

Plaque 4 - Formulations, fichiers de données et filtres (ET, OU, NON)

Tableaux croisés — Seconde

Corrections détaillées

Boîte à outils

Critères, comptages et formulations.

- « A ET B » : les individus qui vérifient les deux critères (une case du tableau).
- « A OU B » : au moins un des deux critères. On compte sans compter deux fois :

$$n(A \text{ ou } B) = n(A) + n(B) - n(A \text{ et } B)$$
- « NON A » : $n(\text{non } A) = N - n(A)$. « Ni A ni B » = NON (A OU B).
- En Python : *and*, *or*, *not*, et les comparaisons == (égal) et != (différent).
- Traduire une phrase : « x % des élèves... » → fréquence marginale ou de case ;
 « parmi les A, x %... » ou « x % des A... » → fréquence conditionnelle f_A .
- Tableur : *NB.SI.ENS(plage1 ; critère1 ; plage2 ; critère2)* compte les lignes qui vérifient les deux critères.

Correction de l'exercice 1 - Traduire des phrases

La population de référence est ce qui suit « des » ou « parmi » ; c'est elle qui donne le dénominateur.

a) Référence : tous les élèves, catégorie « nageurs » : fréquence **marginale**,
 $\frac{75}{400} = 0,1875$. Exact.

b) Référence : les nageurs : fréquence **conditionnelle** $f_{\text{natation}}(\text{filles}) = \frac{45}{75} = 0,6$. Exact.

c) Référence : tous les élèves, individus qui sont à la fois filles et nageurs : fréquence d'une **case**,

$$\frac{45}{400} = 0,1125$$

Exact.

d) Référence : les filles : fréquence **conditionnelle** $f_{\text{filles}}(\text{natation}) = \frac{45}{210} \approx 0,214$.
 Exact.

Les phrases b) et d) ont le même numérateur, 45, mais pas la même référence.

Réponse. a) marginale · b) conditionnelle (parmi les nageurs) · c) case · d) conditionnelle (parmi les filles) ; toutes exactes.

Correction de l'exercice 2 - Filles ou nageurs

a) « Fille OU nageur » regroupe les élèves qui vérifient au moins un des deux critères. Les 45 filles nageuses vérifient les deux : il ne faut pas les compter deux fois.

$$n(A \text{ ou } B) = n(A) + n(B) - n(A \text{ et } B)$$

$$n = 210 + 75 - 45 = 240$$

b) En additionnant 210 et 75, l'élève compte deux fois les 45 filles nageuses (une fois parmi les filles, une fois parmi les nageurs).

c) $f = \frac{240}{400} = 0,6$, soit 60 % des élèves.

Contrôle. $240 = 210$ filles + 30 garçons nageurs.

Réponse. a) 240 élèves · c) 60 %.

Correction de l'exercice 3 - Ni football ni basket

a) « Ni football ni basket » est la négation de « football OU basket ». Ces deux sports n'ont pas d'élève commun (un seul sport par élève) :

$$n(\text{football ou basket}) = 100 + 75 = 175$$

$$n(\text{ni l'un ni l'autre}) = 400 - 175 = 225$$

Ce sont les nageurs et les élèves sans sport : $75 + 150 = 225$.

b) Référence : les 190 garçons. Les garçons qui ne jouent pas au football :

$$190 - 70 = 120, \text{ d'où}$$

$$f_{\text{garçons}}(\text{non football}) = \frac{120}{190} \approx 0,632$$

C'est $1 - f_{\text{garçons}}(\text{football}) = 1 - \frac{70}{190}$.

Réponse. a) 225 élèves · b) environ 63,2 %.

Correction de l'exercice 4 - Un filtre ET en Python

a) La fonction renvoie *True* quand les deux conditions sont vraies : l'élève est une fille (*sexe == "F"*) **et** son sport n'est pas « Aucun » (*!=* signifie « différent de »). Ce sont les **filles qui pratiquent un sport en club**.

b) On additionne les cases de la colonne « Filles » sauf « Aucun » :

$$n = 30 + 40 + 45 = 115$$

(ou $210 - 95 = 115$). Fréquence : $\frac{115}{400} = 0,2875$, soit 28,75 % des élèves.

Réponse. a) Les filles qui pratiquent un sport en club · b) 115 élèves, soit 28,75 %.

Correction de l'exercice 5 - Un filtre OU en Python

a) La fonction renvoie *True* dès qu'**au moins une** des deux conditions est vraie : l'élève joue au football **ou** c'est un garçon. Sont sélectionnés tous les garçons, et les filles footballeuses.

b) On compte sans doublon :

$$n(A \text{ ou } B) = n(A) + n(B) - n(A \text{ et } B)$$

$$n = 100 + 190 - 70 = 220$$

Contrôle. 190 garçons + 30 filles footballeuses = 220. Fréquence : $\frac{220}{400} = 55\%$.

Réponse. a) Tous les garçons et les filles qui jouent au football · b) 220 élèves.

Correction de l'exercice 6 - Compter dans un fichier

On parcourt le fichier ligne par ligne.

a) 2A et latin : élèves 1, 4, 9, soit 3 élèves.

b) Latin : élèves 1, 4, 5, 9 (4) ; demi-pension : 1, 2, 5, 7, 8, 9 (6) ; les deux : 1, 5, 9 (3).

$$n(\text{latin ou DP}) = 4 + 6 - 3 = 7$$

Ce sont les élèves 1, 2, 4, 5, 7, 8, 9.

c) 2B et pas demi-pensionnaire : élèves 6 et 10, soit 2 élèves.

Contrôle de b). Les élèves exclus sont ceux qui ne font ni latin ni demi-pension : 3, 6, 10, soit $10 - 7 = 3$.

Réponse. a) 3 · b) 7 · c) 2.

Correction de l'exercice 7 - Écrire les conditions

a) $classe == "2B"$ and $option == "Latin"$: seul l'élève 5, soit 1 élève.

b) $not dp == "Oui"$, ou plus simplement $dp != "Oui"$: élèves 3, 4, 6, 10, soit 4 élèves.

c) « Ni 2A ni latin » est la négation de « 2A ou latin » :

$$\text{non}(A \text{ ou } B) = (\text{non } A) \text{ et } (\text{non } B)$$

d'où $classe != "2A"$ and $option != "Latin"$, ou $not (classe == "2A" \text{ or } option == "Latin")$.
Élèves de 2B sans latin : 2, 6, 8, 10, soit 4 élèves.

Piège. $classe != "2A"$ or $option != "Latin"$ sélectionnerait presque tout le monde : seuls les élèves de 2A latinistes seraient exclus.

Réponse. a) 1 élève · b) 4 élèves · c) 4 élèves.

Correction de l'exercice 8 - Fréquences dans un fichier

On filtre d'abord la sous-population de référence, puis on compte dedans.

a) Demi-pensionnaires : élèves 1, 2, 5, 7, 8, 9 (6 élèves), dont latinistes 1, 5, 9 (3) :

$$f_{DP}(\text{latin}) = \frac{3}{6} = 0,5$$

b) Latinistes : 1, 4, 5, 9 (4 élèves), dont demi-pensionnaires 1, 5, 9 (3) :

$$f_{\text{latin}}(DP) = \frac{3}{4} = 0,75.$$

c) Tableau :

	Demi-pension	Externe	Total
Latin	3	1	4
Aucune	3	3	6
Total	6	4	10

Réponse. a) 50 % · b) 75 %.

Correction de l'exercice 9 - Vrai ou faux : le télétravail

Chaque phrase a sa population de référence :

$$f_A(B) = \frac{\text{effectif de « A et B »}}{\text{effectif de A}}$$

a) Référence : les cadres. $\frac{220}{400} = 0,55$: **vrai**.

b) Référence : les télétravailleurs réguliers. $\frac{220}{405} \approx 0,543$: **vrai**.

c) Référence : les employés. $\frac{60}{600} = 0,1$: **vrai**.

d) Référence : les télétravailleurs réguliers. $\frac{60}{405} \approx 0,148$, soit environ 14,8 % : **faux**. La phrase reprend le « 10 % » de c) en inversant la population de référence.

Réponse. a) vrai · b) vrai · c) vrai · d) faux (environ 14,8 %).

Correction de l'exercice 10 - Vrai ou faux : les accidents

a) Référence : les 450 accidents hors agglomération : $\frac{150}{450} = \frac{1}{3}$. **Vrai.**

b) On compare deux fréquences conditionnelles :

$$f_{\text{hors}}(\text{grave}) = \frac{150}{450} \approx 0,333 \quad f_{\text{agglo}}(\text{grave}) = \frac{130}{750} \approx 0,173$$

et $\frac{0,333}{0,173} \approx 1,92$: **vrai.**

c) 37,5 % est la part des accidents hors agglomération **parmi tous les accidents** ($\frac{450}{1200}$). Parmi les accidents graves, la part hors agglomération est $\frac{150}{280} \approx 53,6\%$. **Faux.**

Réponse. a) vrai · b) vrai · c) faux (environ 53,6 %).

Correction de l'exercice 11 - Ni fille ni nageur

Directement. « Ni fille ni nageur » signifie « garçon ET non nageur » : ce sont les garçons de la colonne « Garçons », sauf les nageurs :

$$190 - 30 = 160$$

Par la négation. « Ni fille ni nageur » est la négation de « fille OU nageur ». L'exercice 2 a donné $n(\text{fille ou nageur}) = 240$, donc

$$n(\text{ni fille ni nageur}) = 400 - 240 = 160$$

Les deux méthodes concordent : c'est la règle « non (A ou B) = (non A) et (non B) ».

Réponse. 160 élèves.

Correction de l'exercice 12 - Compter au tableur

a) **NB.SI.ENS** compte les lignes qui vérifient **tous** les critères : sexe "F" **et** sport "Natation". Elle renvoie l'effectif de la case « filles nageuses », 45.

b) On divise cet effectif par le nombre de filles, obtenu avec **NB.SI(B2:B401;"F")** :

$$= \text{NB.SI.ENS}(B2:B401;"F";C2:C401;"Natation")/\text{NB.SI}(B2:B401;"F")$$

qui calcule

$$f_{\text{filles}}(\text{natation}) = \frac{45}{210} \approx 0,214$$

Intérêt. Si le fichier est mis à jour, les fréquences se recalculent seules.

Réponse. a) 45 · b) le quotient des deux comptages, environ 0,214.

Correction de l'exercice 13 - Algorithme de fréquence conditionnelle

a) La boucle parcourt tous les élèves. Le premier **if** compte les filles ; le second, **placé à l'intérieur**, ne compte les nageuses que parmi les filles. La fonction renvoie

$$f_{\text{filles}}(\text{natation}) = \frac{\text{nombre de filles nageuses}}{\text{nombre de filles}}$$

b) Avec le fichier de l'exercice 1 : $\frac{45}{210} \approx 0,214$.

Variante. Pour obtenir $f_{\text{natation}}(\text{filles})$, il faudrait compter d'abord les nageurs, puis les filles parmi eux : on échange les rôles des deux tests.

Réponse. a) $f_{\text{filles}}(\text{natation})$ · b) environ 0,214.

Correction de l'exercice 14 - En ligne ou en magasin

a) Sans compter deux fois les 40 clients qui font les deux :

$$n(\text{ligne ou magasin}) = 120 + 90 - 40 = 170$$

et $200 - 170 = 30$ clients n'ont acheté ni en ligne ni en magasin.

b) En ligne seulement : $120 - 40 = 80$; magasin seulement : $90 - 40 = 50$.

	Magasin	Pas de magasin	Total
En ligne	40	80	120
Pas en ligne	50	30	80
Total	90	110	200

c) $f_{\text{ligne}}(\text{magasin}) = \frac{40}{120} \approx 33,3\%$ et $f_{\text{magasin}}(\text{ligne}) = \frac{40}{90} \approx 44,4\%$.

Réponse. a) 170 ; 30 · c) environ 33,3 % ; environ 44,4 %.

Correction de l'exercice 15 - Les lycéens qui lisent

a) Marges : $0,52 \times 800 = 416$ filles, 384 garçons ; $0,35 \times 800 = 280$ lecteurs, dont $0,7 \times 280 = 196$ filles et 84 garçons.

	Filles	Garçons	Total
Lecteurs	196	84	280
Non-lecteurs	220	300	520
Total	416	384	800

b) Dans chaque colonne :

$$f_{\text{filles}}(\text{lecteur}) = \frac{196}{416} \approx 0,471 \quad f_{\text{garçons}}(\text{lecteur}) = \frac{84}{384} \approx 0,219$$

c) 70 % est la part des filles **parmi les lecteurs**. Parmi les filles, la proportion de lectrices est d'environ 47 % : moins de la moitié.

Réponse. b) environ 47,1 % des filles et 21,9 % des garçons · c) environ 47 % des filles lisent un livre par mois.