

OBJECTIFS

- Lire et construire un tableau croisé d'effectifs.
- Calculer une fréquence marginale et une fréquence conditionnelle.
- Préciser clairement la population de référence.
- Filtrer des individus avec des critères de type ET, OU, NON.

ACTIVITE DE DEPART

Dans une enquête, on interroge 120 élèves sur deux caractères : *le moyen de transport* et *le fait de pratiquer ou non un sport en club*.

1. Que signifie : « parmi les élèves venant en bus » ?
2. Pourquoi le pourcentage de sportifs peut-il changer si l'on ne regarde que les élèves venant à vélo ?
3. Quelle table permettrait de résumer toutes les informations ?

1. Tableaux croisés d'effectifs

Définition (Deux variables qualitatives) Une **variable qualitative** décrit une catégorie : couleur, choix, réponse, type de transport, groupe, etc. Croiser deux variables qualitatives consiste à classer les individus selon deux critères simultanément.

Définition (Tableau croisé d'effectifs) Un **tableau croisé d'effectifs** présente les effectifs d'une population selon deux variables qualitatives. Les totaux de lignes et de colonnes sont appelés **marges**.

Méthode Pour construire un tableau croisé, on choisit une variable en lignes et une variable en colonnes, puis on compte les individus dans chaque case. On ajoute ensuite les totaux de lignes et de colonnes pour contrôler le tableau.

Vigilance. Une case du tableau correspond à deux conditions simultanées. Par exemple, « vélo et sport » ne signifie pas « vélo ou sport ».

	Sport	Pas sport	Total
Bus	18	22	40
Vélo	20	10	30
Autre	27	23	50
Total	65	55	120

Exemple Il y a 20 élèves qui viennent à vélo et pratiquent un sport en club. Il y a 30 élèves qui viennent à vélo au total.

⇒ TD N17 – Partie A : lire un tableau croisé

2. Fréquences marginales et conditionnelles

Définition (Fréquence marginale) Dans une population de taille N , la fréquence d'un sous-ensemble d'effectif n est $\frac{n}{N}$. Dans un tableau croisé, une fréquence calculée à partir d'un total de ligne ou de colonne est une **fréquence marginale**.

Définition (Fréquence conditionnelle) Une **fréquence conditionnelle** est une fréquence calculée dans un sous-groupe. Elle répond à la question : « parmi quels individus ? »

Propriété (Formule de fréquence conditionnelle) Si A et B sont deux caractères et si l'effectif de B est non nul, alors la fréquence de A **parmi** les individus de B est

$$f_B(A) = \frac{\text{effectif de } A \cap B}{\text{effectif de } B}.$$

Méthode Avant tout calcul de pourcentage, écrire une phrase du type : « On se place parmi les ... ». Le dénominateur est alors l'effectif de ce groupe.

Exemple La fréquence marginale des élèves sportifs est $\frac{65}{120} \approx 54,2\%$. La fréquence conditionnelle des sportifs parmi les élèves venant en bus est $\frac{18}{40} = 45\%$.

⇒ TD N17 – Partie B : calculer des fréquences conditionnelles

3. Filtrer des données et interpréter

Définition (Filtre) Filtrer une base de données consiste à ne garder que les individus vérifiant un critère. On utilise souvent les mots **ET**, **OU**, **NON**.

Propriété (Sens des connecteurs) Le critère A ET B garde les individus vérifiant les deux conditions. Le critère A OU B garde les individus vérifiant au moins l'une des deux conditions. Le critère NON A garde les individus qui ne vérifient pas A .

Méthode Pour interpréter une phrase médiatique du type « 70 % des personnes de tel groupe vérifient tel critère », repérer :

1. la population totale ;
2. le sous-groupe éventuel ;
3. le caractère dont on mesure la fréquence ;
4. le dénominateur utilisé.

Vigilance. Le mot « parmi » est décisif. Dire « 60 % des sportifs viennent à vélo » et « 60 % des cyclistes sont sportifs » ne signifie pas la même chose.

4. Synthèse

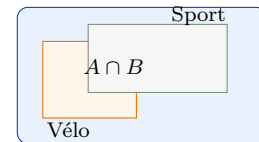
Exemple Dans le tableau précédent :

$$f_{\text{vélo}}(\text{sport}) = \frac{20}{30} = \frac{2}{3} \approx 66,7\%.$$

En revanche,

$$f_{\text{sport}}(\text{vélo}) = \frac{20}{65} \approx 30,8\%.$$

Les deux résultats ne répondent pas à la même question.



Exemple Dans une liste d'individus, le filtre :

Transport = vélo ET Sport = oui

sélectionne les élèves situés dans la case 20 du tableau. Le filtre :

Transport = vélo OU Sport = oui

sélectionne plus d'élèves : tous les cyclistes, tous les sportifs, sans compter deux fois les cyclistes sportifs.

```
selection = []
for i in range(len(transport)):
    if transport[i]=="velo" and sport[i]=="oui":
        selection.append(i)
```

A RETENIR

- Dans un tableau croisé, chaque case correspond à l'intersection de deux critères.
- Une fréquence marginale utilise le total de la population comme référence.
- Une fréquence conditionnelle utilise un sous-groupe comme référence : il faut toujours préciser « parmi quoi ».
- Les filtres ET, OU, NON permettent de traduire des critères sur une liste d'individus.

APPLICATIONS IMMEDIATES

1. Dans le tableau initial, calculer la fréquence des élèves venant en bus parmi tous les élèves.
2. Calculer la fréquence des élèves venant en bus parmi les sportifs.
3. Calculer la fréquence des sportifs parmi les élèves venant par un autre moyen.
4. Expliquer en une phrase la différence entre les deux questions précédentes.