

Statistiques et informatique

appliquées aux sciences sociales

Construction et utilisation d'un tableau croisé

Mardi 26/9/2023 15h30-17h Amphi D2 Censier

Licence de sciences sociales 3^e année
Université Paris 1 Panthéon Sorbonne

Plan de la séance

- Présentation de l'exemple (reprise)
- Variables catégorielles
- proportions, tris à plat et tris croisés
- % lignes et colonnes
- Utilisation des marges
- Lecture des résultats
- Commentaire des résultats

Rappels

- Données
- Présentation : notion de population-cible, individus statistiques
- Notion de variable, types de variables
- Comprendre et utiliser un certain nombre de codes

L'enquête JobMob

Thème : modes de vie, articulation des temps et focus sur des formes de mobilité liées au travail et exerçant des contraintes fortes sur la vie de couple et la vie de famille (temps de trajet, nuitées, déménagement récent, domiciles distincts au sein d'un couple)

Année : 2007

Pays étudiés : 6 pays européens dont la France (ainsi qu'Allemagne, Belgique, Espagne, Pologne, Suisse)

Population cible : individus adultes entre 25 et 54 ans (âges actifs), ayant une ligne (tél) fixe

Type d'échantillon : aléatoire (enquête par sondage)

w_equal	w_mob_equa	employed	working	Mob	mob1	mob2	mob3	mob4	mob5	mobili_01a	country	gender	edu0	genderedu	age10	genderage
0	0,0000	0,00	1,00	0,0	0	0	0	0	0	6,0	1,0	0,0	5,0	3,0	3,0	3,0
1	0,0	0,00	0,00	0,0	0	0	0	0	0	8,0	1,0	1,0	2,0	11,0	1,0	11,0
1	0,0	1,00	1,00	0,0	0	0	0	0	0	8,0	1,0	1,0	2,0	11,0	2,0	12,0
0	0,0	0,00	1,00	0,0	0	0	0	0	0	6,0	1,0	1,0	3,0	12,0	3,0	13,0
0	0,0	0,00	0,00	0,0	0	0	0	0	0	8,0	1,0	1,0	3,0	12,0	3,0	13,0
1	0,0	1,00	1,00	0,0	0	0	0	0	0	8,0	1,0	1,0	1,0	10,0	2,0	12,0
1	0,0	1,00	1,00	0,0	0	0	0	0	1	6,0	1,0	0,0	2,0	1,0	1,0	1,0
0	0,4508421054	1,00	1,00	1,0	1	0	0	0	0	1,0	1,0	1,0	3,0	12,0	3,0	13,0
0	0,3952618747	1,00	1,00	1,0	1	0	0	0	0	1,0	1,0	1,0	3,0	12,0	2,0	12,0
1	0,0	1,00	1,00	0,0	0	0	0	0	0	6,0	1,0	1,0	2,0	11,0	2,0	12,0
0	0,0	0,00	1,00	0,0	0	0	0	0	0	8,0	1,0	0,0	5,0	3,0	2,0	2,0
0	0,0	1,00	1,00	0,0	0	0	0	0	0	6,0	1,0	0,0	5,0	3,0	2,0	2,0
1	0,0	1,00	1,00	0,0	0	0	0	0	0	8,0	1,0	1,0	2,0	11,0	3,0	13,0

Nom	country		
Titre	survey country	pays enquêté	
Format	F4.0		
Mesure	Échelle		
Affiche la largeur		8	
Valeur	Étiquette	Traduction	
	1 Germany	Allemagne	1:Allemagne
	2 France	France	2:France
	3 Spain	Espagne	3:Espagne
	4 Switzerland	Suisse	4:Suisse
	5 Poland	Pologne	5:Pologne
	6 Belgium	Belgique	6:Belgique

Nom	gender		
Titre	sex of respondent	sexe de la personne interrogée	
Format	F8.2		
Mesure	Échelle		
Affiche la largeur	10		
Valeurs manquantes	9000 THRU HIGHEST		
Valeur	Étiquette	Traduction	
0	male	homme	0:homme
1	female	femme	1:femme
9990	Not asked ?	Non demandé (?)	9990:Non demandé (?)
9999	Don't know	Ne sait pas	9999:Ne sait pas

Nom	age10	
Titre	age / cohort, in 10 year categories	tranches d'âge décennales
Format	F4.0	
Valeur	Étiquette	
1	25-34	
2	35-44	
3	45-54	
9990	undocumented missing (autre manquante)	
9998	Don't want to say (ne souhaite pas répondre)	
9999	Don't know (ne sait pas)	

Nom	CarOr	
Titre	Career Orientation	
traduction	importances respectives pour l'enquêté·e de sa carrière professionnelle et de sa famille	
Format	F8.2	
Mesure	Échelle	
Affiche la largeur	10	
Valeur	Étiquette	Traduction
1	career +	Privilégie sa carrière professionnelle
2	career=family	Accorde une égale importance à sa famille et à sa carrière professionnelle
3	Family +	Privilégie sa famille
9	DK or DWTS	Ne sait pas ou ne souhaite pas répondre

Quelques précisions sur la question

- 18 parties, la seizième s'intitulant « attitudes regarding job, job mobility, family » (questions 14.01 à 14.06.03)

14.03 v1403

Filter rules: –

Do you consider yourself to be more family-oriented or more career-oriented?

***[if clarification is necessary:]* Independent from the hours you actually spend with paid work or with housework family, with which of the two would you like to spend the bigger share of the day with if you could choose?**

[do not read out answer categories]

- (1) more family-oriented
- (2) more career-oriented
- (3) both equally
- (9998) don't want to say
- (9999) don't know

Extrait du questionnaire (version française)

Question 14.03

Variable v1403

Filter rules: –

Vous considérez-vous comme quelqu'un qui aurait plutôt tendance à privilégier sa vie de famille ou sa vie professionnelle ?

[Si des éclaircissements sont nécessaires :] Indépendamment des heures que vous consacrez dans la réalité à vous occuper de votre travail rémunéré et à vous occuper de votre famille et des tâches domestiques, quelle est la part de votre temps que vous aimeriez consacrer à l'un ou l'autre si vous aviez le choix ?

Ne pas citer

(1) vie de famille

(2) carrière professionnelle

(3) tout autant l'un que l'autre

(9998) ne veut pas répondre

(9999) ne sait pas

Détail du codage de la variable CarOr

#Syntaxe SPSS (utilisable sous PSPP) :

```
recode v1403 (2=1) (3=2) (1=3) (9998,9999=9) into  
CarOr.
```

```
var lab CarOr "Career Orientation".
```

```
val lab CarOr
```

```
1 "career +" 2 "career=family" 3 "Family +" 9 "DK or  
DWTS".
```

#Remarque : l'ordre des modalités est susceptible d'avoir favorisé la première d'entre elles (la famille)

Tableau croisé dynamique (brut)

TraitementJobmob1.ods - LibreOffice Calc

Fichier Édition Affichage Insertion Format Styles Feuille Données Outils Fenêtre Aide

Calibri 11 G I S A % 00 7 Supprimer des cellules 00 00

H12 fx Σ =

	A	B	C	D	E	F	G
1	country	2,0000					
2							
3	Somme - w_equal	CarOr					
4	gender	1,0000	2,0000	3,0000	9,0000	Total Résultat	
5	0,0000	64	49	329	1	442	
6	1,0000	56	48	352		456	
7	(vide)	2				2	
8	Total Résultat	122	96	681	1	900	
9							
10							
11							
12							
13							
14							
15							
16							
17							

jobmob1 tcd Tableau QuelquesVariables

Rechercher Tout rechercher Affichage mis en forme Respecter la casse

Feuille 2 sur 4 Par défaut Français (France) Moyenne: ; Somme: 0 160%

Tris « à plat »

genre	Effectifs	%
0:homme	442	49,2 %
1:femme	456	50,8 %
Ensemble	898	100,0 %

$$\begin{aligned}
 \text{proportion} &= \frac{\text{effectif des hommes}}{\text{effectif total}} \\
 \text{d'hommes} &= \frac{442}{898} \approx 0,492 = 49,2\%
 \end{aligned}$$

« pour tout »

$$p_i = \frac{n_i}{n}, \forall i=1,2,3$$

$$p_1 + p_2 + p_3 = 1 = 100\%$$

$$\sum_{i=1}^3 p_i = 1 = 100\%$$

$$p_1 = \frac{n_1}{n} = \frac{120}{898} \approx 13,4\%$$

Priorités entre carrière et vie familiale	Effectif	%
1:Privilégie sa carrière professionnelle	120	13,4%
2:Accorde une égale importance...	97	10,8%
3:Privilégie sa famille	681	75,8%
Ensemble	898	100,0%

Calcul d'une proportion dans un tableur

The image shows a spreadsheet with the following data:

	A	B	C
1	genre		%
2	0:homme	442	49,2%
3	1:femme	456	50,8%
4	Ensemble	898	100,0%
5			
6	Priorités entre carrière et vie familiale	Effectif	%
7	1:Privilégie sa carrière professionnelle	120	13,4%
8	2:Accorde une égale importance...	97	10,8%
9	3:Privilégie sa famille	681	75,8%
10	Ensemble	898	100,0%
11			

Construction d'un tableau

On dispose généralement (par convention) :

- en ligne, une variable déjà connue, par exemple de type socio-démographique, considérée comme « explicative »
- en colonne, une variable nouvelle, sur une opinion, attitude, valeur, pratique, nouvelle, considérée comme « expliquée » ou « d'intérêt »

Le tableau ainsi obtenu ressemble quelque peu à une version simplifiée du tableau *brut* initial, en substituant aux individus statistiques les groupes socio-démographique distingués par la variable explicative, en gardant en colonne les modalités de la variable d'intérêt et en indiquant dans chaque case l'effectif du sous-groupe présentant deux caractéristiques à la fois.

Tableau d'effectifs (« tri croisé »)

Effectifs	Priorités entre carrière et vie familiale			
genre	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille	Ensemble
0:homme	64	49	329	442
1:femme	56	48	352	456
Ensemble	120	97	681	898

Nombre d'hommes privilégiant leur carrière professionnelle : 64 (notation : n_{01})

Nombre total d'hommes : 442 (notation des diapositives suivantes : n_0 .)

Nombre total de personnes privilégiant leur carrière professionnelle : 120 (notation : $n_{.1}$)

Effectif total : 898 (notation : n)

Structure du tableau : le total de chaque ligne est sur sa dernière colonne (marge en ligne)

Le total de chaque colonne est sur sa dernière ligne (marge en colonne)

Le total des sous-totaux en ligne ou en colonne redonne bien le total global (898)

Tableau des % du total

% du total	Priorités entre carrière et vie familiale			
	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille	Ensemble
Genre				
0:homme	7,1%	5,5%	36,6%	49,2%
1:femme	6,2%	5,3%	39,2%	50,8%
Ensemble	13,4%	10,8%	75,8%	100,0%

proportion d'hommes «carriéristes» = $\frac{\text{nb hommes privilégiant...}}{\text{Effectif total}} = \frac{64}{898} \approx 0,071 = 7,1\%$
 parmi l'ensemble de la population

$$p_{ij} = \frac{n_{ij}}{n}, \forall i=0,1, j=1,2,3 \text{ avec } \sum_{i=0}^1 \sum_{j=1}^3 p_{ij} = \sum_{j=1}^3 \sum_{i=0}^1 p_{ij} = 100\%$$

Calcul d'un % du total dans un tableur

Formula bar: $=B3/ΣE\$5$

	A	B	C	D	E
1	Effectifs	Priorités entre carrière et vie familiale			
2	genre	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille	Ensemble
3	0:homme	64	49	329	442
4	1:femme	56	48	352	456
5	Ensemble	120	97	681	898
6					
19	% du total	Priorités entre carrière et vie familiale			
20	genre	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille	Ensemble
21	0:homme	7,1%	5,5%	36,6%	49,2%
22	1:femme	6,2%	5,3%	39,2%	50,8%
23	Ensemble	13,4%	10,8%	75,8%	100,0%

Annotations: A red circle highlights the formula bar. A green circle highlights the value 64 in cell B3. A blue circle highlights the value 898 in cell E5. A green division symbol is shown in cell E2, with arrows pointing to the 64 in B3 and the 898 in E5.

Tableau des pourcentages en ligne

Effectifs	Priorités entre carrière et vie familiale			
	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille	Ensemble
0:homme	14,5%	11,1%	74,4%	100,0%
1:femme	12,3%	10,5%	77,2%	100,0%
Ensemble	13,4%	10,8%	75,8%	100,0%

proportion de «carriéristes» parmi les hommes $= \frac{\text{nb hommes priv...}}{\text{total des hommes}} = \frac{64}{442} \approx 0,145 = 14,5\%$

$$p_{ij|i} = \frac{n_{ij}}{n_i}, \forall i=0,1, j=1,2,3 \text{ avec } \sum_{j=1}^3 p_{ij|i} = p_{i1|i} + p_{i2|i} + p_{i3|i} = 1 = 100\% \quad \forall i=0,1$$

« pour tout »
↓

Calcul d'un % ligne dans un tableur

fx Σ = $=B3/ΣE3$

	A	B	C	D	E
1	Effectifs	Priorités entre carrière et vie familiale			
2	genre	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille	Ensemble
3	0:homme	64	49	329	442
4	1:femme	56	48	352	456
5	Ensemble	120	97	681	898
6					
7	% en ligne	Priorités entre carrière et vie familiale			
8	genre	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille	Ensemble
9	0:homme	14,5%	11,1%	74,4%	100,0%
10	1:femme	12,3%	10,5%	77,2%	100,0%
11	Ensemble	13,4%	10,8%	75,8%	100,0%

Tableau des % en colonnes

% en colonne	Priorités entre carrière et vie familiale			
genre	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille	Ensemble
0:homme	53,3%	50,5%	48,3%	49,2%
1:femme	46,7%	49,5%	51,7%	50,8%
Ensemble	100,0%	100,0%	100,0%	100,0%

proportion d'hommes parmi les «carriéristes» = $\frac{\text{nb hommes privilégiant...}}{\text{total des «carriéristes»}} = \frac{64}{120} \approx 53,3\%$

$$p_{ij|j} = \frac{n_{ij}}{n_{.j}}, \forall i=0,1, j=1,2,3 \text{ avec } \sum_{i=0}^1 p_{ij|j} = p_{0j|j} + p_{1j|j} = 100\% \forall j=1,2,3$$

Calcul d'un % colonne dans un tableur

B15 fx Σ = **=B3/B\$5**

	A	B	C	D	E
1	Effectifs	Priorités entre carrière et vie familiale			
2	genre	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille	Ensemble
3	0:homme	64	49	329	442
4	1:femme	56	48	352	456
5	Ensemble	120	97	681	898
6					
13	% en colonne	Priorités entre carrière et vie familiale			
14	genre	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille	Ensemble
15	0:homme	53,3%	50,5%	48,3%	49,2%
16	1:femme	46,7%	49,5%	51,7%	50,8%
17	Ensemble	100,0%	100,0%	100,0%	100,0%

Ce que vous devez pouvoir faire :

- Calculer (sans vous tromper) les % ligne, colonne ou du total dans un tableau d'effectif
- Sur papier (en général avec calculatrice)
- Sur tableur (au second semestre au plus tard)
- En utilisant une méthode « visuelle » (ou « tactile ») : par exemple, diviser chaque effectif d'une ligne par le total de la ligne, vérifier que la somme des % en ligne sur une ligne fait bien 100 %
- Un bon tableau de % indique dans son titre de quels % il s'agit et donne également les marges, qui permettent de vérifier sommairement si il s'agit d'un tableau de % en ligne ou en colonne
- En utilisant les formules sur un tableur (au second semestre au plus tard)
- En comprenant les formules mathématiques générales (à utiliser si elles vous aident, à comprendre lorsqu'elles sont utilisées dans d'autres contextes, par exemple en master)

Lecture des résultats

- Principe d'organisation : du général au particulier
- Principe de parcimonie : sélectionner et synthétiser l'information
- => omettre certaines informations déjà connues ou non essentielles
- => insister sur ce qui est important ou nouveau, en particulier la variable « d'intérêt »
- Traduire en français, simplifier les fractions et les arrondis (petits calculs mentaux parfois nécessaires)
- Transformer « ce que dit le tableau » en *résultats*

Interprétation

- « ce qu'on dit du tableau »
- Hypothèses ou éléments de connaissance permettant d'expliquer d'où viennent les résultats
- Travail par approximations successives : on explique d'abord les marges avant de détailler la façon dont chaque profil ligne s'écarte de la moyenne
- On présente les résultats évidents ou attendus avant de présenter des résultats plus fins, surprenants voire fragiles

Information simplifiée

Profil	Priorités entre carrière et vie familiale		
genre	1:Privilégie sa carrière professionnelle	2:Accorde une égale importance...	3:Privilégie sa famille
0:homme	+	+	-
1:femme	-	-	+

Suite

- Évaluer la force des écarts
- Quelle information est significative ?
- Choix d'autres variables