRACE DE LA COURBE DE LORENTZ"
gen effrev = 1
collapse (count) effrev, by(s94c2) /*Détermination des effectifs de chaque de revenu*/
list in 1/100 /*Affichage des 100 premiers éléments*/
gen someffrev=sum(effrev) /*Détermination du total des effectifs de chaque revenu*/
gen freqrev=effrev/someffrev[_N] /*Détermination de la fréquence des revenus*/
gen cumfreqrev=sum(freqrev) /*Détermination de la fréquence cumulée*/
list in 1/100
gen massesal=effrev[_n]*s94c2[_n] /*Détermination de la masse salariale*/
gen somms=sum(massesal) /*Montant de la masse salariale*/
gen freqms=massesal/somms[_N] /*Détermination de la fréquence de la masse salariale*/
gen cumfreqms=sum(freqms) /*détermination de la fréquence cumulée*/
list in 1/100
summarize
graph cumfreqms cumfreqrev cumfreqms /*réalisation du graphique*/
display "                COEFFICIENT DE GINI"
gen sominter=sum(freqrev[_n-1]*(cumfreqms[_n]+cumfreqms[_n-1]))
gen coefgini=1-sominter[_N]
list in 1/100
summarize
drop effrev someffrev freqrev cumfreqrev massesal somms freqms cumfreqms sominter
/*suppression des variables de travail*/
list in 1/100
summarize/*tableau du résumé*/
/*-----*/
exit

```


3.9 La réalisation de graphiques

Plusieurs types de graphiques peuvent être réalisés avec Stata à savoir :

- les diagrammes à bâton (ou bar charts ou bar)
- les diagrammes circulaires (ou pie charts ou pie)
- les diagrammes de boîtes (ou box-and-whiskers plots ou box)
- les graphiques en formes de matrices (ou twoway scatterplots matrix ou matrix)
- les graphiques d'analyse de variance à un critère de classification (ou oneway scatterplots ou oneway)
- les graphiques d'analyse de variance à deux critères de classification (ou twoway scatterplots ou twoway)
- les histogrammes (ou histograms)
- les graphiques schéma en étoile (ou star charts ou star)

Regardons maintenant quelques représentations :

- les diagrammes à bâton ou bar

Ces diagrammes servent à représenter les chiffres statistiques à l'aide de bâtonnets.

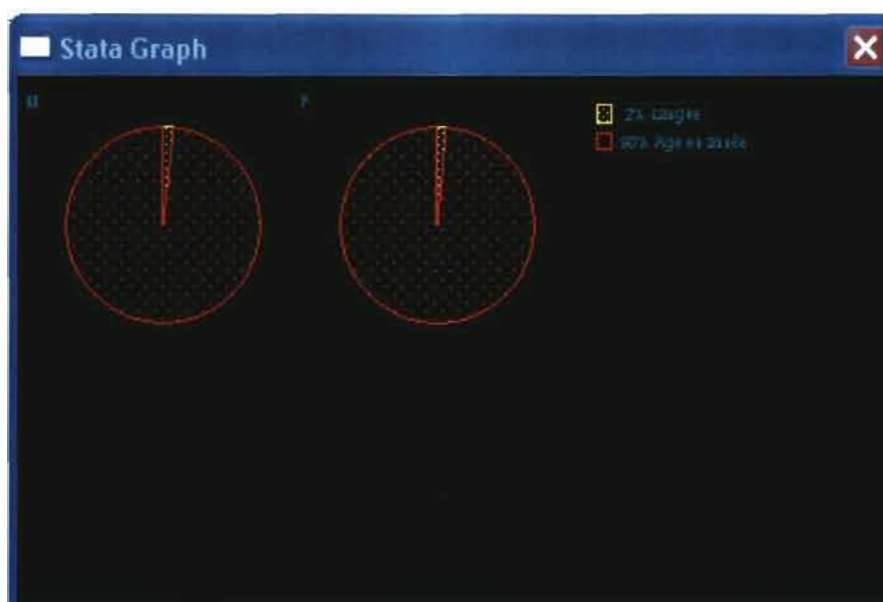
Ex : `graph s4q2 s1q10, bar by(s1q6)`

s4q2 : langue s1q10 : âge en année s1q6 : sexe

- les diagrammes circulaires ou pie

Ceux-ci servent à représenter les données statistiques à l'aide de forme cylindrique.

Ex : `graph s4q2 s1q10, pie by(s1q6)`



- les diagrammes de boîtes ou box

Ces diagrammes servent à représenter les données statistiques à l'aide boîte à piston.

Ex : `graph s4q2 s1q10, box(s1q6)`



IV. Bibliographie

4.1 les livres et articles

- Référence manuel release 4 (volume 1-2-3)
- Review of stata7 de stanislav Kolenikov
- Brief Guide of stata commands de Eric M. Uslaner
- Etc.

4.2 les sites Internet

- www.timberlake.co.uk/software/stata/Stata.html
- www.cpc.unc.edu/services/computer/presentations/statatutorial.html
- www.princeton.edu/_erp/stata/stata_woframes.htm
- www.dss.princeton.edu/online_help/stats_packages/stata/stata.htm
- www.princeton.edu/_erp/stata/main.html
- www.essex.ac.uk/economics/talif-stata/Navigation.htm
- www.stata.com
- etc.

Conclusion

Après avoir effectué une analyse fonctionnelle du traitement des données statistique, cette phase d'activité constituait la dernière partie de notre stage de fin d'étude. Elle a été bénéfique pour nous à différents niveaux. C'était une phase d'apprentissage des logiciels SPSS et STATA. La manipulation des bases de données à l'aide de ces logiciels nous a fait comprendre que la gestion des bases de données est un exercice bien délicat et complexe exigeant une excellente organisation surtout dans le cadre d'un travail d'équipe, mettant en jeu le PNUD et l'INSD.

En plus des connaissances acquises durant ce stage sur le plan du traitement informatique des données, nous avons pu relever la grande différence qu'il y a entre le traitement des projets cas d'école et les besoins des utilisateurs en entreprise. D'où la nécessité d'une capacité d'adaptation peu importe les situations.