Probabilités conditionnelles et indépendance – Fiche de cours

1. Probabilités – Variable aléatoire

1.1. Vocabulaire et propriétés des événements

$\varnothing$ est appelé événement impossible
Ω (univers) est appelé événement certain
$A \cap B$ est l'événement « A et B »
$A \cup B$ est l'événement « A ou B »
$\overline{A}$ est appelé événement contraire de A
2 événements sont incompatibles lorsque $A \cap B = \varnothing$

1.2. Probabilité d'un événement

Une expérience aléatoire est constituée par plusieurs issues possibles qui dépendent du hasard.

Soit Ω un univers mathématique représentant l'ensemble des issues possibles d'une expérience aléatoire.

Définition
Soit A un événement tel que $A \subset \Omega$
On définit la probabilité d'un événement A par :

$$p(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} \qquad \text{avec } 0 \leq p(A) \leq 1$$

Equiprobabilité
Si l'univers Ω est constitué de n issues qui ont la même probabilité alors :

$$p = \frac{1}{n}$$

Formules
Soit A et B deux événements tel que $A \subset \Omega$ et $B \subset \Omega$
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ et $p(\overline{A}) = 1 - p(A)$

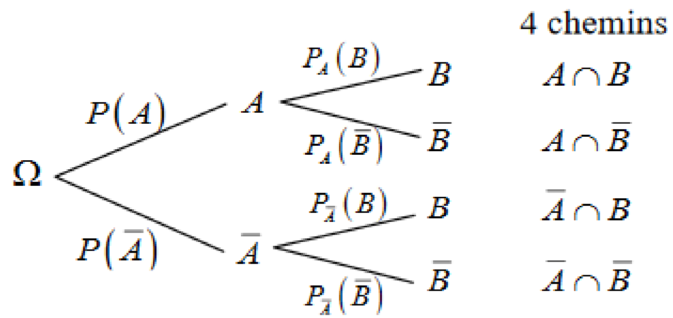
2. Conditionnement

a. Probabilités conditionnelles

Définition
Soient A et B deux événements d'un univers Ω
On appelle probabilité conditionnelle de l'événement A sachant B le réel :

$$P_B(A) = \frac{P(A \cap B)}{P(B)}$$

b. Arbre pondéré



Les branches sont pondérées par des probabilités conditionnelles
La somme des probabilités issues d'un même sommet (ou partition) est égale à 1

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

3. Indépendance

a. Événements indépendants

Définition :

Soit A et B deux événements de probabilité non nulle
A et B sont indépendants lorsque la réalisation de l'un ne change pas la réalisation de l'autre :

Deux événements A et B sont indépendants ssi :

$$P(A \cap B) = P(A) \times P(B)$$

Lorsque A et B sont deux événements indépendants, alors :

- $\overline{A}$ et B sont indépendants
- A et $\overline{B}$ sont indépendants
- $\overline{A}$ et $\overline{B}$ sont indépendants

b. Probabilités totales

Partition :

Les événements A_1, A_2, … A_n forment une partition d'un univers Ω ssi :

- les événements A_1, A_2, … A_n sont 2 à 2 disjoints
- $A_1 \cup A_2 \cup \ldots A_n = \Omega$

Définition :

Soit A_1, A_2, … A_n une partition d'un univers Ω, alors :

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_n)$$

$$P(B) = P(A_1) \times P_{A_1}(B) + P(A_2) \times P_{A_2}(B) + \ldots + P(A_n) \times P_{A_n}(B)$$

Chapitre 2:
Statistiques descriptives bivariées

- Recueil simultané de deux variables chez les même sujets : deux variables X et Y
- L'intérêt se porte le plus souvent sur la relation entre les deux variables, recherche de corrélation qui n'implique pas un lien de causalité.
- Les modalités sont donc des couples et les données sont présentées dans des tableaux à double entrée, encore appelés tableaux de contingence.
- L'analyse de ces tableaux vise à mettre en évidence d'éventuelles relations ou corrélations entre les deux variables.

Présentation des données

Il existe deux façons de présenter une série bivariée :

- Dans un tableau simple, constitué d'une colonne pour les observations (population) et de deux colonnes pour les variables Ainsi, chaque ligne comporte l'identifiant de l'observation (l'individu) dans la première colonne et les modalités observées pour chacune des deux variables dans les deux colonnes suivantes.
- Dans un tableau croisé ou tableau de contingence, où en ligne se trouvent les modalités (xi) d'une variable X et en colonne les modalités (yi) d'une autre variable (Y). On utilise un tableau de contingence quand on étudie une population sous l'angle de deux variables que l'on croise. un tableau de contingence croise ainsi les modalités des deux variables X et Y; On obtient des couples de modalités (xi ; yj) avec
  - xi la ième modalité de la variable X et
  - yj la jième modalité de la variable Y

Tableau simple

Le tableau suivant indique, pour chacune des trois circonscriptions, le nombre d'étudiants et le nombre d'enseignants.

circonscription	X: nombre d'étudiants	Y: nombre d'enseignants
Ali-Sabieh	759	18
Arta	141	4
Dikhil	725	19

Cette série double, ou bivariée, comporte trois individus. Si on mesure sur ces trois individus (population) les variables « X : le nombre d'étudiants » et « Y : le nombre d'enseignants », l'observation ou l'individu « Arta » est représentée par la modalité $(x2 ; y2) = (141; 4)$.

Tableau de contingence

Ce tableau de contingence se présente ainsi

- Les modalités xi de la variable X apparaissent dans la première colonne
- Les modalités yj de la variable Y apparaissent dans la première ligne
- L'effectif partiel nij de la modalité (ou couple de modalités) (xi, yj) est inscrit au croisement de la ligne i et de la colonne j.
- L'effectif marginal Ni+ de X est reporté dans la dernière colonne du tableau.
- L'effectif marginal N+j de Y est reporté sur la dernière ligne du tableau. La dernière ligne et la dernière colonne du tableau de contingence contiennent la distribution marginale de X et de Y. Elles représentent les effectifs des séries simples X et Y.
- L'effectif total n++ est indiqué au croisement des deux distributions marginales de X et de Y. D'où la présentation suivante du tableau de contingence :

X \ Y	y1	y2	.....	yj	......	yq	Effectif marginal de X: Ni+
X1	n11	n12	......	n1j	....	n1q	N1+
X2	n21	n22	......	n2j	....	n2q	N2+
.....							
xi	ni1	ni2	......	nij	....	niq	Ni+
......							
xp	np1	np2	......	npj	....	npq	Np+
Effectif marginal de Y : N+j	N+1	N+2	......	N+j	....	N+q	N++

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

Présentation des effectifs

L’effectif partiel du couple de modalités (xi, yj) est

Le nombre d’observations présentant simultanément les deux modalités xi et yj. Il se note nij.

L’effectif marginal de la modalité xi se note Ni+, ou encore Ni•. Cet effectif désigne la somme des effectifs de la ligne i.

Les modalités (valeurs) xi de X associées à Leurs effectifs marginaux Ni+ s’appelle distribution marginale de X.

$$N_{i+} = \sum_{j=1}^{p} n_{ij}$$

Présentation des effectifs

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

L'effectif marginal de la modalité yj se note N+j, encore N•j. Cet effectif désigne la somme des effectifs de la Colonne j. Les modalités (valeurs) yj de Y associées à leurs effectifs marginaux N+j s'appelle distribution marginale de Y.

$$N_{+j} = \sum_{i=1}^{q} n_{ij}$$

L'effectif total de la série double est la somme des effectifs marginaux de la série X (ou Y). Il est noté N++, N•• ou simplement n,

$$N_{++} = \sum_{j=1}^{q} \sum_{i=1}^{p} n_{ij}$$

Tableau de contingence

Suite à la dernière reforme sur la formation professionnelle, une enquête de satisfaction sur un échantillon de 200 étudiants de deux sexes a donné les résultats suivants, avec X la variable « sexe » et Y la variable « satisfaction ». La signification de chaque valeur est donnée dans le diapos suivant :

X \ Y	Satisfait	Insatisfait	Ni satisfait, ni insatisfait	Total = Ni+
Masculin	40	20	30	90
Féminin	55	35	20	110
Total = N+j	95	55	50	N++ = 200

Tableau de contingence

Effectifs partiels

- 40 : nombre d'individus du couple de modalités (masculin, satisfait).
- 20 : nombre d'individus du couple de modalités (masculin, insatisfait).
- 30 : nombre d'individus du couple de modalités (masculin, ni satisfait ni insatisfait).
- 55 : nombre d'individus du couple de modalités (féminin, satisfait).
- 35 : nombre d'individus du couple de modalités (féminin, insatisfait).
- 20 : nombre d'individus du couple de modalités (féminin, ni satisfait ni insatisfait).

Effectifs marginaux de X

- 90 : nombre d'individus de sexe masculin ou de la modalité « masculin » de la variable X
- 110 : nombre d'individus de sexe féminin ou de la modalité « féminin » de la variable sexe
- 30 : nombre d'individus du couple de modalités (masculin, ni satisfait ni insatisfait).

Effectifs marginaux de Y

- 95 : nombre d'individus satisfaits ou de la modalité « satisfait » de la variable Y
- 55 : nombre d'individus insatisfaits ou de la modalité « insatisfait » de la variable Y
- 50 : nombre d'individus ni satisfaits ni insatisfaits ou de la modalité « ni satisfait ni insatisfait » de la variable Y

Effectif total

200 : taille de la population (somme des effectifs partiels) ou somme des effectifs marginaux de X ou somme des effectifs marginaux de Y

Présentation des fréquences

À partir du tableau de contingence composé des effectifs, il est possible de calculer les fréquences (fréquences relatives). Il existe trois types de fréquences :

- les fréquences partielles ;
- les fréquences marginales ;
- les fréquences conditionnelles.

Présentation des fréquences

La fréquence partielle du couple de modalité (xi, yj) est notée fij et est définie par. Il est clair que ΣΣfij = 1 . On retrouve le Concept d'intersection, ces Individus appartenant à la Modalité xi de X et à la modalité yj de Y.

$$f_{ij} = \frac{n_{ij}}{N_{++}}$$

La fréquence marginale de la modalité xi est notée fi+ et est définie.

$$f_{i+} = \frac{N_{i+}}{N_{++}}$$

De même, la fréquence marginale de la modalité yj est notée f+j et est définie par

$$f_{+j} = \frac{N_{+j}}{N_{++}}$$

Tableau de contingence des fréquences

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

Un tableau de contingence des fréquences est obtenue en divisant les effectifs partiels et marginaux par l'effectif total

X \ Y	y1	y2	.....	yj	......	yq	Fréquence marginal de X: Ni+
X1	n11 / N++	n12 / N++	......	n1j / N++	....	n1q / N++	N1+ / N++
X2	n21 / N++	n22 / N++	......	n2j / N++	....	n2q / N++	N2+ / N++
.....							
xi	ni1 / N++	ni2 / N++	......	nij /N++	....	niq / N++	Ni+ / N++
......							
xp	np1 / N++	np2 / N++	......	npj / N++	....	npq / N++	Np+ / N++
Fréquence marginal de Y : N+j	N+1 / N++	N+2 / N++	......	N+j / N++	....	N+q / N++	N++ / N++

Présentation des fréquences

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

Reprenons les données de l'enquête de satisfaction

sur un échantillon de 200 étudiants de deux sexes. Le calcul des fréquences partielles et marginales est le suivant :

X \ Y	Satisfait	Insatisfait	Ni satisfait, ni insatisfait	Fréquence marginale de X = Fi+
Masculin	40 / 200 = 0,2	20 / 200 = 0,1	30 / 200 = 0,15	90 / 200 = 0,45
Féminin	55 / 200 = 0,275	35 / 200 = 0,175	20 / 200 = 0,1	110 / 200 = 0,55
Fréquence marginale de Y = F+j	95 / 200 = 0,475	55 / 200 = 0,275	50 / 200 = 0,25	N++ = 200 / 200 = 1

Fréquences conditionnelles

Fréquences conditionnelles de X sachant Y : La fréquence conditionnelle de la modalité xi sachant yj est donnée par

$$f(X = x_i / Y = y_j) = \frac{n_{ij}}{N_{+j}}$$

Fréquences conditionnelles de Y sachant X : La fréquence conditionnelle de la modalité yj sachant xi est donnée par

$$f(Y = y_j / X = x_i) = \frac{n_{ij}}{N_{i+}}$$

Exemple

- les fréquences conditionnelles de X sachant Y.

X \ Y	Satisfait	Insatisfait	Ni satisfait, ni insatisfait
Masculin	40 / 95 = 0,42	20 / 55 = 0,36	30 / 50 = 0,6
Féminin	55 / 95 = 0,57	35 / 55 = 63	20 / 50 = 0,4
Total = f+j	1	1	1

- les fréquences conditionnelles de Y sachant X.

X \ Y	Satisfait	Insatisfait	Ni satisfait, ni insatisfait	Total = fi+
Masculin	40 / 90 = 0,44	20 / 90 = 0,22	30 / 90 = 0,3	1
Féminin	55 / 110 = 0,5	35 / 110 = 0,318	20 / 110 = 0,18	1

Caractéristiques des séries à deux caractères

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

Les fréquences, indicateurs qui se calculent dans le cadre des séries univariées, se calculent également sur des séries bivariées. Il en va de même pour les autres caractéristiques des séries statistiques que sont la moyenne, la variance et l’écart-type. Ces caractéristiques peuvent être calculées sur des variables quantitatives, à partir :

- des distributions marginales : il s’agit de caractéristiques marginales ;
- des distributions conditionnelles : il s’agit de caractéristiques conditionnelles.

Caractéristiques marginales

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

Les séries marginales sont des séries univariées. Les calculs des moyennes, variances et écarts types marginaux se font donc de la façon habituelle, Après extraction de la série marginale.

La moyenne marginale de X est égale à

p : nombre de modalités de X

$$\bar{X} = \frac{1}{N_{++}} \times \sum_{i=1}^{p} N_{i+} \times x_i$$

La moyenne marginale de Y est égale à

q : nombre de modalités de Y

$$\bar{Y} = \frac{1}{N_{++}} \times \sum_{j=1}^{q} N_{+j} \times y_j$$

Exemple

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Soit un échantillon d'entreprises sur lequel sont observées les variables X, investissement annuel en milliers d'euros, et Y, chiffre d'affaires annuel en millions d'euros :

La moyenne marginale de X est $\bar{X} = \dfrac{(10 \times 380 + 35 \times 320 + 50 \times 300)}{1000} = 30$

La moyenne marginale de Y est $\bar{Y} = \dfrac{(5 \times 390 + 25 \times 310 + 65 \times 300)}{1000} = 29{,}2$

X \ Y	5	25	65	Ni+
10	300	80	0	380
35	70	200	50	320
50	20	30	250	300
	390	310	300	1000

Caractéristiques marginales

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

La variance marginale de X est égale à
p : nombre de modalités de X

$$V(X) = \frac{1}{N_{++}} \times \sum_{i=1}^{p} N_{i+} \times (x_i - \bar{X})^2$$

L'écart-type marginale de X est égale à

$$\sigma_X = \sqrt{V(X)}$$

La variance marginale de Y est égale à

q : nombre de modalités de Y

$$V(Y) = \frac{1}{N_{++}} \times \sum_{j=1}^{q} N_{+j} \times (y_j - \bar{Y})^2$$

L'écart-type marginale de Y est égale à

$$\sigma_Y = \sqrt{V(Y)}$$

Exemple

Soit un échantillon d'entreprises sur lequel sont observées les variables X, investissement annuel en milliers d'euros, et Y, chiffre d'affaires annuel en millions d'euros :

La variance marginale de X est $V(X) = \dfrac{380 \times (10-30)^2 + 320 \times (35-30)^2 + 300 \times (50-30)^2}{1000} = 280$

La variance marginale de Y est $V(Y) = \dfrac{390 \times (5-29{,}2)^2 + 310 \times (25-29{,}2)^2 + 300 \times (65-29{,}2)^2}{1000} = 618{,}36$

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

Notion de la Covariance

Dans le cas des séries doubles, nous disposons d'un indicateur, appelé covariance, qui permet de mesurer Les fluctuations simultanées de chaque variable par rapport à sa moyenne. Il est important de noter que, contrairement à la variance (moyenne de carrés) qui est toujours positive ou nulle, la covariance peut être de signe quelconque.

La covariance : Soit X et Y deux caractères quantitatifs. La covariance du couple (X ; Y) est définie par :

$$Cov(X;Y) = \frac{1}{n_{++}} \sum_{j=1}^{q} \sum_{i=1}^{p} n_{ij} (x_i - \bar{x})(y_j - \bar{y}) = \sum_{j=1}^{q} \sum_{i=1}^{p} f_{ij} (x_i - \bar{x})(y_j - \bar{y}).$$

Signe de la covariance:

- Une covariance positive indique que les caractères X et Y varient globalement dans le même sens.
- Une covariance négative indique que les caractères X et Y varient globalement en sens contraire.

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

Exemple

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

À partir du tableau observé ou donné ci-dessous, on détermine les moyennes de X et Y. Moyenne de X est $\ddot{X}$ =30 et Moyenne de Y est $\bar{Y}$ =29,2 (voir comment calculer sur le diapos 18)

X \ Y	5	25	65
10	300	80	0
35	70	200	50
50	20	30	250

Pour chaque couple de modalité (xi, yj), on calcule la valeur vij = nij×(xi - $\ddot{X}$) ×(yi - $\bar{Y}$)

La covariance est obtenue en divisant la somme des vij par l'effectif total N++
COV(X , Y) =(1/N++) × ∑vij = (1/N++) × ∑nij × (xi - $\ddot{X}$) × (yi - $\bar{Y}$) = 315000/1000 = 315

X \ Y	5	25	65	Total
10	300x(10-30)x(5-29,2)	80x(10-30)x(25-29,2)	0x(10-30)x(65-29,2)	151920
35	70x(35-30)x(5-29,2)	200x(35-30)x(25-29,2)	50x(35-30)x(65-29,2)	-3720
50	20x(50-30)x(5-29,2)	30x(50-30)x(25-29,2)	250x(50-30)x(65-29,2)	166800
Total	127050	0	187950	315000

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

Coefficient de corrélation linéaire

Le degré de liaison linéaire entre X et Y est étudie par le coefficient de corrélation linéaire noté r.

$$\rho = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}}$$

- encore noté r(x,y) le coefficient de corrélation est compris entre -1 et +1. •
- si r = 1 ou si r =–1, x et y sont parfaitement corrèles et les points sont alignés sur une droite.

- Si les variables X et Y sont linéairement indépendantes alors

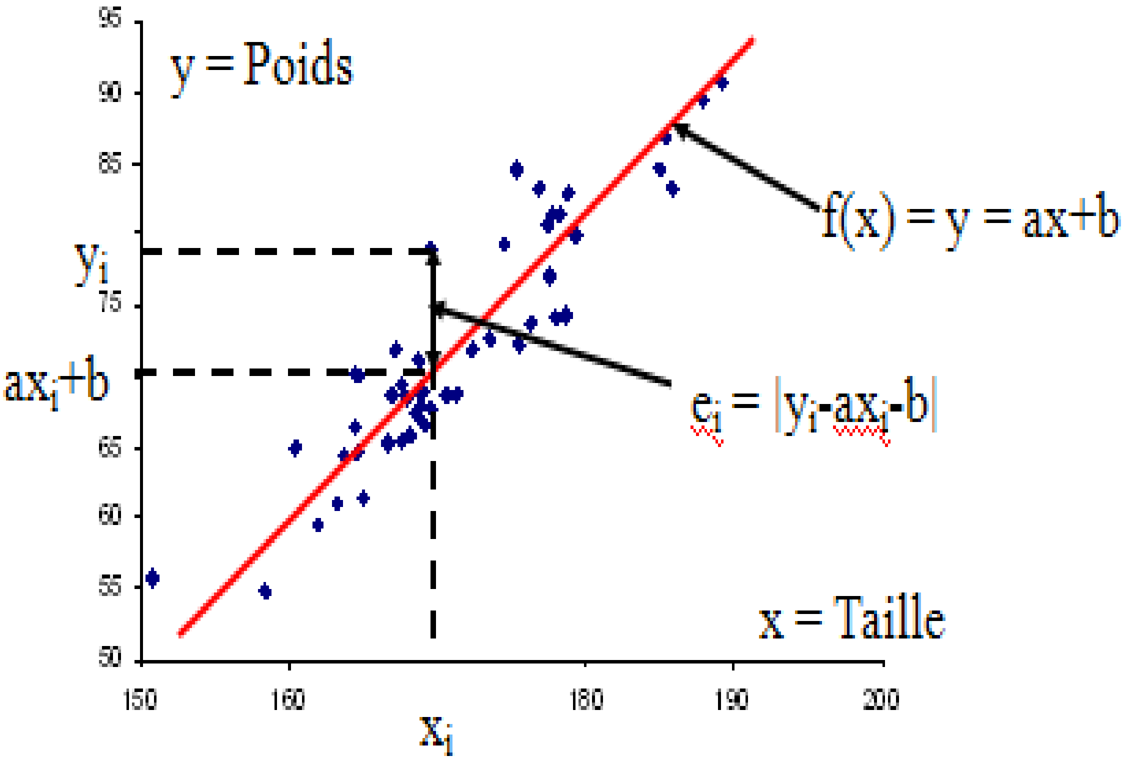
$r = 0$.

- Si le coefficient de corrélation $r = 0$ et que les variables X et Y suivent une loi normale, alors elles sont linéairement indépendantes.

- mais si $r = 0$ cela n'implique pas l'absence de relation entre x et y → si $f(x) = x2$: relation, mais non linéaire et $r = 0$.

Notion de droite de régression

La *droite de régression* est la droite qui ajuste au mieux un nuage de points au sens des moindres carres.



- Soit D une droite d'ajustement. Soit Mi (xi ; yi) un point du nuage. Pi est le point de même abscisse xi que Mi situé sur la droite D d'équation y = ax + b.

- On appelle droite de régression de y en x, la droite D tell que :

$$\sum MiPi^2 = \sum[y_i - (ax_i + b)]^2$$

- Le principe est de trouver a et b, c'est le critère des moindres carrés

Etude de liaison entre deux variables

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

La notion de liaison entre deux variables est un premier stade incontournable vers une éventuelle imputation causale qu'il est fondamental de mettre en évidence. Dans le cas particulier de deux caractères quantitatifs, le degré d'association Peut Varier entre deux extrêmes : d'un côté la liaison fonctionnelle et de l'autre l'indépendance.

Un caractère X est lié fonctionnellement au caractère Y si à chaque modalité de Y Correspond une seule modalité de X.

X est fonctionnellement lié à Y, car pour chaque modalité de Y résulte une seule modalité de

X ; ainsi, un consommateur satisfait est nécessairement une femme. Par contre, Y n'est pas

fonctionnellement lié à X, car à la modalité Masculin de X correspondent deux modalités

pc

sy

X \ Y	Insatisfait	Ni satisfait, ni insatisfait	Satisfait	Somme
Masculin (M)	20	70	0	90
Féminin (F)	0	0	110	110
Somme	20	70	110	200

Etude de liaison entre deux variables

Deux variables statistiques X et Y sont indépendantes si les distributions conditionnelles de X sachant Y sont identiques et égales à la distribution marginale de X ce qui équivaut à :

$$f(X = x_i / Y = y_j) = \frac{n_{ij}}{N_{+j}} = f_{i+} = \frac{N_{i+}}{N_{++}}$$

Le concept d'indépendance étant symétrique, Deux variables statistiques X et Y sont indépendantes si les distributions conditionnelles de Y sachant X sont identiques et égales à la distribution marginale de Y ce qui équivaut à :

$$f(Y = y_j / X = x_i) = \frac{n_{ij}}{N_{i+}} = f_{+j} = \frac{N_{+j}}{N_{++}}$$

Etude de liaison entre deux variables

Les fréquences marginales de X (0,45 et 0,55)

X \ Y	Satisfait	Insatisfait	Ni satisfait, ni insatisfait	Total = Fi+
Masculin	40 / 200 = 0,2	20 / 200 = 0,1	30 / 200 = 0,15	90 / 200 = 0,45
Féminin	55 / 200 = 0,275	35 / 200 = 0,175	20 / 200 = 0,1	110 / 200 = 0,55

Les fréquences conditionnelles de X selon Y sont dans le tableau ci-dessous

X \ Y	Satisfait	Insatisfait	Ni satisfait, ni insatisfait
Masculin	40 / 95 = 0,42	20 / 55 = 0,36	30 / 50 = 0,6
Féminin	55 / 95 = 0,57	35 / 55 = 0,63	20 / 50 = 0,4
Total = f+j	1	1	1

On remarque que les fréquences conditionnelles de X selon Y ne sont pas identiques (égales) aux fréquences marginales de X.
On conclut que X et Y ne sont pas indépendantes c'est-à-dire que la réponse d'un individu à la satisfaction n'est pas indépendante de son sexe

Liaison entre deux variables qualitatives : introduction au test de khi-deux

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

Le test d'indépendance du khi-deux (khi2) permet de se prononcer sur l'indépendance de deux variables qualitatives, observées sur un échantillon. Il s'effectue en deux étapes :

1) Étape 1

Calculer la distance entre le tableau des effectifs observés et le tableau des effectifs théoriques calculés sous l'hypothèse d'indépendance en déterminant le khi-deux calculé qui est un indicateur permettant d'accepter ou de refuser l'hypothèse d'indépendance entre les variables X et Y

Étape 2

La deuxième étape, consiste à comparer la valeur du khi-deux calculé à la valeur du khi-deux théorique lue sur la table de la loi de khi-deux pour un degré de liberté (ddl) égale à ν=(nombre de ligne - 1)*(nombre de colonne - 1) et un risque d'erreur α donné. Ou bien comparer la probabilité associée à la décision d'accepter ou de refuser l'hypothèse d'indépendance au risque d'erreur choisit.

Tableau de contingence théorique

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

C'est un tableau des effectifs théoriques calculés sous l'hypothèse d'indépendance entre les deux variables croisées X et Y. l'effectifs est noté cij et est égale à : cij = (Ni+ x N+j) / N++

X \ Y	y1	y2	.....	yj	......	yq	Effectif marginal de X: Ni+
X1	(N1+ × N+1) / N++	(N1+ × N+2) / N++	......	(N1+ × N+j) / N++	....	(N1+ × N+q) / N++	N1+
X2	(N2+ × N+1) / N++	(N2+ × N+2) / N++	......	(N2+ × N+j) / N++	....	(N2+ × N+q) / N++	N2+
.....							
xi	(Ni+ × N+1) / N++	(Ni+ × N+2) / N++	......	(Ni+ × N+j) / N++	....	(Ni+ × N+q) / N++	Ni+
......							
xp	(Np+ × N+1) / N++	(Np+ × N+2) / N++	......	(Np+ × N+j) / N++	....	(Np+ × N+q) / N++	Np+
Effectif marginal de Y : N+j	N+1	N+2	......	N+j	....	N+q	N++

Tableau de contingence théorique

A partir du tableau de contingence observé ci-dessous, on calcule les effectifs théoriques notés cij et est égale à : cij = (ni+ x n+j) / N++

X \ Y	Satisfait	Insatisfait	Ni satisfait, ni insatisfait	Total = Ni+
Masculin	40	20	30	90
Féminin	55	35	20	110
Total = N+j	95	55	50	N++ = 200

Le tableau de contingence théorique obtenu est le suivant

X \ Y	Satisfait	Insatisfait	Ni satisfait, ni insatisfait	Total = Ni+
Masculin	(90 x 95)/200 = 42,75	(90 x 55)/200 = 24,75	(90 x 50)/200 = 22,5	90
Féminin	(110 x 95)/200 = 52,25	(110 x 55)/200 = 30,25	(110 x 50)/200 = 27,5	110
Total = N+j	95	55	50	N++ = 200

Distance de khi-deux et degré de liberté d'un tableau de contingence

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Après détermination des effectifs calculés cij, il est possible de déterminer un indicateur de distance entre le tableau observé, composé des nij, et le tableau théorique, composé des cij. Cette « distance » est appelée distance du khi-deux.

Distance du khi-deux : La distance entre les tableaux observé et théorique est appelée khi-deux calculé, notée χ_c^2, et définie par $\chi_c^2 = \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{(n_{ij} - c_{ij})^2}{c_{ij}}$, les coefficients c_{ij} désignant les effectifs théoriques ou calculés et les n_{ij} les effectifs observés.

Degré de liberté d'un tableau de contingence : Soit un tableau de contingence formé de n lignes et de p colonnes. Son degré de liberté, noté ddl, est donné par : ddl = (n–1)(p–1), ou encore ddl = (nombre de lignes – 1) × (nombre de colonnes – 1).

Test d'indépendance

On considère

- α = le risque de rejeter l'indépendance entre les variables X et Y sachant qu'elle est vraie.
- $\chi^2_{(\alpha;\nu)}$ = la valeur du khi-deux critique, dépendant du seuil de signification α et du degré de liberté ν (ddl)

- Si $\chi_c^2 \geq \chi^2_{(\alpha;\nu)}$, l'hypothèse H_0 d'indépendance entre les deux variables est rejetée et l'hypothèse H_1 est acceptée : les deux caractères seront considérés comme statistiquement associés.
- Si $\chi_c^2 \leq \chi^2_{(\alpha;\nu)}$, l'hypothèse H_0 d'indépendance entre les deux variables n'est pas rejetée : il est impossible de conclure de façon significative à l'existence d'un lien statistique entre les variables.

Distance de khi-deux

Le degré de liberté est égale à ddl (v) = (2-1) x (3-1) = 2

La distance du khi-deux obtenu est égale à χ_c^2 = 6,251

La valeur critique pour un risque d'erreur 5% et un degré de liberté 2 est $\chi^2_{(\alpha;v)}$ = 5,991

On conclut qu'il existe une relation entre les variables X et Y car $\chi_c^2 > \chi^2_{(\alpha;v)}$

X \ Y	Satisfait	Insatisfait	Ni satisfait, ni insatisfait	Total = Ni+
Masculin	(-2,75)^2 / 42,75 = 0,176	(-4,75)^2 / 24,75 = 0,911	(7,5)^2/22,5 = 2,5	3,587
Féminin	(2,75)^2 / 52,25 = 0,144	(4,75)^2/30,25 = 0,475	(-7,5)^2/27,5 = 2,045	2,664
Total = N+j	0,32	1,386	4,545	6,251

ANALYSE DES DONNEES AVEC LE LOGICIEL SPSS

Présenté par
Dr. IDRISS OKIYE WAISS

09/12/2024

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Objectif

Ce cours a pour objectif de former les étudiants de la filière COTM2 en statistiques et traitement de données sur le logiciel SPSS. Il s'agit de découvrir, par l’application, le principe de fonctionnement du logiciel, ses fonctionnalités, leur intérêt, leur mise en œuvre, la présentation et l’interprétation des résultats des analyses.

Les participants vont prendre connaissance d'exemples pratiques d'appui et d'illustration. Ainsi, un module de formation est dédié à l’application et l’exploitation des données.

Canevas de la formation

Les différents parties de ce cours sont organisées en se basant sur le type de variables disponibles et à analyser.

Que peut-on utiliser comme outil statistique lorsqu'on dispose de :

- Une seule variable quantitative
- Une seule variable qualitative
- Plusieurs variables quantitatives
- Une variable quantitative selon au moins une variable qualitative
- Une variable qualitative selon au moins une variable quantitative
- Plusieurs variables qualitatives

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Chapitre 1:
Statistiques descriptives

L'objectif de ce chapitre est d'apprendre les principales techniques de statistique descriptive uni variée, bi variée et multidimensionnelle.

Etre capable de mettre en œuvre ces techniques de manière appropriée dans un contexte donné.

I. Introduction :

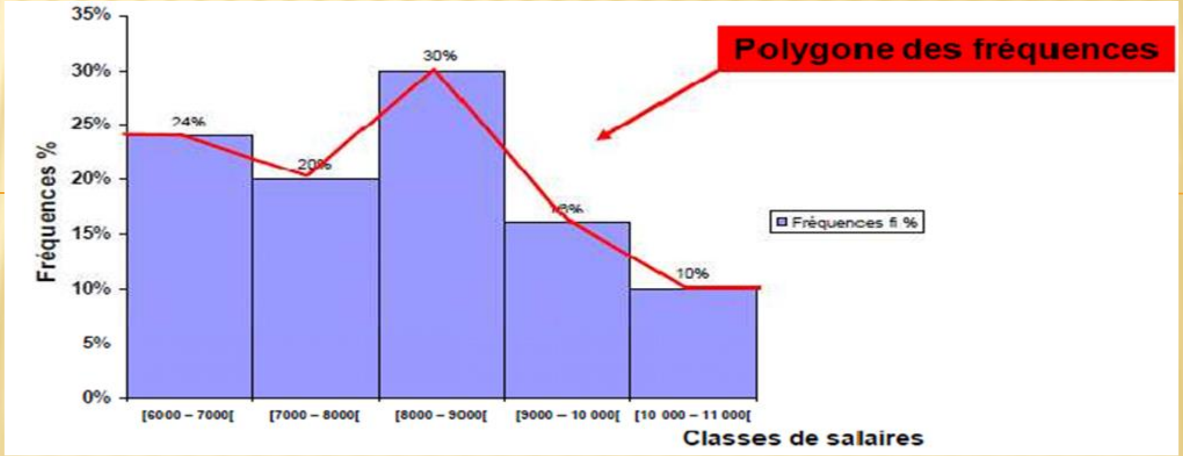
→ La statistique descriptive est un ensemble de méthodes de description, d’analyse de façon quantifiée des ensembles nombreux.

→ Elle comprend un ensemble d’outils d’investigation et de mesure de données chiffrées (tableaux, graphiques, calcul des moyennes….etc.) d’un caractère quantitatif ou qualitatif (variable) donné afin de ressortir la signification.

Quelles réponses offre la statistique?

1) Des outils descriptifs, pour **résumer** les donnés et les représenter graphiquement.
Résumé numérique pour estimer des paramètres d'une distribution : *moyenne, variance, médiane, quantiles, skewness, kurtosis, corrélation…*

Graphiques statistiques pour montrer / analyser les structures sous-jacentes aux données :
diagramme en barres, histogramme, boxplot…



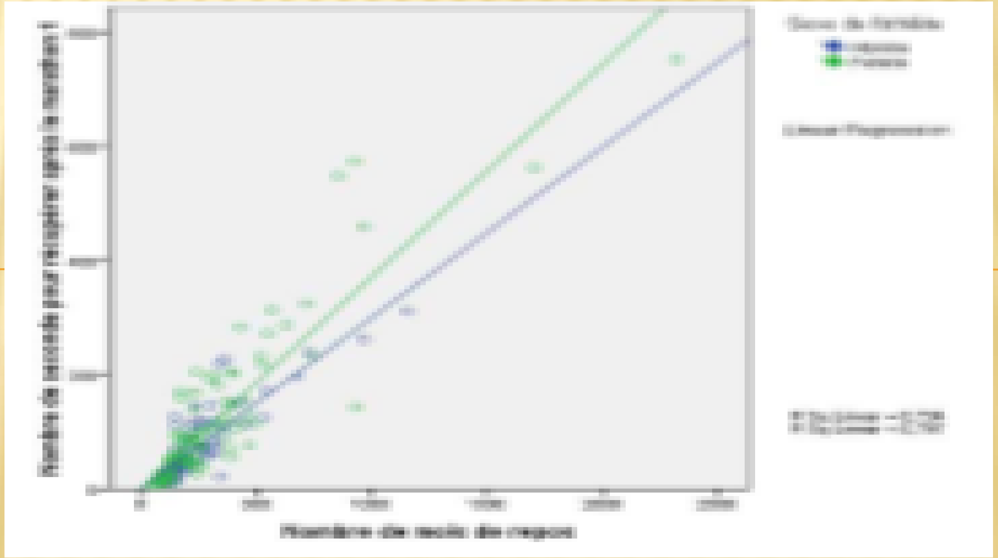
→ Des outils de modélisation, pour **expliquer** certaines variables à partir d' autres variables.

Les outils de modélisation expliquent et prédisent une ou plusieurs variables par une fonction mathématique d'autres variables.

Exemple: L' échec scolaire des élèves ruraux est-il liée à la distance parcouru pour se rendre à l'école?

Quelques exemples de modélisation :

Régression simple, multiple, régression logistique, analyse de variance



Statistique descriptive

Statistique descriptive à 1 variable

Tableau statistique

Graphiques (histogramme, diagramme en barre, circulaire)

Caractéristiques centrales (Moyenne, mode, médiane quartiles, déciles, centiles)

Caractéristiques de dispersion (étendue, écart interquartile, écart moyen absolu, variance et écart-type, coefficient de variation)

Statistique descriptive à 2 variables

Tableau de contingence

Caractéristiques marginales (moyenne, variance et écart-type marginales)

Caractéristiques conditionnelles (moyenne, variance et écart-type conditionnelles)

Covariance

Graphiques (boxplot, barplot, nuage de point)

Étude de liaison (corrélation de pearson, test de khi-deux, test exact de fisher, test de student, test de wilcoxon, analyse de variance, test de kruskall wallis, corrélation de spearman)

Statistique descriptive à 3 et plus de variables

Analyse factorielle (ACP, AFC, ACM, AFD, AFM)

Classification (CAH, classification mixte)

Segmentation

II- Concept de base :

Population : l'ensemble des éléments sur lesquels porte une étude. La population constitue l'univers de l'étude.

Individu : unité statistique de base ou élément de la population étudiées.

Caractère ou variable : c'est le critère d'étude d'une population. C'est ce qui est observé ou mesuré sur une population.

Echantillon: c'est un sous-ensemble de la population. La notion d'échantillon est fondamentale, en règle générale, la population entière n'est pas disponible ou observable.

Modalité: ce sont les différents résultats (ou valeurs) que peut prendre la variable statistique.

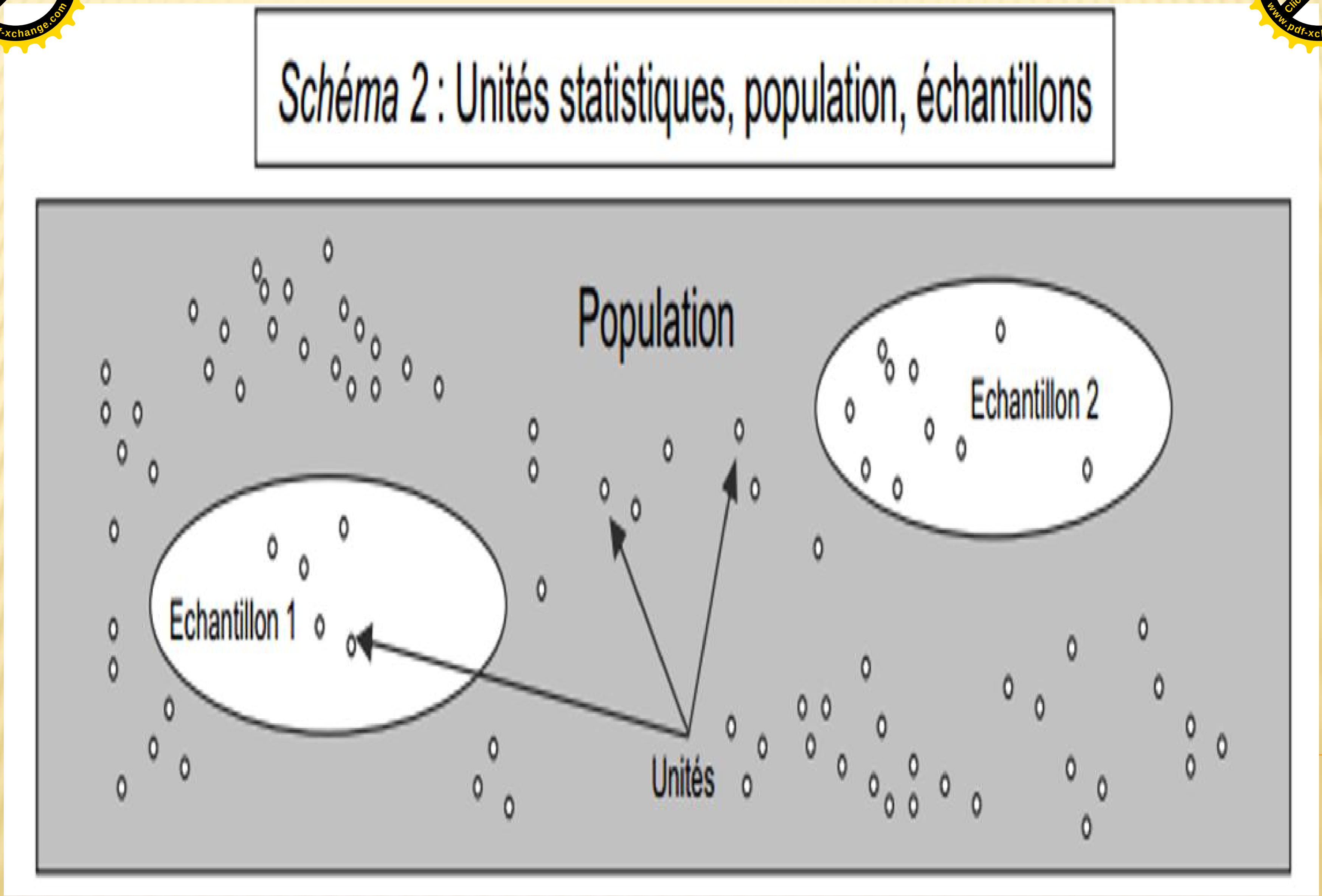


Schéma 2 : Unités statistiques, population, échantillons
Population
Echantillon 2
Echantillon 1
Unités

Exemple de population

Les éléments suivants sont autant d'exemples de population sur laquelle peut porter une étude.

- Les élèves inscrits en 1 ère année de l'OTI,
- Les infrastructures scolaires du ministère,
- Le personnel administratif d'un collège,
- Les habitants d'une ville ou commune,
- La surface des établissement scolaire,
- Les ménages ou les foyers d'une ville (région),
- Les écoles de plus 1000 élèves inscrits,
- Etc.

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Exemple de variable

Les éléments suivants sont autant d’exemples de critères (variables) d’étude d’une population

- Le prix
- L’effectif
- La quantité
- L’âge
- Le sexe
- L’opinion ou l’avis
- La qualité
- Le niveau d’étude
- La répartition géographique
- Le diplôme
- Etc.

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Application

Exemple 1: On réalise une étude sur la répartition selon le sexe de 400 étudiants d'un collège.

Traduisons ces informations dans le vocabulaire de la statistique descriptive.

Population	400 étudiants
Unité Statistique	Chaque étudiant du collège
Variable ou caractère statistique	Sexe
Modalités	Féminin ou Masculin
Nature de la variable	Qualitative nominale

Exemple 2: On soumet aux étudiant de la filière COTM2 un questionnaire de satisfaction concernant la formation suivie. Les réponses des élèves sont les suivantes : nulle, (b) médiocre, (c) moyenne, (d) assez bonne, (e) très bonne, (f) excellent

Population	Élèves de la formation professionnelle
Unité Statistique	Chaque élèves
Variable ou caractère statistique	La satisfaction
Modalités	(a) nulle, (b) médiocre, (c) moyenne, (d) assez bonne, (e) très bonne, (f) excellente
Nature de la variable	Qualitative ordinale

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Exemple 3: Le responsable des ressources humaines a relevé l'ancienneté, exprimée en année, de 2000 salariés de l'entreprise. Elle s'étale sur un intervalle de 0 à 35 ans.

Traduisons ces informations dans le vocabulaire de la statistique descriptive.

Population	2000 salariés de l'entreprise
Unité Statistique	Chaque salarié
Variable ou caractère statistique	Ancienneté
Modalités	[0 an ; 35 ans[
Nature de la variable	Quantitative continue

Exemple 4: Un groupe de 20 étudiants a répondu sur le nombre de films que chacun d'entre eux a vus au cours du mois dernier. Le résultat est le suivant : {0, 1, 2, 3}.

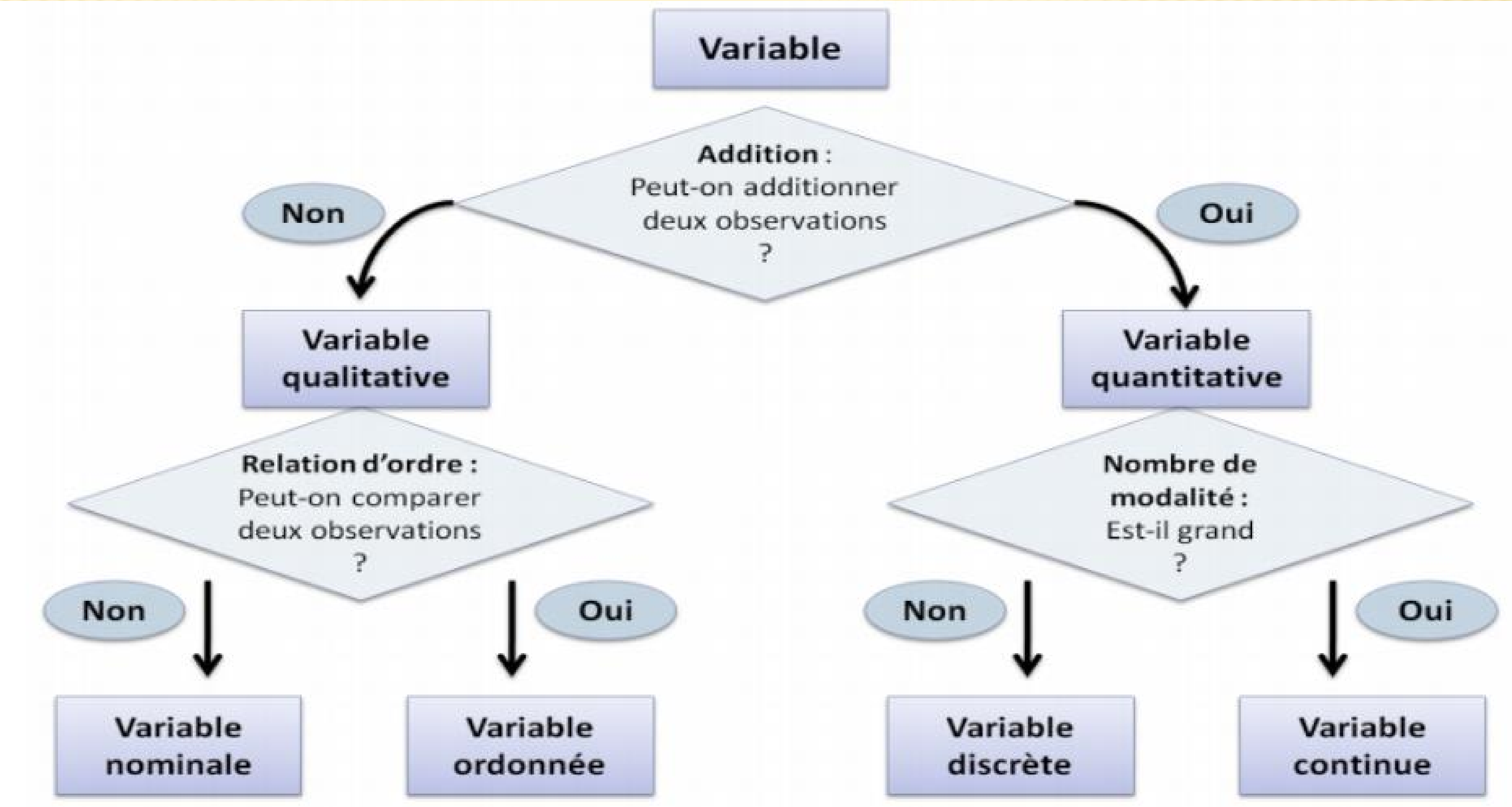
Traduisons ces informations dans le vocabulaire de la statistique descriptive.

Population	Groupe de 20 étudiants
Unité Statistique	Chaque étudiant
Variable ou caractère statistique	Nombre de films regardés
Modalités	{0, 1, 2, 3}
Nature de la variable	Quantitative discrète

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

III- Nature d’une variable :

La nature d'une variable détermine le type d'outil statistique qu'on pourra utiliser sur la variable.



Variable qualitative

Variable qualitative ordinale	Variable qualitative nominale
EXEMPLES	**EXEMPLES**
Classement à un examen (1^{er} , $2^{ème}$, $3^{ème}$, $4^{ème}$)	Couleur (marron, noire, bleu, grise, rouge)
Comportement (très mauvais, mauvais, bon, très bon, excellent)	Sexe (féminin/masculin)
Niveau scolaire (analphabète, primaire, secondaire, unive rsitaire)	Situation familiale (célibataire/marié)
Intensité de la douleur (aucune, faible, moyenne, forte, très forte)	Groupe sanguin (A/B/O/AB)
Consommation de tabac (absence, rarement, modéré, souvent)	

Variable quantitative

Variable quantitative discrète

Exemples

Nombre d'enfant par famille

La note à un contrôle

Variable quantitative continue

Exemples

Age

Poids

Taille

IV- Représentation des données:

Quand les données sont recueillies sur la population, il existe plusieurs niveaux de description statistique :

- la représentation brute des données,
- des représentations par des tableaux statistiques
- des représentations par des graphiques
- et des résumés numériques en paramètres caractéristiques

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Tableau statistique

Variable	Effectif	Fréquence	Pourcentage	Effectif cumulé croissant	Effectif cumulé croissant
Modalité 1					
Modalité 2					
……..					

- ***L'effectif*** *de la modalité xi désigne le nombre d'individus de la population présentant la modalité xi. Il est noté ni.*
- ***La fréquence*** *de la modalité xi exprime la proportion d'individus présentant la modalité xi. Elle est notée fi et est définie par : fi= ni/ N*

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

- ***Pourcentage*** *de la modalité xi est la fréquence multipliée par 100. Le pourcentage évalue la part de chaque modalité dans l'effectif total*

- ***L'effectif cumulé croissant*** *de la modalité xi est la somme des effectifs de toutes les modalités inférieures ou égales à xi, c'est-à-dire le nombre d'individus de la population pour lesquels le caractère étudié a une modalité inférieure ou égale à xi.*

- ***L'effectif cumulé décroissant*** *de la modalité xi est la somme des effectifs de toutes les modalités supérieures ou égales à xi, c'est-à-dire le nombre d'individus de la population pour lesquels le caractère étudié a une modalité supérieure ou égale à xi.*

Application

On reprend l'exemple où le responsable des ressources humaines étudie la structure de l'ancienneté de ses 2000 salariés. A partir des données brutes recueillies, Le responsable RH découpe l'ancienneté qui s'étale sur [0 ; 35[en 7 classes d'ancienneté toutes de même amplitude. Le tableau ci dessous est le tableau statistique déterminé après le découpage en classes.

Ancienneté	Effectif	Fréquence	Pourcentage	Effectif cumulé croissant	Fréquence cumulée croissante	Effectif cumulé décroissant	Fréquence cumulée décroissante
[0 ; 5[	800	0,4	40%	800	0,4	2000	1
[5 ; 10[	710	0,355	35,5%	1510	0,755	1200	0,6
[10 ; 15[	190	0,095	9,5%	1700	0,850	490	0,245
[15 ; 20[	120	0,06	6%	1820	0,910	300	0,15
[20 ; 25[	90	0,045	4,5%	1910	0,955	180	0,09
[25 ; 30[	50	0,025	2,5%	1960	0,980	90	0,045
[30 ; 35[	40	0,02	2%	2000	1	40	0,02
Total	2000						

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Interprétation

- *710 salariés ont une ancienneté comprise entre 5 ans et 10 ans*
- *1910 salariés ont une ancienneté de moins de 25 ans.*
- *490 salariés ont une ancienneté d'au moins 10 ans (10 ans ou plus).*

Représentations graphiques des séries à une variable

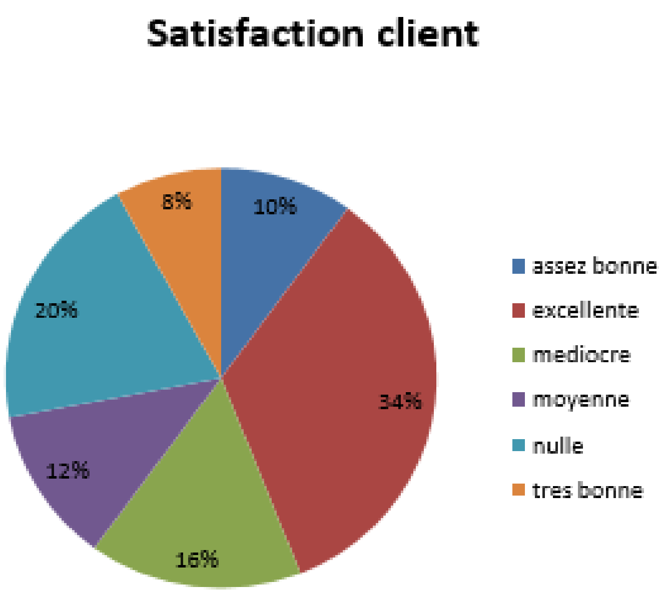
Les graphiques utilisés dépendent de la nature de la variable.

Les variables qualitatives (nominales ou ordinales) – peuvent être représentées au choix à l'aide d'un diagramme circulaire ou à l'aide d'un diagramme en barres.

- Le diagramme circulaire « camembert » est un graphique qui divise un disque en secteur angulaires dont les angles au centre sont proportionnels aux effectifs de chaque modalité. Pour une modalité mi, l'effectif ni, l'angle au centre αi, correspondant est donné (en degré) par : $a_i = \frac{ni}{N} \times 360$
- Un diagramme en barres : c'est un graphique qui à chaque modalité d'une variable qualitative associe un rectangle de base constante dont la hauteur est proportionnelle à l'effectif (ou à la fréquence).

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Les diagrammes circulaire et en barres ci-dessous sont une représentation graphique de la distribution statistique de la variable étudiée qui est « la satisfaction client ».



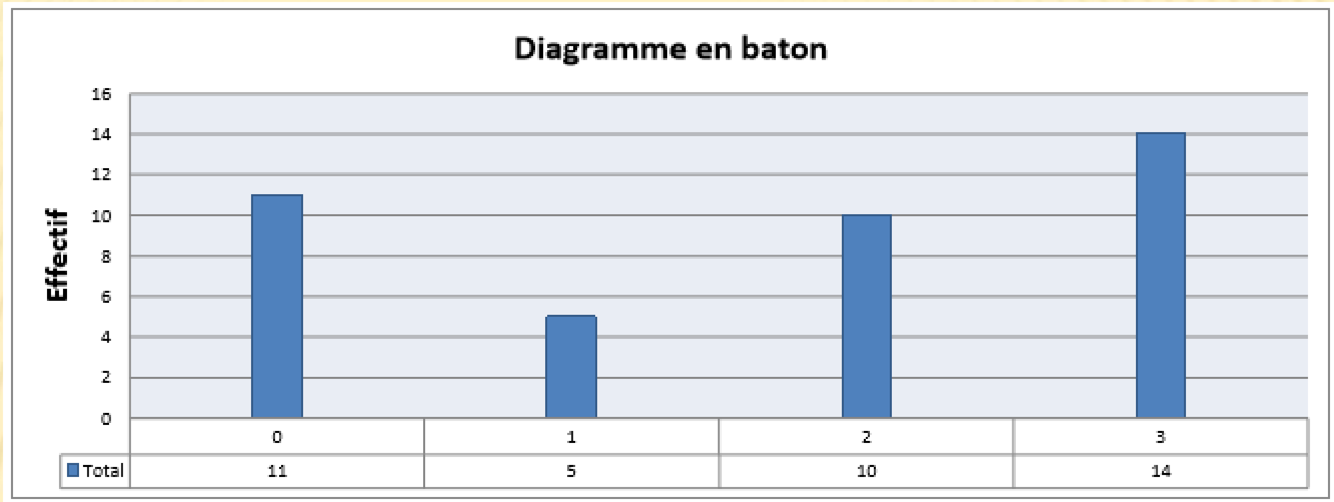
PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

La représentation graphique d'une variable quantitative dépend de sa nature : discrète ou continue.

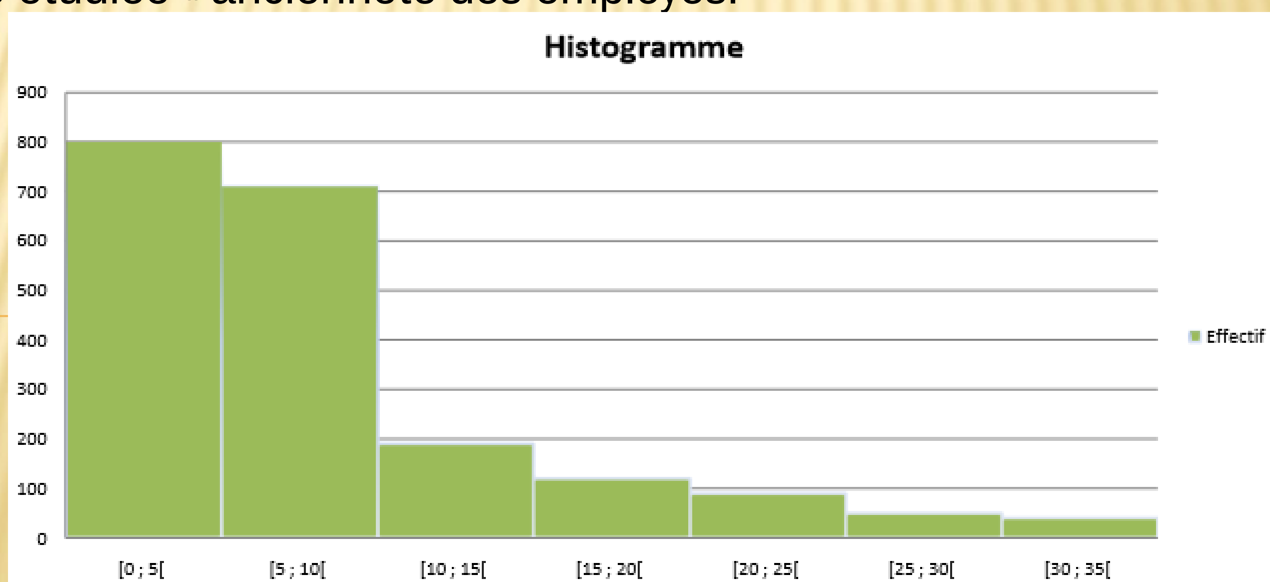
- La distribution d'une variable quantitative discrète est représentée par un diagramme en bâtons (en barres)

- La distribution d'une variable quantitative continue est représentée par un histogramme

La distribution statistique de la variable étudiée « nombre de films regardés » est représentée graphiquement par le diagramme en bâton ci-dessous.



L'histogramme ci-dessous est la représentation graphique de la distribution statistique de la variable étudiée « ancienneté des employés.



PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

IV- paramètres de tendance centrale:

But : Donner une valeur centrale aux donnée

Le mode

Le mode est la valeur la plus fréquente dans une série des donnés.

La moyenne

La moyenne est l'indicateur le plus simple pour résumer l'information fournie par un ensemble de données statistiques : elle est égale à la somme de ces données divisée par leur nombre (moyenne arithmétique).

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Médiane et quantiles

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

La médiane d’une série statistique est une valeur qui partage en deux parties égales l’effectif total de cette série.

La notion de quantile d'ordre p ($0 < p < 1$) généralise la médiane.

On utilise souvent le premier quartile et le troisième quartile

V- paramètres de dispersion:

But : Savoir comment les données varient autour du centre

La variance

La variance sert à mesurer la dispersion, ou l'étalement, d'un ensemble de valeurs autour de leur moyenne. Elle n'est nulle que si tous les individus ont la même valeur pour la variable considérée.

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\longleftrightarrow)^2$$

$$= \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})^2$$

L'écart-type,

égal à la racine carrée de la variance, est souvent proposé comme une mesure des inégalités au sein de la distribution. Plus l'écart-type est faible, plus la population est homogène.

L'étendu

L'étendue est simplement la différence entre la plus grande et la plus petite valeur observée.

E = Xmax - Xmin

Intervalle inter-quantiles

L' intervalle interquartile est la différence entre les valeurs du troisième et du premier quartile : Q3 – Q1

C'est donc l'intervalle qui contient 50% des observations, en laissant 25%à droite et 25% à gauche.

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Coefficient de variation

Naturellement la dispersion autour de la moyenne augmente lorsque la moyenne augmente. On calcule pour t $cv = \frac{\sigma}{\overline{x}}$ e cet effet le coefficient de variation

Exemple : on compare deux groupes des élèves à moyennes différentes

Groupe 1 $\begin{array}{lcl} \overline{x}_1 & = & 11 \\ \sigma_1 & = & 1 \end{array}$ Une grande partie des étudiants a une note comprise entre 10 et12

Groupe 2 $\begin{array}{lcl} \overline{x}_2 & = & 13 \\ \sigma_2 & = & 6 \end{array}$ Une grande partie des étudiants a une note comprise entre 7 et19

Dans ce cas de figure, les étudiants du groupe 2 sont en moyenne meilleurs que ceux du groupe 1 mais l'hétérogénéité est plus forte.

PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

Étude de cas :

Exercice : supposons que l'on veut faire une étude statistique sur les 50 notes attribuées par un jury à l'examen du baccalauréat.

On dispose pour cette étude de la liste des notes obtenues :

2	14	10	16	20	19	7	5	13	14
6	9	16	13	12	3	7	8	18	12
4	8	15	10	8	11	13	9	9	13
5	8	14	5	11	12	2	1	7	1
6	12	3	11	19	17	18	3	0	4

Questions:

1. Déterminez la population étudiée et la variable d'intérêt
2. Précisez la nature de la variable
3. Représentez par un graphique
   - La distribution des effectifs
   - La distribution des fréquences (des pourcentages)
4. Combien d'élèves ont obtenu une note:
   - au plus 5 (inférieur ou égale à 5)
   - au moins 10 (supérieur ou égale à 10)
   - de moins de 12
   - de plus de 18
5. Quelle est la note la plus fréquente?
6. Quelle est la note moyenne de ces 50 élèves?
7. Quelle est la noté médiane?
8. Comment les notes des élèves sont-elles dispersées?

PDF-XChange Editor Click to BUY NOW! www.pdf-xchange.com

On peut regrouper ces notes par ordre croissant : 0,1,1,2,2,3,3,3 ....., et construire le tableau suivant : (dans ce cas la distribution est discrète)

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	2	2	3	2	3	2	3	4	3	2	3	4	4	3	1	2	1	2	2	1

Ou bien regrouper ces notes par intervalle (classe) : (dans ce cas la distribution est continue) Exemple de regroupement par classe :

[0 ; 5[	10
[5 ; 8[	8
[8 ; 12[	12
[12 ; 15[	11
[15 ; 20[	9
	50

Chapitre 2:
Introduction au logiciel SPSS

L’objectif de ce chapitre est de découvrir l’environnement SPSS

SPSS C'EST...

- Saisir, décrire, visualiser les données ...
- Un outil statistique puissant qui vous permet de tester des hypothèses et d'effectuer des analyses simples comme des calculs de moyennes, ou plus complexes comme des analyses multivariées, à partir d'un fichier de données préalablement constitué.

SPSS C'EST…

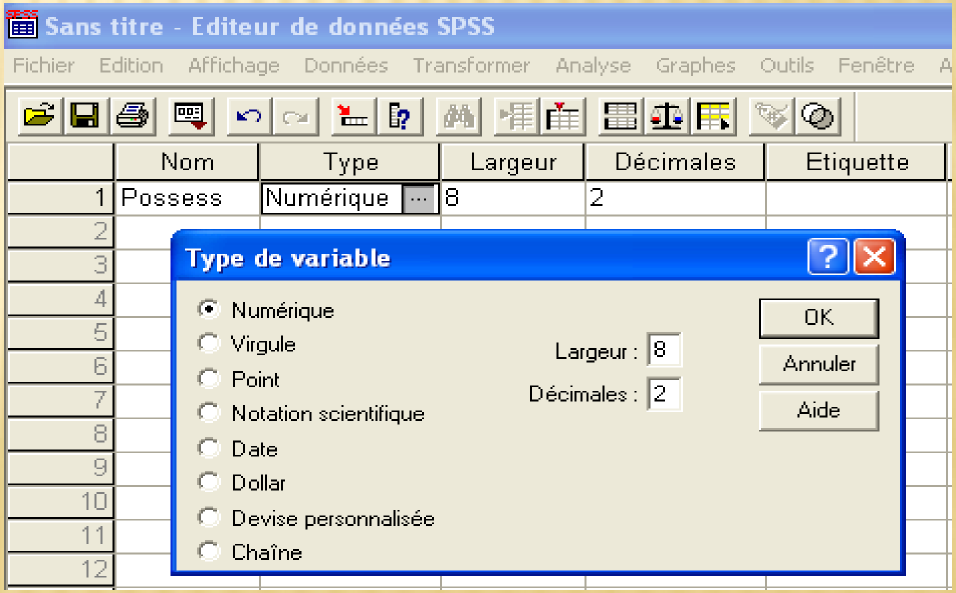
- Capturer toutes les informations dont vous avez besoin sur les attitudes et opinions de vos clients.
- Prévoir les résultats d'interactions avant qu'elles ne se produisent.
- Agir en fonction de ces informations, en intégrant les résultats d'analyse.

NOM DE VARIABLE

- Le nom doit commencer par une lettre.
- Les noms de variable ne peuvent pas se terminer par un point.
- Les espaces et les caractères spéciaux ne peuvent pas être utilisés (par exemple, !, ?,', et *).
- Chaque nom de variable doit être unique ; aucune duplication n'est admise. *Les noms NOUVVAR, NouvVar et nouvvar sont tous identiques.*

TYPES DE VARIABLES

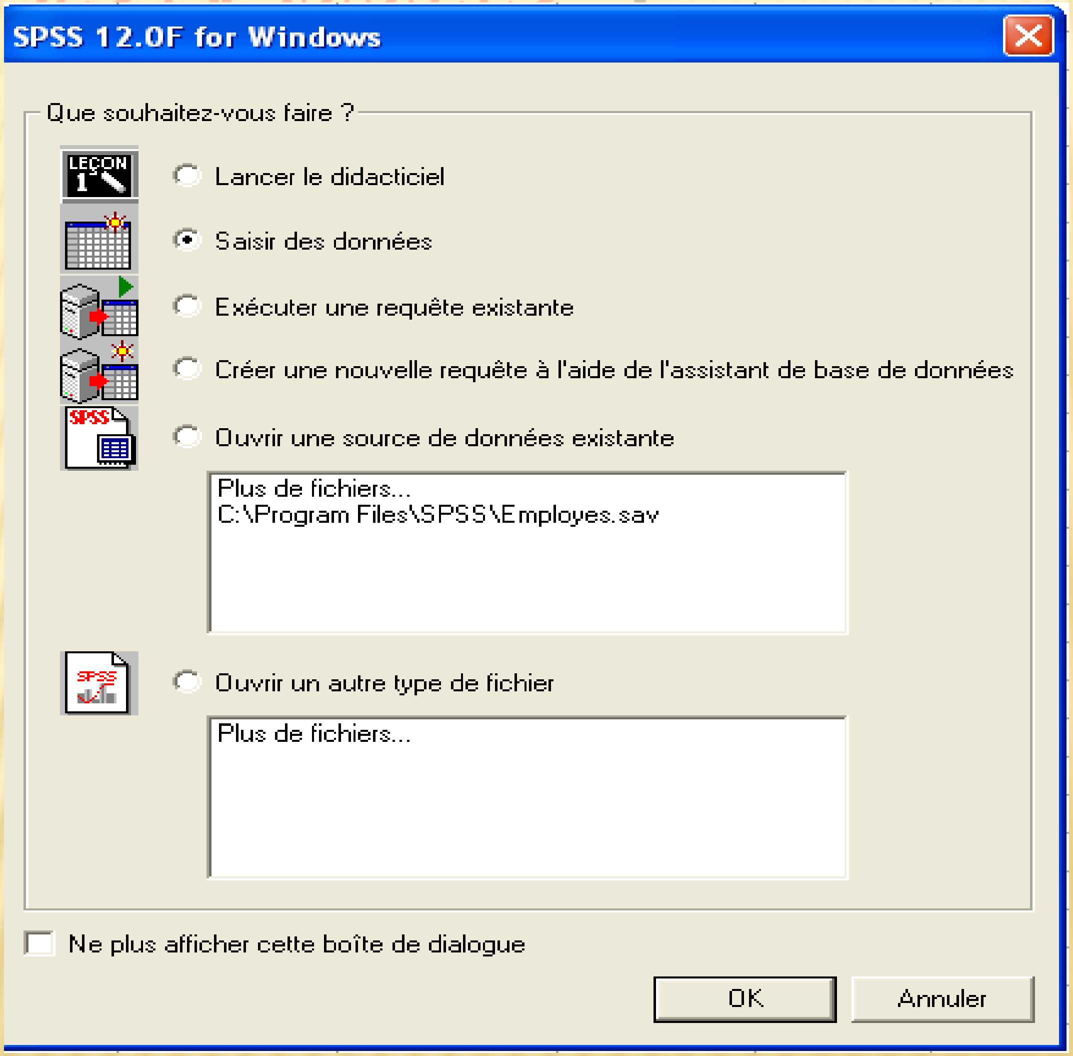
- Numérique
- Virgule (Exemple : 123,425.67)
- Point (Exemple 123.425,6)
- Notation scientifique (1.23E2, 1.23D2, 1.23E+2)
- Date
- Dollar
- Devise personnalisée
- Chaîne.



Niveau de mesure des variables

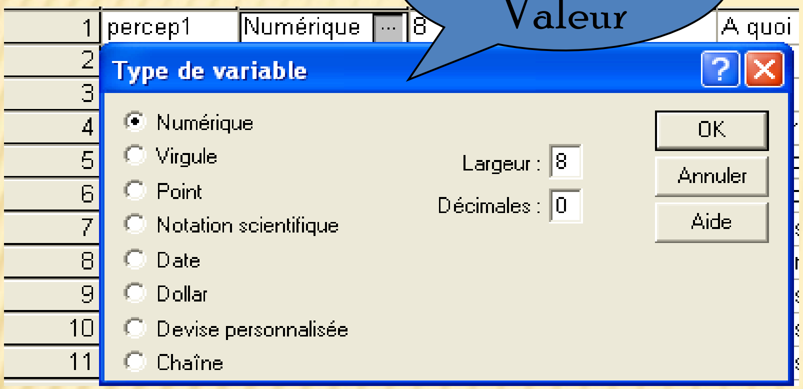
- Nominale : Sexe, CSP, …
- Ordinale : Mention, degré de préférence: *faible, moyen, élevé ; tout à fait d'accord, plutôt d'accord, plutôt pas d'accord, pas du tout d'accord*
- Échelle (*quantitative*) : age, salaire, C.A, …

SAISIE DES DONNÉES -1-

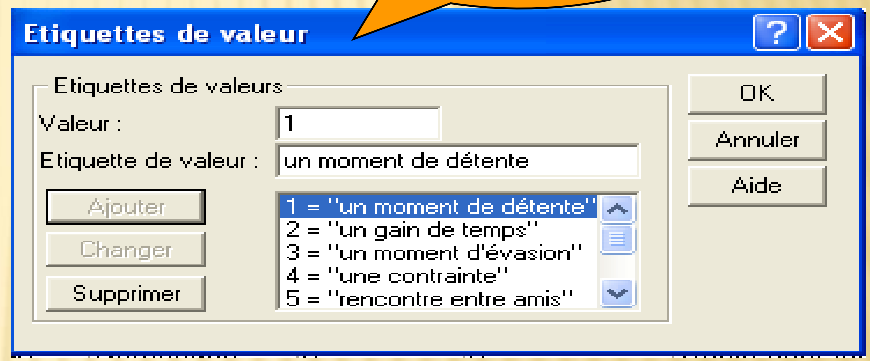


SAISIE DES DONNÉES -2-

Type de Valeur



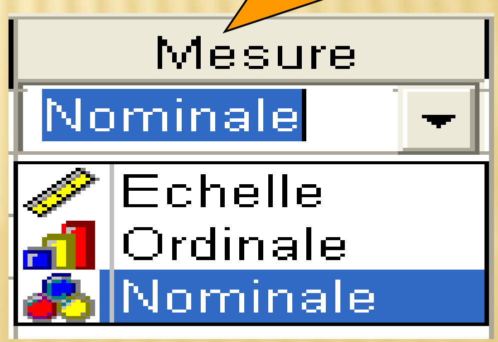
Valeurs



Étiquette

A quoi associez-vous le fait de manger à l'exterieur?

Mesure

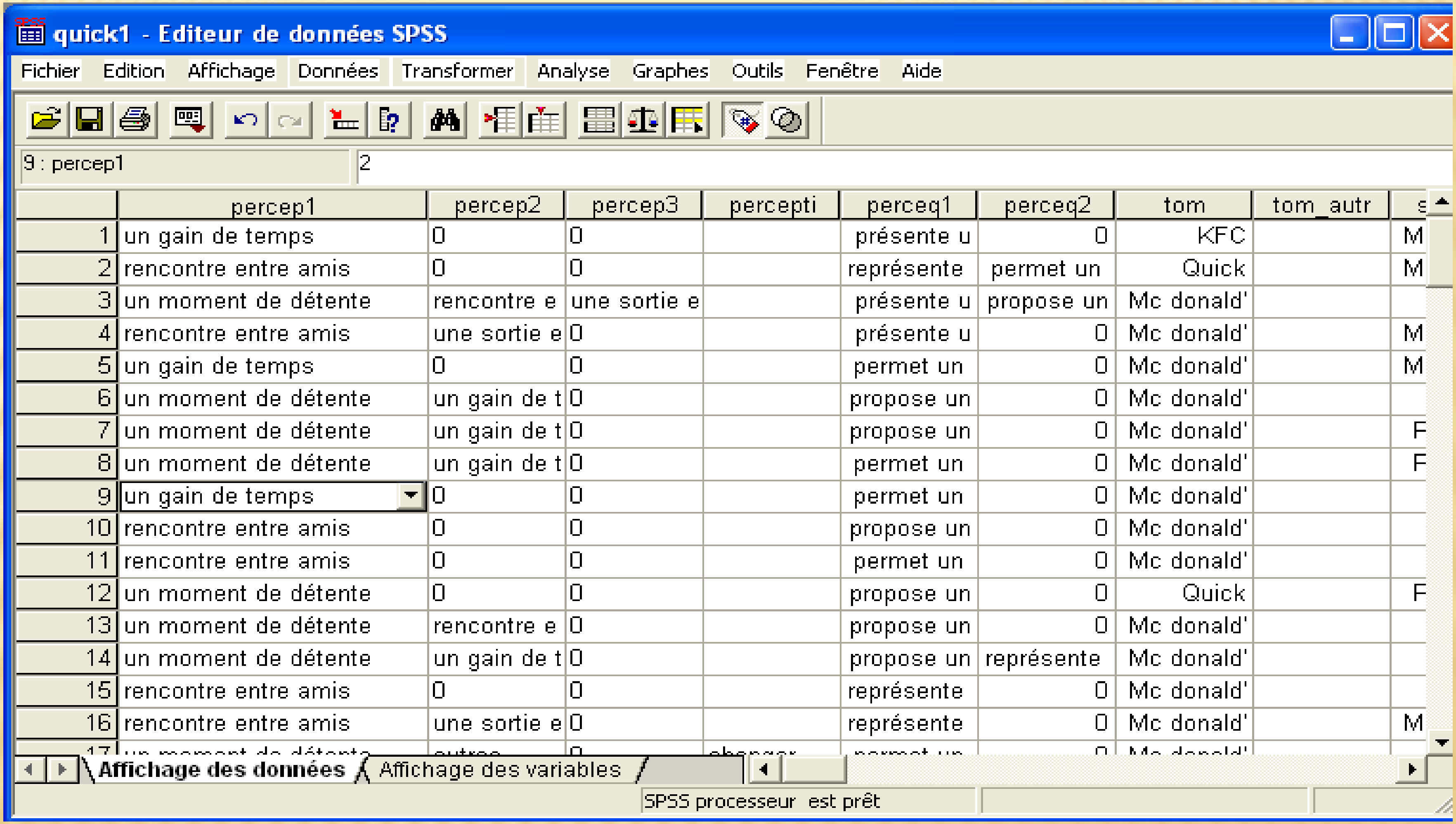


UTILISER SPSS IMPLIQUE LA GESTION DE 4 FICHIERS (FENÊTRES)

- Un fichier d'édition des données (data.sav) : *fenêtre données*
- Un fichier de commandes (syntax.sps) : *fenêtre syntaxe*
- Un fichier de résultats (outpout.spo) *fenêtre résultats*
- Un fichier Script pour personnaliser l'environnement de travail *fenêtre script*

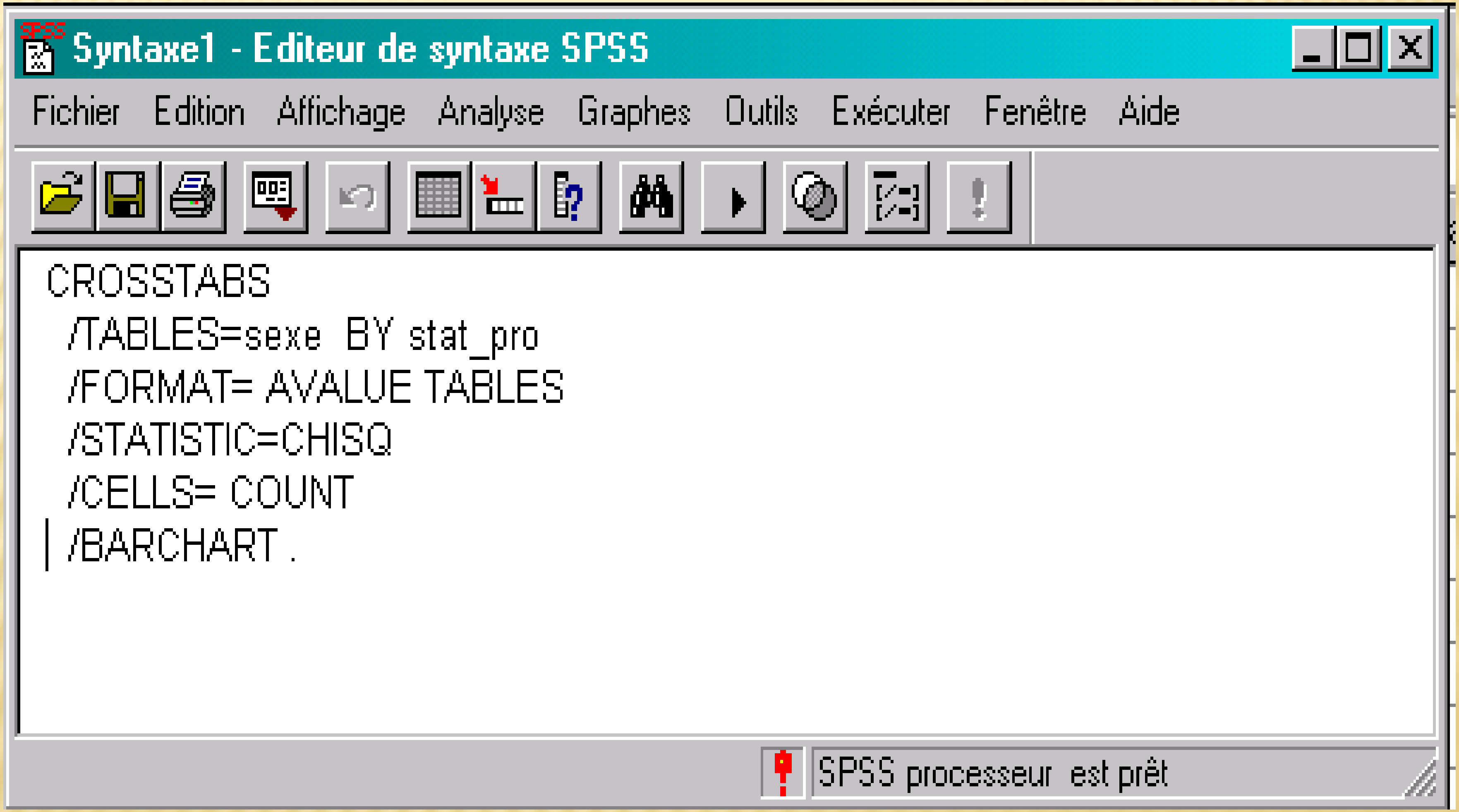
PDF-XChange Editor
Click to BUY NOW!
www.pdf-xchange.com

FENÊTRE DONNÉES

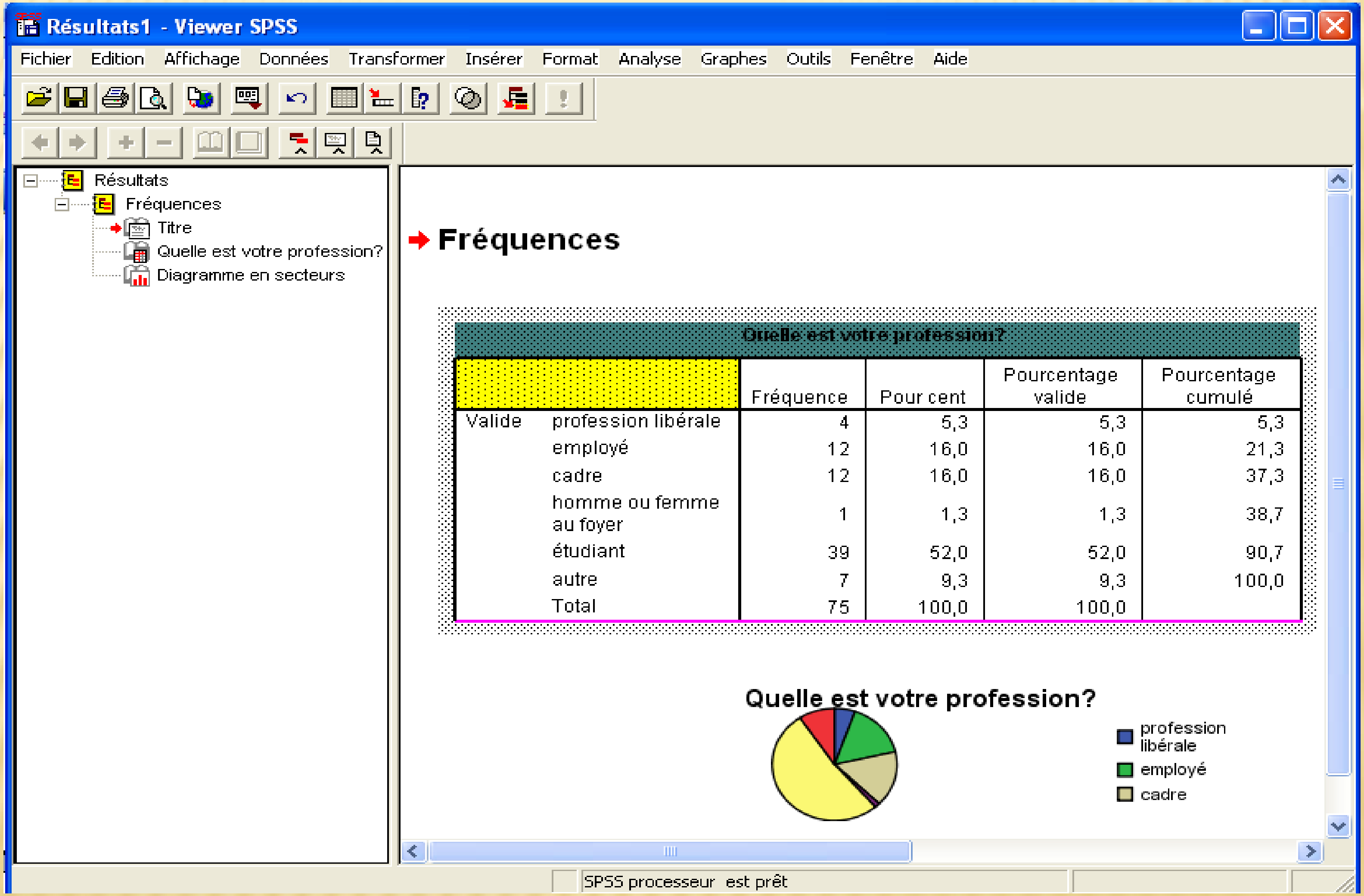


	percep1	percep2	percep3	percepti	perceq1	perceq2	tom	tom_autr	s
1	un gain de temps	0	0		présente u	0	KFC		M
2	rencontre entre amis	0	0		représente	permet un	Quick		M
3	un moment de détente	rencontre e	une sortie e		présente u	propose un	Mc donald'		
4	rencontre entre amis	une sortie e	0		présente u	0	Mc donald'		M
5	un gain de temps	0	0		permet un	0	Mc donald'		M
6	un moment de détente	un gain de t	0		propose un	0	Mc donald'		
7	un moment de détente	un gain de t	0		propose un	0	Mc donald'		F
8	un moment de détente	un gain de t	0		permet un	0	Mc donald'		F
9	un gain de temps	0	0		permet un	0	Mc donald'		
10	rencontre entre amis	0	0		propose un	0	Mc donald'		
11	rencontre entre amis	0	0		permet un	0	Mc donald'		
12	un moment de détente	0	0		propose un	0	Quick		F
13	un moment de détente	rencontre e	0		propose un	0	Mc donald'		
14	un moment de détente	un gain de t	0		propose un	représente	Mc donald'		
15	rencontre entre amis	0	0		représente	0	Mc donald'		
16	rencontre entre amis	une sortie e	0		représente	0	Mc donald'		M
17	un moment de détente	autres	0	changer	permet un	0	Mc donald'		

FENÊTRE SYNTAXE



FENÊTRE RÉSULTATS



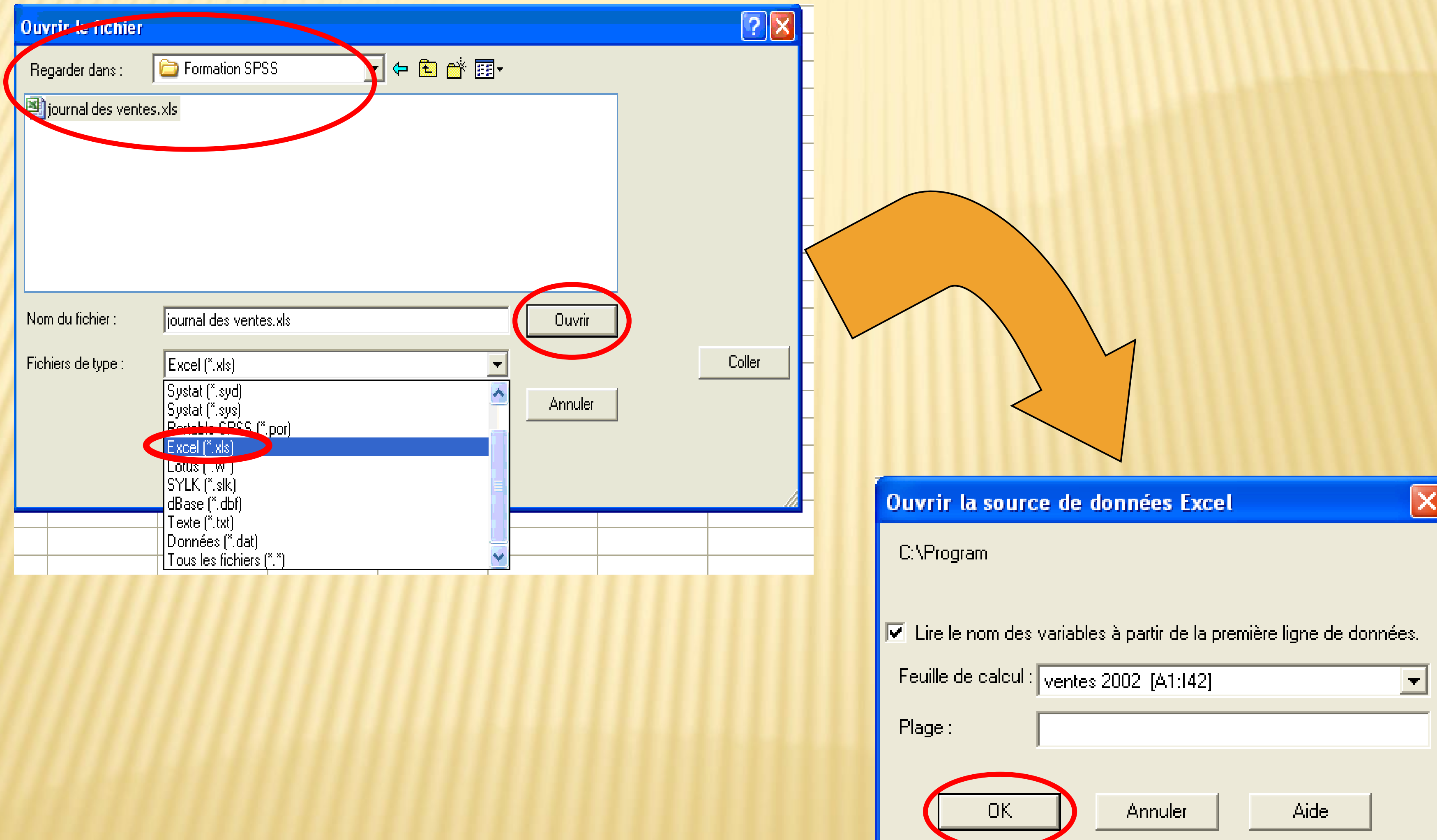
Quelle est votre profession?

		Fréquence	Pour cent	Pourcentage valide	Pourcentage cumulé
Valide	profession libérale	4	5,3	5,3	5,3
	employé	12	16,0	16,0	21,3
	cadre	12	16,0	16,0	37,3
	homme ou femme au foyer	1	1,3	1,3	38,7
	étudiant	39	52,0	52,0	90,7
	autre	7	9,3	9,3	100,0
	Total	75	100,0	100,0	

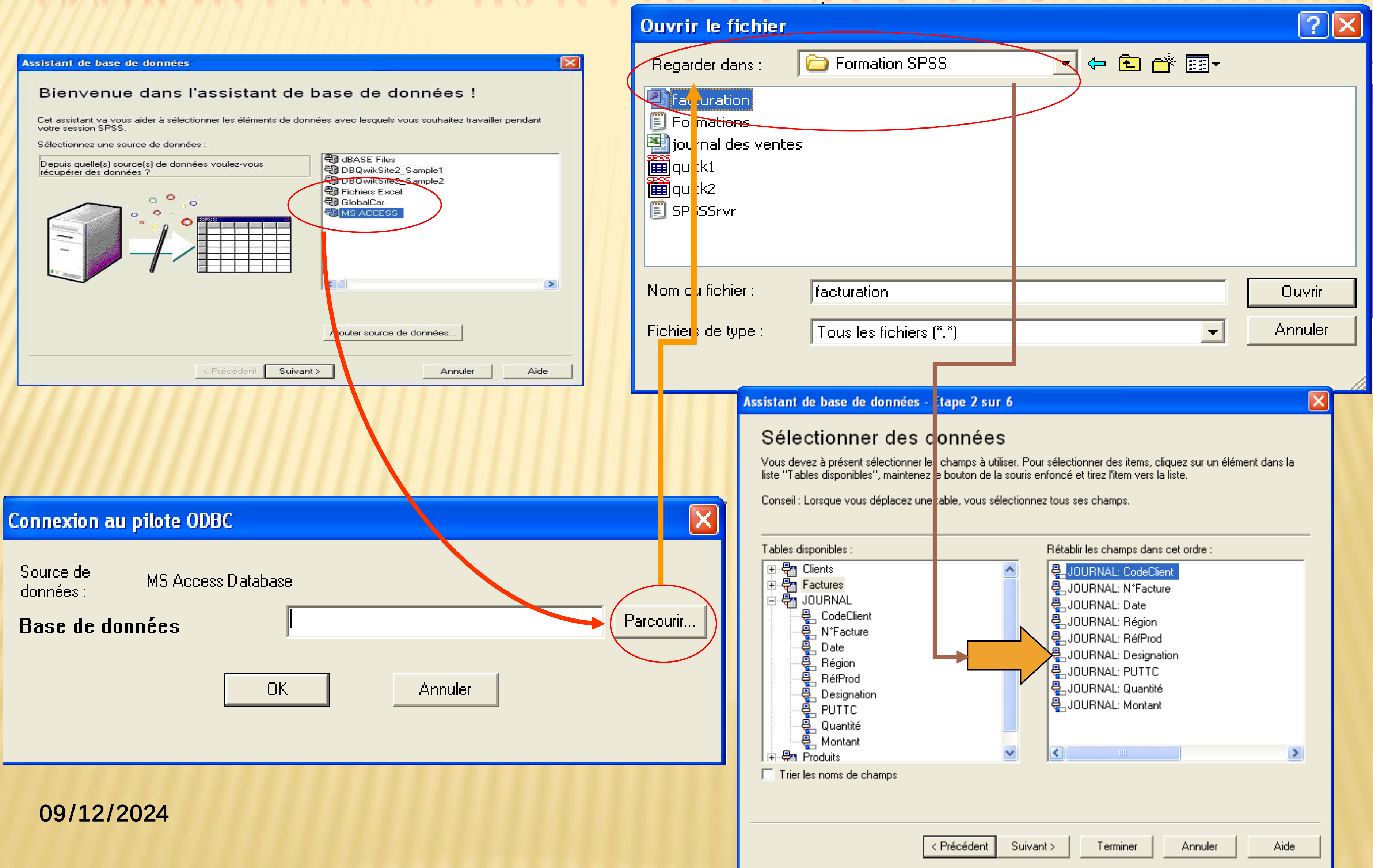
IMPORTER LES DONNÉES

Parfois les données sont disponibles sous différents formats : Fichier Excel; Base de données Access ; Fichier texte formaté...ou autres. SPSS est compatible avec un certain nombre de formats. Nous allons nous intéresser ici aux formats suivants : Fichier Excel, Base de données Access et Fichier Texte Formaté.

IMPORTER À PARTIR D'EXCEL



IMPORTER À PARTIR D'ACCESS



09/12/2024