

Utilisation du Logiciel de statistique SPSS 8.0

1 Introduction

Etude épidémiologique transversale en population générale dans 4 pays d'Afrique pour comprendre les différences de prévalence du VIH.

2000 personnes dans chaque ville.

Questionnaire, urine, serum et prélèvement vaginal.

2 Installation de SPSS 8.0 (Pour ceux qui possèdent un PC avec un lecteur de CD)

Mettre le CD-ROM

Cliquer sur Install SPSS 8.0

...

Numéro de série 2734031

License code: 30671 14980 77488 31093 73614 2222

3 Lancement du programme

>Démarrer >Programmes >SPSS 8.0 for Windows

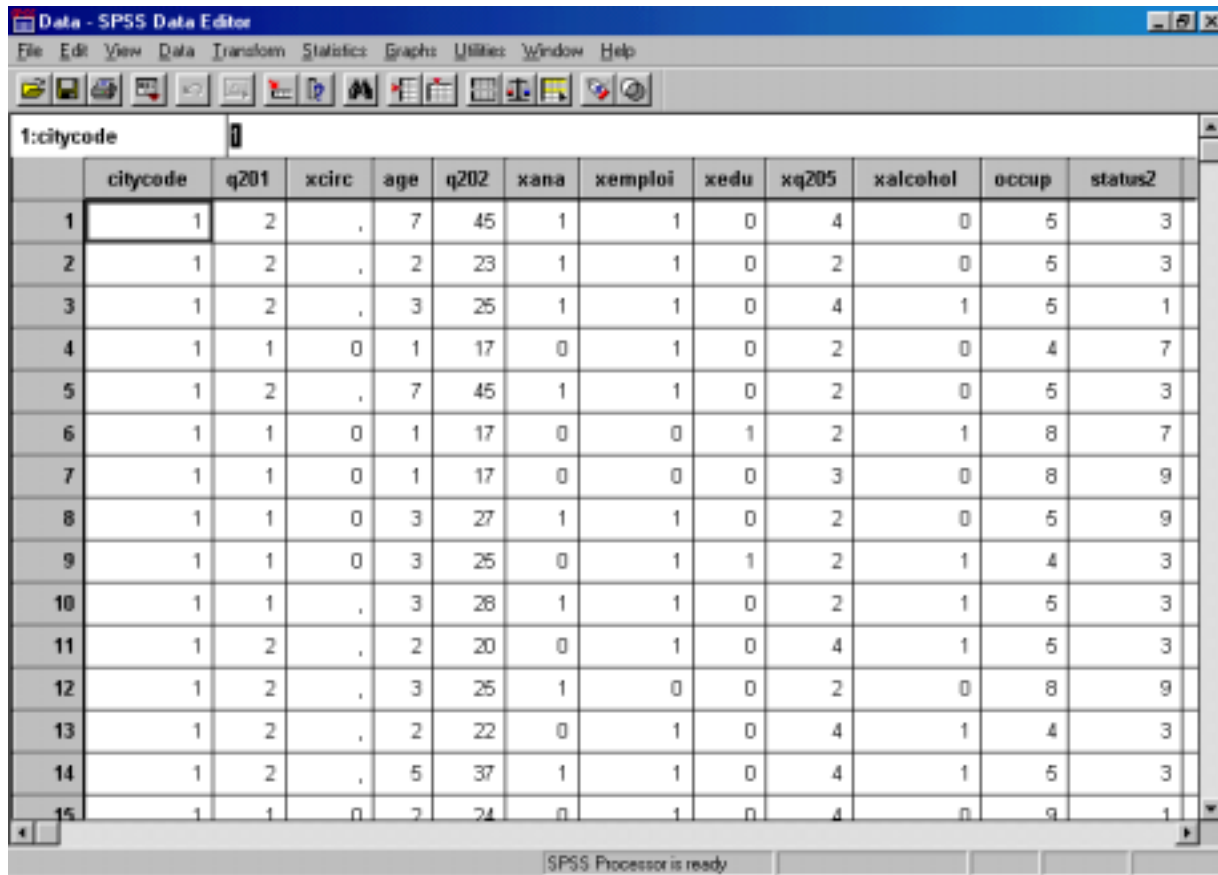
What would you like to do > Cancel

4 Ouvrir un fichier de données

File>Open>E:\data.sav (pour ceux qui ont le Cdrom)
File>Open>A:\data.sav (pour ceux qui ont la disquette)

Ecran

Se déplacer dans les données: la souris (7821 lignes)
valeurs manquantes: ‘.’



The screenshot shows the SPSS Data Editor window with a data table. The table has 15 rows and 13 columns. The columns are labeled: citycode, q201, xcirc, age, q202, xana, xemploi, xedu, xq205, xalcohol, occup, and status2. The data is as follows:

	citycode	q201	xcirc	age	q202	xana	xemploi	xedu	xq205	xalcohol	occup	status2
1	1	2	.	7	45	1	1	0	4	0	5	3
2	1	2	.	2	23	1	1	0	2	0	5	3
3	1	2	.	3	25	1	1	0	4	1	5	1
4	1	1	0	1	17	0	1	0	2	0	4	7
5	1	2	.	7	45	1	1	0	2	0	5	3
6	1	1	0	1	17	0	0	1	2	1	8	7
7	1	1	0	1	17	0	0	0	3	0	8	9
8	1	1	0	3	27	1	1	0	2	0	5	9
9	1	1	0	3	25	0	1	1	2	1	4	3
10	1	1	.	3	28	1	1	0	2	1	5	3
11	1	2	.	2	20	0	1	0	4	1	5	3
12	1	2	.	3	25	1	0	0	2	0	8	9
13	1	2	.	2	22	0	1	0	4	1	4	3
14	1	2	.	5	37	1	1	0	4	1	5	3
15	1	1	0	2	24	0	1	0	4	0	9	1

5 Statistique descriptive

a Répartition de Status2

Statistics> Summarize> Frequencies... Variable Status2... OK

Ecran

Statistics

STATUS2		
N	Valid	7818
	Missing	3

STATUS2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Job (full time)	1520	19,4	19,4	19,4
	Self employed	1929	24,7	24,7	44,1
	Homemaker	1205	15,4	15,4	59,5
	Student	1492	19,1	19,1	78,6
	Other	1672	21,4	21,4	100,0
	Total	7818	100,0	100,0	
Missing	System	3	,0		
	Total	7821	100,0		

SPSS Processor is ready

b Remarque

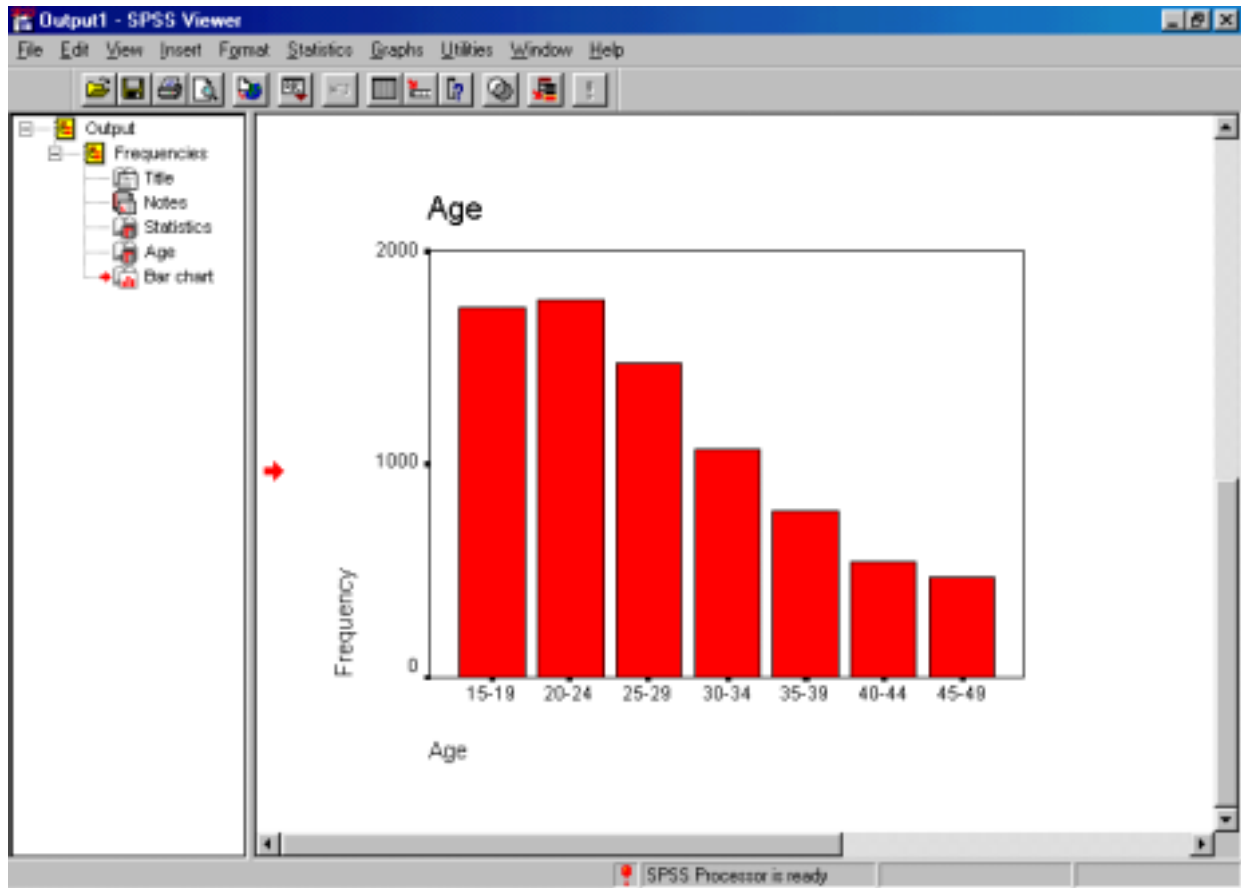
Double-clic dans le fenêtre Statistics
Simple clic sur Range puis clic à droite > What is this?

c Navigation entre la fenêtre Data et Output1

d Répartition par age (en classe: variable age)

Statistics> Summarize> Frequencies... Variable : Age... Charts> Bar charts ...OK

Ecran



e Description de l'age (en années: variable q202)

Statistics> Summarize> Frequencies... Variable : Q202

Décocher la case 'Display frequency table'

Statistics...

Cocher: Quartiles, Mean, Median, Mode, Std deviation, Variance, Minimum, maximum, OK

Ecran

The screenshot shows the SPSS Output1 - SPSS Viewer window. The 'Frequencies' section is expanded, showing a table of statistics for variable Q202. The table includes counts for valid and missing cases, and various statistical measures such as Mean, Median, Mode, Std. Deviation, Variance, Range, Minimum, Maximum, and Percentiles (25, 50, 75).

Statistics		
Q202		
N	Valid	7820
	Missing	1
Mean		27,45
Median		26,00
Mode		20
Std. Deviation		8,95
Variance		80,12
Range		34
Minimum		15
Maximum		49
Percentiles	25	20,00
	50	26,00
	75	34,00

f Lien entre HIV et sexe

Statistics > Summarize > Crosstabs...

Row(s) : Sex

Column(s) : HIV

Cells : Percentages: Row > Continue

Statistics : Chi-square > Continue

OK

Ecran

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Sex * HIV	6893	88,1%	928	11,9%	7821	100,0%

Sex * HIV Crosstabulation

Count

		HIV		Total
		Négatif	Positif	
Sex	Men	2729	335	3064
	Women	3157	672	3829
Total		5886	1007	6893

4 items selected (1 hidden/collapsed) | SPSS Processor is ready

Output1 - SPSS Viewer

File Edit View Insert Format Statistics Graphs Utilities Window Help

Output

- Crosstabs
 - Title
 - Notes
 - Case Process
 - Sex * Gono. d
 - Chi-Square T

Sex * Gono. Crosstabulation

			Gono.		Total
			Negatif	Positif	
Sex	Men	Count	3002	28	3030
		% within Sex	99,1%	,9%	100,0%
	Women	Count	3693	65	3758
		% within Sex	98,3%	1,7%	100,0%
Total		Count	6695	93	6788
		% within Sex	98,6%	1,4%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8,056 ^a	1	,005		
Continuity Correction ^a	7,470	1	,006		
Likelihood Ratio	8,355	1	,004		
Fisher's Exact Test				,004	,003
Linear-by-Linear Association	8,054	1	,005		
N of Valid Cases	6788				

a. Computed only for a 2x2 table

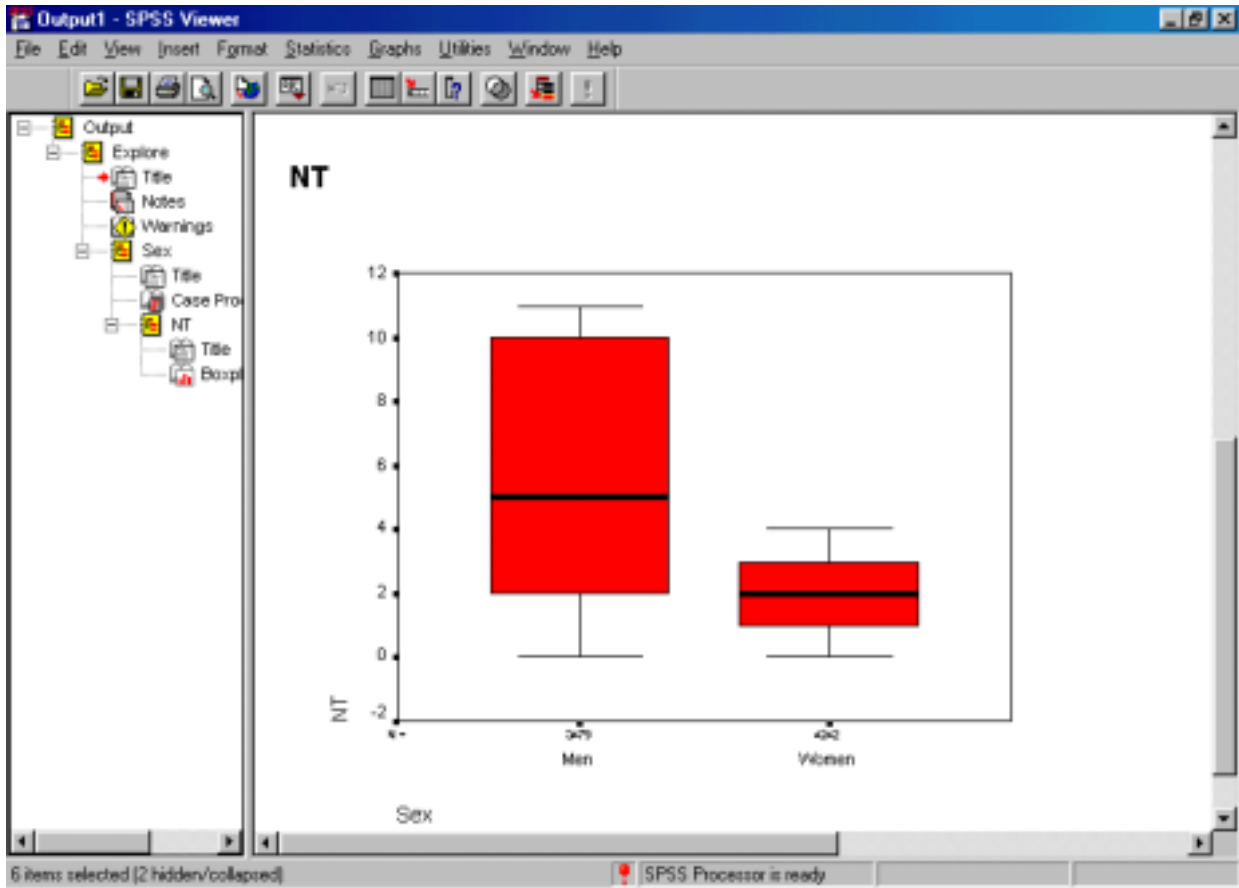
b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 41,51.

SPSS Processor is ready

g Lien entre nt et Sex

Graphs> Boxplot> Define
Variable : nt
Category Axis : Sex
Statistics > Compare means > Independent Samples T-Test
Test variables : nt
grouping variable Sex
Define group: 1 et 2
OK

Ecran



Statistics > Compare means > Independent Samples T-Test
 Test variables : nt
 grouping variable Sex
 Define group: 1 et 2
 OK

Ecran

T-Test

Group Statistics

Sex	N	Mean	Std. Deviation	Std. Error Mean
Men	3479	5,65	3,95	6,70E-02
Women	4242	2,17	1,32	2,03E-02

Independent Samples Test

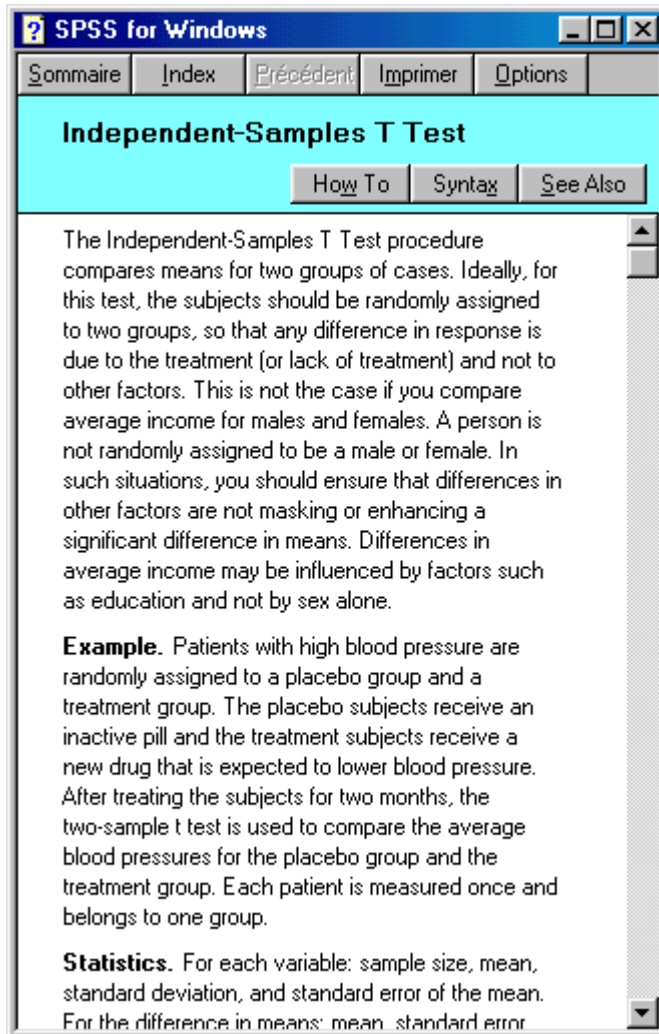
		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. E Differe
NT	Equal variances assumed	6095,540	,000	53,815	7719	,000	3,48	6,471
	Equal variances not assumed			49,724	4118,901	,000	3,48	7,001

SPSS Processor is ready

Remarque

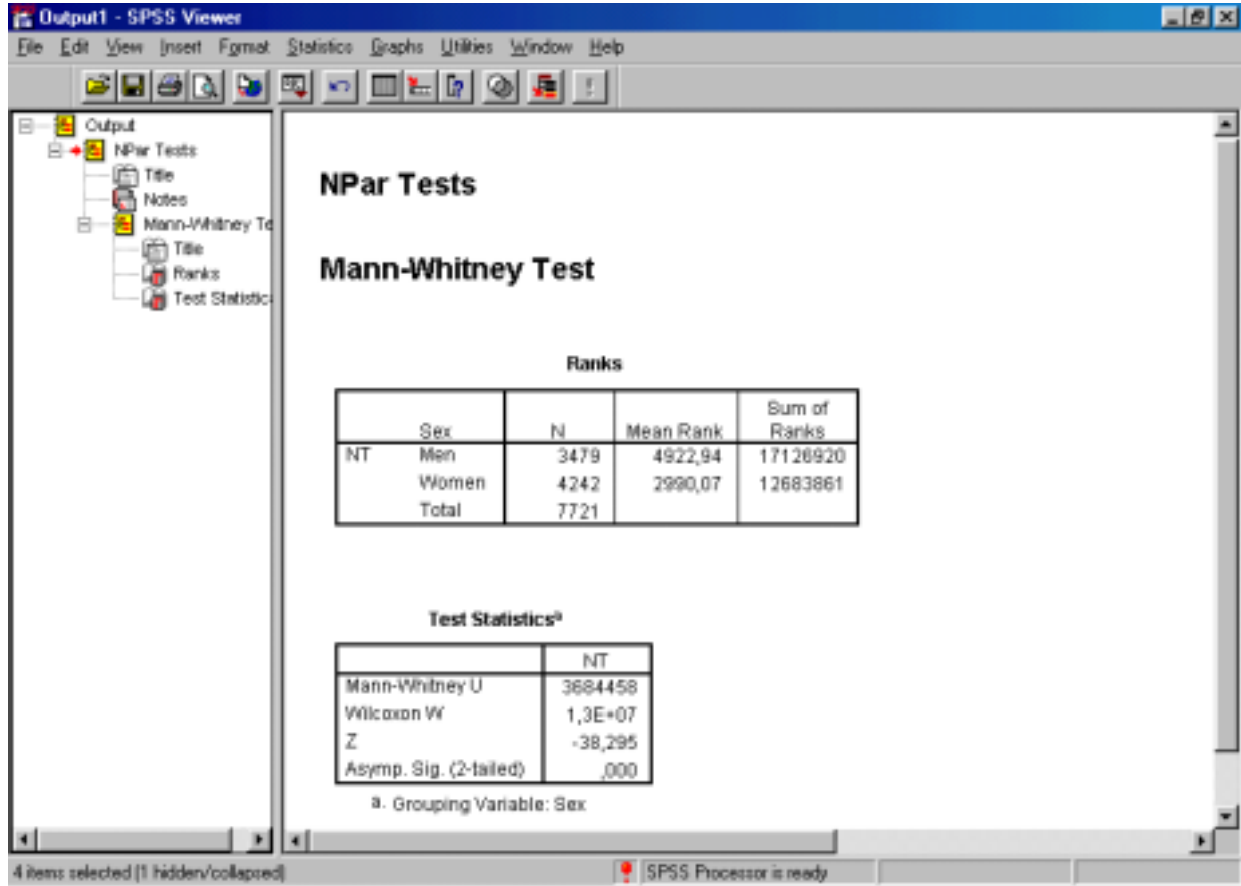
Statistics > Compare means > Independent Samples T-Test
Help

Ecran



Statistics > Nonparametrics tests > 2 Independent Samples
 Test variables List : nt
 grouping variable Sex (1,2)
 OK

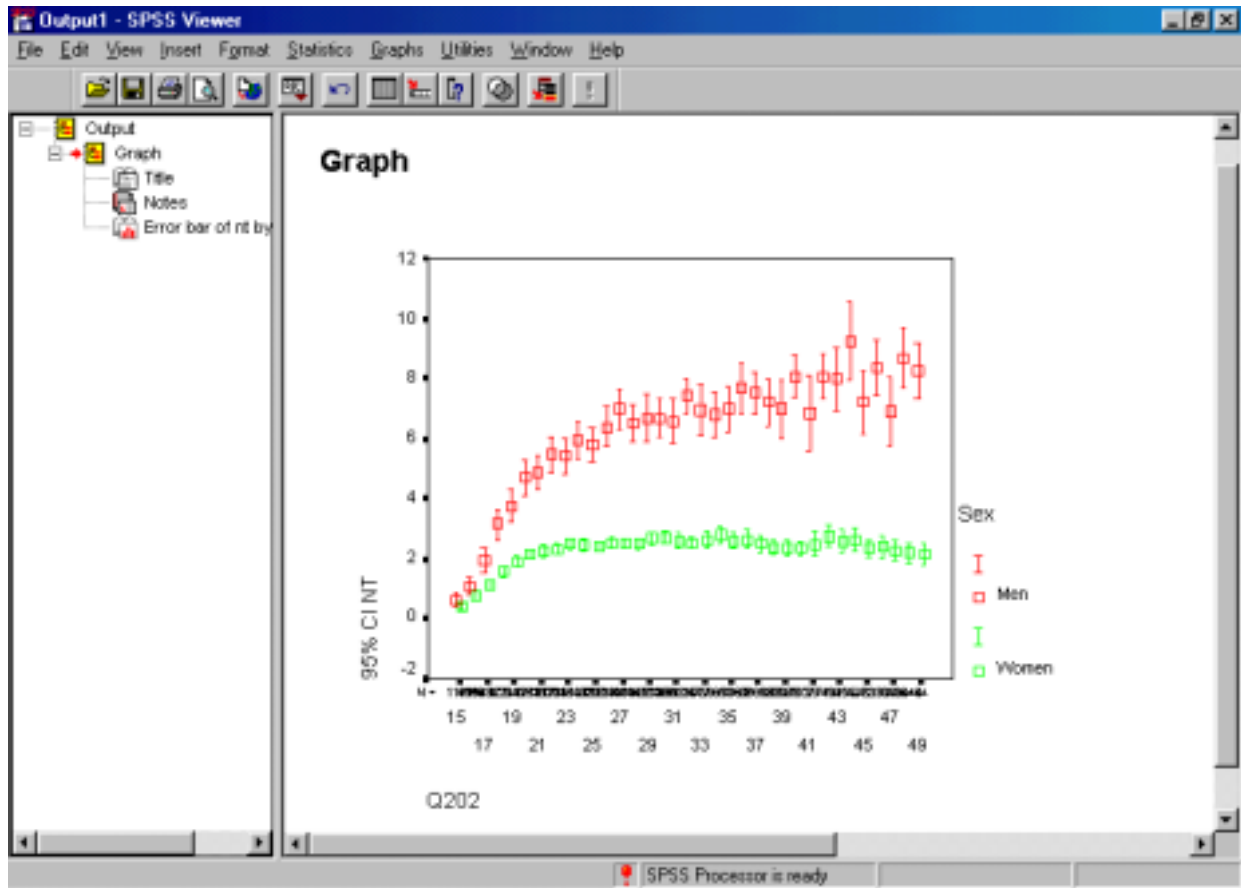
Ecran



h Lien entre nt et age (q202 ou Age) et sexe

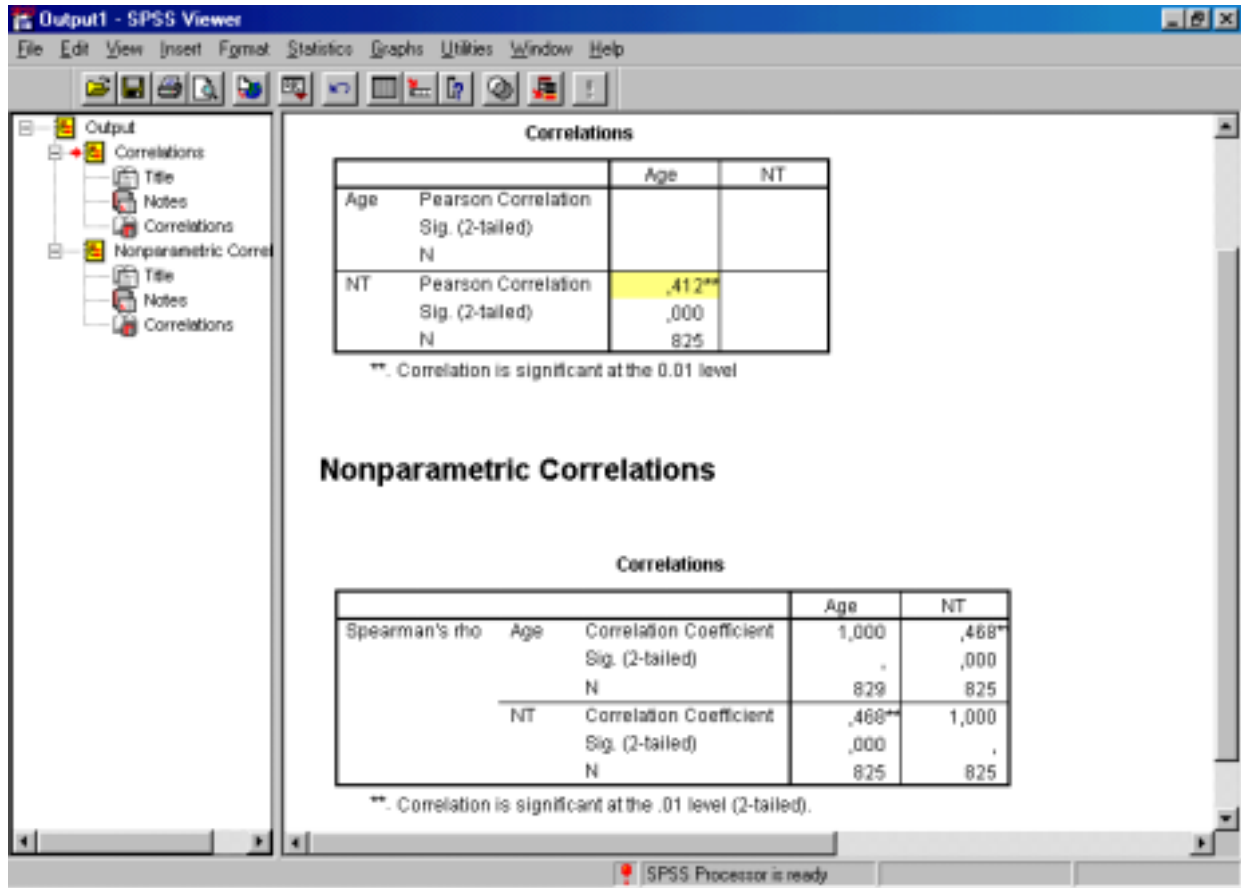
Graph > Error bars > Clustered
Variable : nt
Category Axis : q202 (ou Age)
Define Cluster by : Sex
OK

Ecran



Correlation entre Age et nt
 Statistics > Correlate > Bivariate
 Variables Age et nt
 Cocher Spearman
 OK

Ecran



i Sélection d'un sous-ensemble: Hommes à Kisumu (citycode=3)

Data > Select cases > If condition is satisfied > If....

citycode=3 and q201=2

Continue

OK

Vérification avec Statistics>Summarize> (cocher display frequency tables)

OK

Ecran

The screenshot shows the SPSS Output1 - SPSS Viewer window. The left pane displays a tree view of the output, including 'Citycode in pindiv', 'Sex', and 'HIV'. The main area contains three frequency tables:

Citycode in pindiv

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 3	829	100,0	100,0	100,0

Sex

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Men	829	100,0	100,0	100,0

HIV

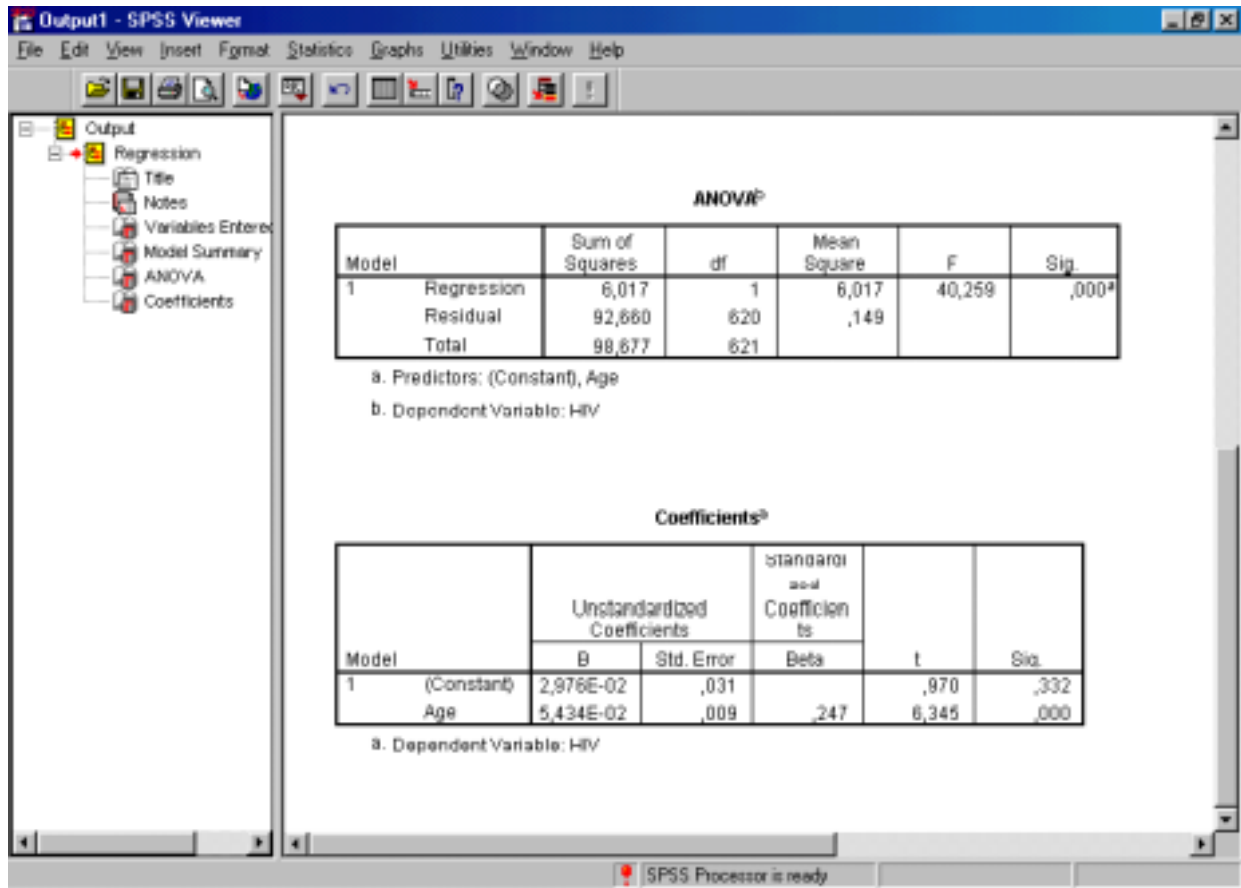
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Negatif	499	60,2	80,2	80,2
Valid Positif	123	14,8	19,8	100,0
Valid Total	622	75,0	100,0	
Missing System	207	25,0		
Total	829	100,0		

4 items selected (0 hidden/collapsed) | SPSS Processor is ready

j Etude du lien avec HIV de age, xbirth, xhsv2, seecsw

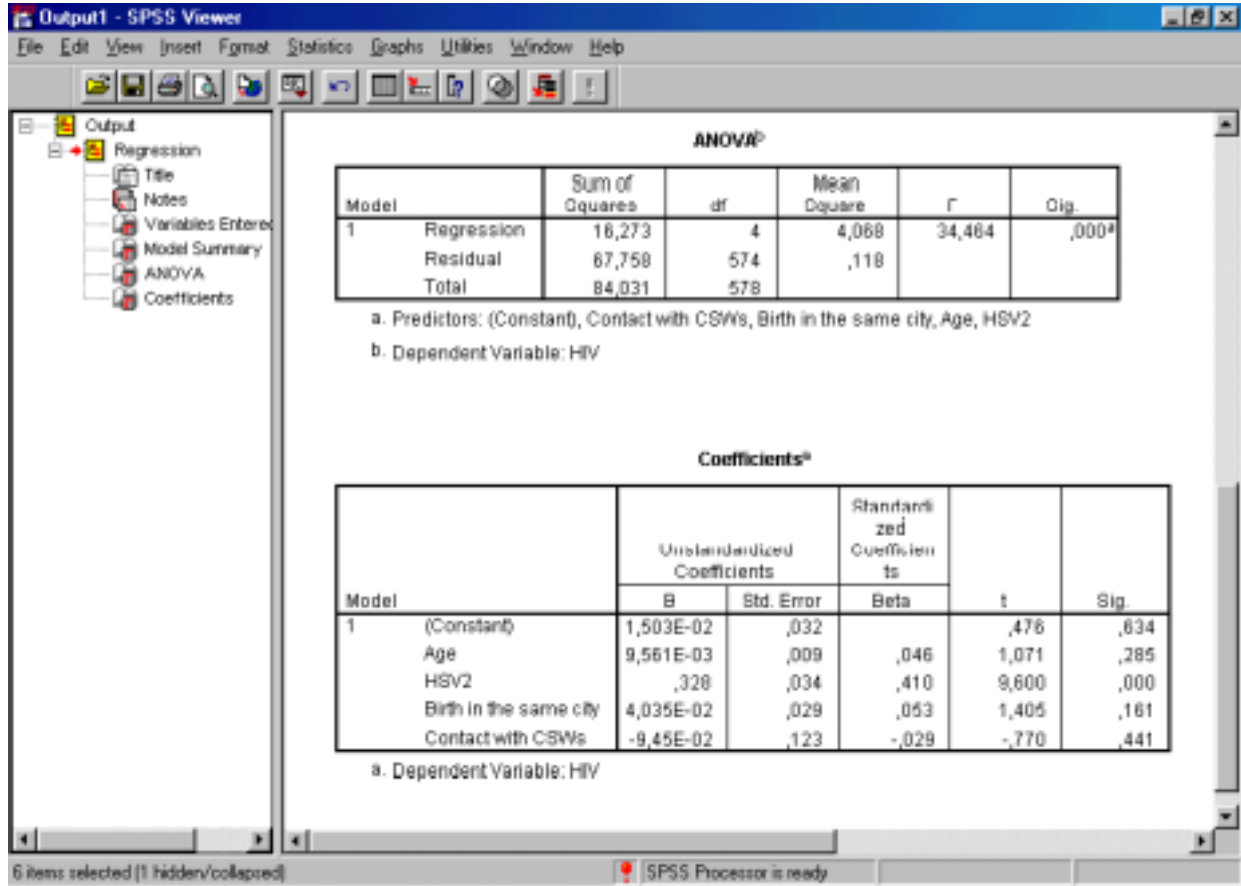
Statistics > Regression > Linear
 Dependent : HIV
 Independent(s): Age
 OK

Ecran



Refaire avec (variable par variable) age, xbirth, xhsv2, seecsw

Refaire avec les 4 variables ensembles



Statistics > Regression > Logistic
 Dependent : HIV
 Covariate(s): age, xbirth, xhsv2, seecsw
 Categorical: Age > continue
 OK

Ecran

Output1 - SPSS Viewer

File Edit View Insert Format Statistics Graphs Utilities Window Help

Observed		Predicted			Percent Correct
		Negatif	Positif		
Negatif	N	I 461	I 16	I	96,65%
Positif	P	I 81	I 21	I	20,59%
Overall					83,25%

Variables in the Equation							
Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
XHSV2	2,1572	,2951	53,4455	1	,0000	,3089	8,6468
AGE			12,8438	6	,0456	,0396	
AGE (1)	-1,6413	,6858	5,7280	1	,0167	-,0832	,1937
AGE (2)	-,2472	,5187	,2272	1	,6336	,0000	,7810
AGE (3)	,2676	,5132	,2719	1	,6021	,0000	1,3068
AGE (4)	,1721	,4920	,1223	1	,7265	,0000	1,1876
AGE (5)	-,1638	,5207	,0989	1	,7531	,0000	,8489
AGE (6)	-,5780	,5598	1,0662	1	,3018	,0000	,5610
SEECSW	-,7588	1,1812	,4126	1	,5207	,0000	,4682
XBIRTH	,4661	,2504	3,4656	1	,0627	,0521	1,5938
Constant	-2,6454	,5018	27,7912	1	,0000		

SPSS Processor is ready

6 Sauvegarde des résultats

Activer la fenêtre Output1
 Supprimer les sorties inutiles (par sélection avec la souris dans la partie gauche puis Edit > Cut
 File> Save as > Nom: a:\output1.spo

7 Impression d'un résultat

Activer la fenêtre Output1
 Sélectionner le résultat à imprimer avec la souris dans la fenêtre de droite ou de gauche (cf flèche rouge)
 File> Print > OK

8 Quitter

File > Exit
 Répondre Non à toutes les questions

Travail à faire par les étudiants:

Chaque étudiant devra sélectionner les données en fonction de son numéro en utilisant la sélection X...=1.

Ainsi l'étudiant 5 utilisera le filtre X05=1, l'étudiant 24 le filtre X24=1.

Avec le sous-ensemble ainsi obtenu:

1 Descriptions des variables sex et citycode et de 3 variables au choix parmi toutes les variables dont 1 qualitative à 2 classes, 1 qualitative à >2 classes et une quantitative

2 Etude du lien entre une variable biologique (à choisir parmi HIV ou HSV2 ou xtrich ou xchlam ou xgono) en fonction de variables à choisir dans la liste suivante: xcirc age q202 xana xemploi xedu xq205 xalcohol occup status2 ethnic2 xtravel xbirth nt ntot_s q321 neverm xpolyg seecsw.

3 Ecrire en 1 ou 2 pages les analyses qui ont été faites et leurs résultats principaux

4 Lors de la soutenance de l'article, indiquer oralement ce qui a été fait.

5 facultatif Imprimer les principales analyses qui ont été faites

Remarques

Variables qualitative à 2 classes

HIV HSV2 xtrich xchlam xgono

q201 xcirc xana xemploi xedu xalcohol xtravel xbirth q321 neverm xpolyg seecsw

Variables qualitative à >2 classes

citycode age ethnic2 occup status2 xq205

Variables quantitatives

q202 nt ntot_s