

Analyse des données quantitatives

I

Cours 2

Notions fondamentales: éléments constitutifs d'une base de données

Plan

1. Déroulement des cours
2. Éléments constitutifs d'une base de données
 1. Qu'est-ce qu'une base de données
 2. Variables et valeurs
 3. Types de variables
 4. Le codebook
3. La base de données secondaires PSM
4. Le codebook PSM

1. Déroulement du cours

1) Avec les vraies données secondaires

Avantages

- Données existantes, bonne qualité
- Gain de temps: concentration sur procédures

Désavantages

- Coupure avec les techniques de l'approche quantitative, notamment la construction et utilisation d'un questionnaire

1. Déroulement du cours

2) Recours à un logiciel spécialisé

- SPSS: **S**tatistic **P**ackage for the **S**ocial **S**ciences

Avantages

- L'utilisation facilitée des procédures
- Priorité aux « logiques » des procédures *cf.* expertises mathématiques
- Acquisition d'une nouvelle compétence

1. Déroulement des cours avec SPSS

2) Utilisation de l'SPSS

- a) Collective*: introduction aux procédures à l'aide du beamer & du manuel (gestens)
 - b) Individuelle* (pendant cours): répétition des procédures avec son propre portable
 - c) Individuelle* ou en *petits groupes*: pratique des procédures via exercices
- Exigence**: avoir en main les docs: codebook, manuel

1. Déroulement des cours

Concrètement...

- 1) Obtenir SPSS via Micromus
- 2) Accéder aux données, documents, manuels
 - via gestens & dossier protégé si inscrit au cours
 - instructions pour accéder au dossier protégé: gestens
- 3) Faire les quatre exercices (facultatif)
 - a) Détails et dates: gestens

2. Éléments constitutifs d'une base de données

- 1) Qu'est-ce qu'une base de données?
- 2) Variables et valeurs
- 3) Types de variables
- 4) Le codebook

LIEN cours Techniques: l'élaboration d'un questionnaire

2. Éléments d'une base de données

1) Définition

- La **base de données** englobe l'ensemble des données provenant d'une enquête
- Une base de données englobe des indications portant sur:
 - **les répondants** (un échantillon d'une population): informations sociodémographiques: âge, sexe, nationalité, formation...
 - **leurs réponses**: selon le thème de l'enquête: travail, loisirs, attitudes...

2. Éléments d'une base de données

2) Variables et valeurs

Une **variable** correspond à une question; et
(plus précisément) au thème de la question

Les **valeurs** correspondent aux réponses
possibles

Une **variable** est donc un indicateur (indice)
dans l'enquête dont les **valeurs** permettent
de distinguer les réponses contrastées des
répondants

2. Éléments d'une base de données

2) Variables et valeurs

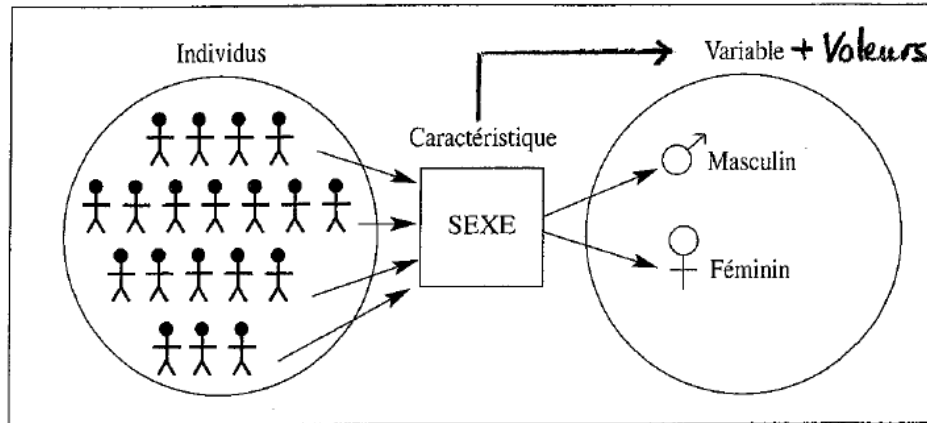
- Exemples...

Variable: Thème	Variable: Question	Valeurs
Sexe	De quel sexe êtes-vous?	1. Homme 2. Femme
Âge	Quel âge avez-vous?	--- ans 9. Pas de réponse
Niveau de formation NivFormat	Quel est le niveau de votre formation (dernière qualification obtenue?)	1. École obligatoire 2. Certificat d'apprentissage 3. Collège 4. HES 5. BA, MA 6. Doctorat 7. Autre 8. N/A 9. Pas de réponse

Variable: Thème	Variable: Question	Valeurs
Loisirs	Quels loisirs pratiquez-vous?	1. Les sports d'hiver 2. Les randonnées à pied 3. Le tennis 4. Les sorties dans les bars 5. Le shopping 6. Autre(s) 7. Je n'ai pas de loisirs
Sentiment de sécurité	Que pensez-vous de la sécurité en Suisse?	1. Niveau très élevé 2. Plutôt bon niveau 3. Plutôt bas niveau 4. Absence de sécurité 5. Je ne sais pas 9. Pas de réponse
Satisfaction de vie	Êtes-vous globalement satisfait avec votre vie?	Réponses selon échelle allant de 0 à 10 0 = pas de tout satisfait 10 = complètement satisfait

Alonzo, p. 17

Les variables sont les outils privilégiés du chercheur. Il y a recours constamment. Il est donc fondamental de les définir. Une variable est une caractéristique qui prend des formes différentes pour chaque unité statistique observée. Le mouvement est donc le suivant :



Les caractéristiques des individus portent le nom de variables parce qu'elles peuvent varier d'un individu à l'autre. Le sexe, la profession, l'âge, le revenu, la situation matrimoniale, le niveau de diplôme et l'opinion politique sont des exemples de variables

2. Éléments d'une base de données

3) Types de variables

- Les variables **numériques**
- Les variables **catégorielles**
- Les variables **ordinales**

2. Éléments d'une base de données

3) Types de variables

Les variables numériques:

- Quantitatives, mesurables, continues, discrètes
- Les valeurs des variables numériques mesurent la **magnitude** (grandeur) de la réponse:
 - âge, poids, revenu, nombre de langues parlées, échelles de satisfaction si au moins 11 mesures
- Les variables numériques peuvent être transformées en **classes**:
 - classes d'âges: 20-29, 30-39, 40-49, 50-59, 60-79, 80+
 - classes de revenus

2. Éléments d'une base de données

3) Types de variables

Les variables catégorielles

- Qualitatives, ordinales
- Les valeurs des variables catégorielles constituent un ensemble de catégories. Elles permettent de **qualifier** les sujets selon les catégories proposées:
 - état-civil: marié, célibataire, veuf, divorcé
 - taux de travail: plein-temps, temps partiel
- Les valeurs **ordinales** correspondent à un cas spécifique de variables catégorielles

2. Éléments d'une base de données

3) Types de variables

Les variables **ordinales**

- Les catégories forment un ordre « *logique* »
 - Exemple de niveau de satisfaction
 - très satisfait
 - plutôt satisfait
 - plutôt pas satisfait
 - pas du tout satisfait
 - Exemple: le niveau de formation:
 - secondaire, secondaire supérieur, tertiaire HES, universitaire (quoique...)
 - BA, MA, doctorat

2. Éléments d'une base de données

3) Types de variables

Selon la **fonction** pour l'analyse

- Les variables sociodémographiques
 - *Qui?* Femmes, célibataires, âgées de 20-29 ans
- Les variables de « contenu »
 - *Quoi?* Satisfaction avec système politique

2. Éléments d'une base de données

3) Types de variables

**Pourquoi
faut-il distinguer
entre les types de
variables?**

2. Éléments d'une base de données

4) Codebook

Le **codebook** (livret de codes): rassemble les instructions permettant de désigner les variables et les valeurs en « chiffres »

Q 3 Quel est votre état civil?

1. Célibataire
2. Marié-e
3. Divorcé-e
4. Veuf-ve

2. Éléments d'une base de données

4) Codebook

Q 4 Quel âge avez-vous?

--- (3 places)

Q 21 Quel est votre revenu brut annuel?

----- (9 places)

[DataSet1] /Users/kerrie/Documents/SHP04 Version StatsI 2008_1.SAV

File Information : Variable Information

Variable	Position	Label
IDPERS	1	Identification number of person
sex	2	Sex
AGE04	3	Age in year of interview
COHAST04	4	Cohabitor Status
CIVSTA04	5	Civil status in year of interview
PERMIT04	6	Foreigners residence permit
P04D167	7	Residential permit
P04C44	8	Satisfaction with life in general
P04L101	9	Feeling of security
P04L100	10	Satisfaction with personal security
P04C01	11	Health status
P04C01A	12	Health in general
P04C02	13	Satisfaction with health status
P04C05A	14	Health problems: Weakness, weariness: Last 4 weeks
P04C19A	15	Chronic illness or long-term health problem

Valeurs - Étiquettes					
1	IDPERS	0		9	P04L101
				-8(a)	other error
				-7(a)	filter error
				-3(a)	inapplicable
2	sex	-8(a)	other error	-2(a)	no answer
		-7(a)	filter error	-1(a)	does not know
		-3(a)	inapplicable	1	Very secure
		-2(a)	no answer	2	Relatively secure
		-1(a)	does not know	3	Fairly insecure
		1	man	4	Very insecure
		2	woman		
3	AGE04	-8(a)	other error	10	P04L100
		-7(a)	filter error	-8(a)	other error
		-3(a)	inapplicable	-7(a)	filter error
		-2(a)	no answer	-3(a)	inapplicable
		-1(a)	does not know	-2(a)	no answer
				-1(a)	does not know
				0	not at all satisfied
				10	completely satisfied
4	COHAST04	-8(a)	other error	11	P04C01
		-7(a)	filter error	-8(a)	other error
		-3(a)	inapplicable	-7(a)	filter error
		-2(a)	no answer	-3(a)	inapplicable
		-1(a)	does not know	-2(a)	no answer
		1	married	-1(a)	does not know
		2	not married	1	very well
				2	well
				3	so, so (average)
5	CIVSTA04	-8(a)	other error	4	not very well
		-7(a)	filter error	5	not well at all
		-3(a)	inapplicable		

2. Éléments d'une base de données

4) Codebook

En consignant les instructions dans un code book à disposition de tous les membres d'une équipe de recherche, on s'assure que:

- Toutes les données seront entrées selon une procédure uniforme
- Les décisions ne seront pas « oubliées »
- La base de données peut être utilisée par d'autres équipes de recherche

3. L'enquête *Panel suisse des ménages* (PSM)

- Voir: <http://www.swisspanel.ch>
- Objectif principal:
 - « Observer le changement social, notamment la dynamique de l'évolution des conditions de vie de la population en Suisse »

3. L'enquête *Panel suisse des ménages* (PSM)

Thèmes en rapport avec les modes de vie (Vivre en Suisse)

- Éducation et travail rémunéré
- Revenu et situation financière
- Santé physique et psychique
- Organisation du ménage et partage des tâches
- Vie associative, religion et loisirs
- Partenariats et vie familiale
- Attitudes, opinions politiques et valeurs

4. Les documents PSM

- Accessibilité aux données et au codebook PSM
- Confidentialité assurée via dossier protégé
- Informations sur accès au dossier protégé:
gestens