

Activité introductive :

Voici les choix des élèves d'une classe 10^{ième} commune pour la 11^{ième} : AL ; S.Eco ; SE ; AL ; SE ; S.Exp ; SS ; SE ; SS ; AL ; S.Eco ; SE ; AL ; SE ; S.Exp ; SS ; SE ; SS ; AL ; S.Eco ; SE ; AL ; SE ; S.Exp ; SS ; SE ; SS ; AL ; S.Eco ; SE ; AL ; SE ; S.Exp ; SS ; SE ; SS ;

Complete le tableau suivant :

	Effectif de la classe	AL	SE	S.Eco	S.Exp	SS
Effectif						
Pourcentage(%))						

I. Introduction au vocabulaire et aux symbolismes graphiques représentant d'une série :

Une étude statistique comporte généralement quatre étapes

- **Le recueil des données** : Il se fait à l'aide d'enquêtes, de recensements, etc. Nous supposons ce travail effectué.
- **La présentation des résultats** : L'enquête dépouillée, on donne une première présentation des résultats sous forme de tableaux. Quelquefois on préfère une représentation graphique plus « parlante ».
- **Calcul des paramètres caractéristiques** : Ils résument une série statistique en quelques nombres afin d'en faciliter l'interprétation.
- **L'exploitation des données** : Il s'agit, au vu des résultats des calculs des différents paramètres, de réviser une stratégie commerciale, de modifier un traitement médical, de prévoir un budget, etc. Selon la nature de la série étudiée.

Exemple :

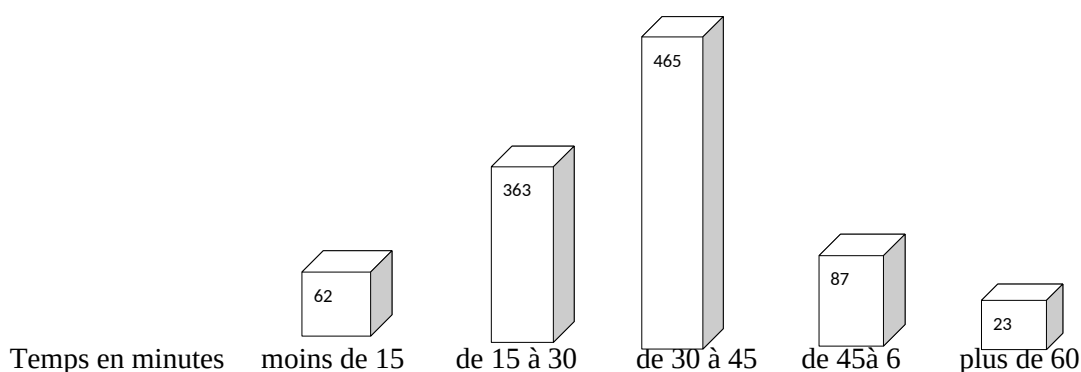
a. Questionnaire :

On a interrogé 1000 clients sortant d'un magasin d'alimentation. Voici les quatre questions posées et les réponses obtenues :

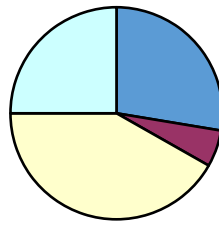
1. Venez-vous régulièrement dans ce magasin ?

oui	non
912	88

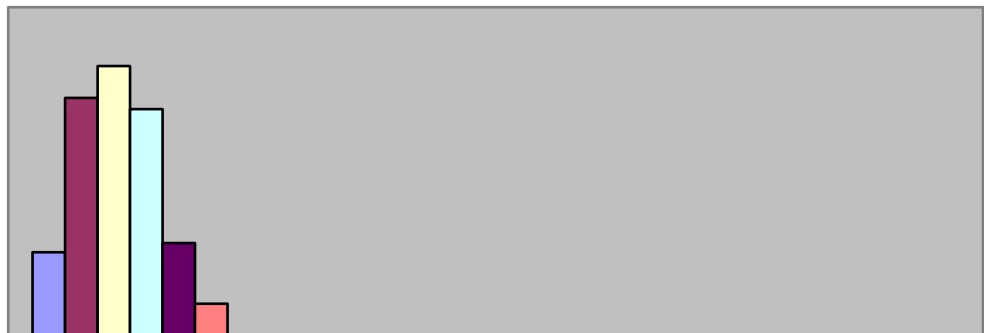
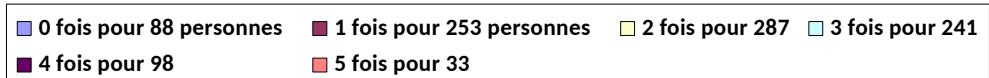
2. Combien de temps venez-vous de passer dans ce magasin ?



3. Êtes-vous satisfait du choix qui vous est offert ?



4. Combien de fois par semaine en moyenne venez-vous habituellement dans ce magasin ?



NB : la réponse « 0 » signifie que la personne interrogée vient moins d'une fois.

b. **Population :**

Dans l'exemple ci-dessus, la **population** a un effectif de 1000. Les **individus** (ou éléments de la population) sont les clients interrogés.

c. **Caractères :**

Nous pouvons étudier quatre **caractères** que nous appelons :

- Client régulier (question 1⁰) ;
- Temps passé (question 2⁰) ;
- Satisfaction (question 3⁰) ;
- Fréquentation (question 4⁰).

Les valeurs prises par les caractères sont les **modalités**. On distingue :

- Les **caractères quantitatifs** lorsque les modalités sont numériques ;
- Les **caractères qualitatifs** dans le cas contraire.

Exemples de **caractères quantitatifs** :

- o Temps passé : c'est un **caractère continu (à modalités regroupées en classes)**. Les **modalités** peuvent prendre toutes les valeurs d'un intervalle. On les regroupe en classes (ici de 0 à 15, de 15 à 30 etc.) ;
- o Fréquentation : c'est un **caractère discret**. Les modalités ne peuvent prendre qu'un nombre fini de valeurs (0, 1, 2, 3, 4, ou 5).

Exemples de **caractères qualitatifs** :

- ❖ client régulier : deux modalités (« oui » et « non »).
- ❖ Satisfaction : les modalités sont prévues dans la question afin de faciliter le dépouillement.

d. **Effectifs et fréquences** :

On appelle :

- **Effectif** d'une modalité : le nombre d'individus ayant cette modalité ;
- **Fréquence** d'une modalité : le quotient de l'effectif de la modalité par l'effectif total de la population concernée, ce qui donne un nombre entre 0 et 1. Les fréquences sont généralement exprimées en pourcentages.

Question : satisfaction des clients

L'effectif de la population est 1000

Modalité	Très satisfait	satisfait	Moyennement satisfait	Pas satisfait
Effectif	418	250	276	56
Fréquence	41.8 %	25 %	27.6 %	5.6 %

Lorsque les modalités sont ordonnées, les effectifs (fréquences) peuvent être cumulés.

Effectifs cumulés, fréquences cumulées :

Exemple : **Temps passé**.

Nombre de fois : x_i	0	1	2	3	4	5
Effectifs : n_i	88	253	287	241	98	33
Valeurs	$x_i \leq 0$	$x_i \leq 1$	$x_i \leq 2$	$x_i \leq 3$	$x_i \leq 4$	$x_i \leq 5$
Effectifs cumulés croissants	88	341	628	869	967	1000
Fréquences cumulées croissantes	8.8 %	34.1 %	62.8 %	86.9 %	96.7 %	100 %
Valeurs	$x_i \geq 0$	$x_i \geq 1$	$x_i \geq 2$	$x_i \geq 3$	$x_i \geq 4$	$x_i \geq 5$
Effectifs cumulés décroissants	1000	912	659	372	131	33
Fréquences cumulées décroissantes	100 %	91.2 %	65.9 %	37.2 %	13.1 %	3.3 %

Par exemple : Il y a 628 personnes qui viennent au plus deux fois par semaine.

Il y a 659 personnes qui viennent au moins deux fois par semaine dans le magasin.

Définitions :

Soit $(x_i; n_i)_{1 \leq i \leq p}$ une série statistique dont les modalités sont rangées par ordre croissant.

- On appelle effectif cumulé croissant (resp. décroissant) de la modalité x_k la somme des effectifs des modalités inférieures ou égales (resp. décroissant) à x_k .
- On appelle fréquence cumulée croissante (resp. décroissante) de la modalité x_k le quotient de son effectif cumulé croissant (resp. décroissant) par l'effectif total de la série.

Donc, en notant N_k (resp. N'_k) l'effectif cumulé croissant (resp. décroissant), F_k (resp. F'_k) la fréquence cumulée croissante (resp. décroissante) de la modalité x_k et N l'effectif total, on a :

$$N_k = \sum_{i=1}^k n_i ; \quad N'_k = \sum_{i=k}^p n_i ; \quad F_k = \frac{N_k}{N} ; \quad F'_k = \frac{N'_k}{N}$$

Evaluations :

- On a relevé les salaires mensuels, exprimés en milliers de francs CFA, des ouvriers d'un atelier de mécanique :
65 ; 83 ; 53 ; 115 ; 94 ; 51 ; 75 ; 71 ; 102 ; 76 ; 105 ; 81 ; 83 ; 77 ; 58 ; 55 ;
90 ; 64 ; 100 ; 108 ; 68 ; 53 ; 51 ; 75 ; 83 ; 80 ; 95 ; 78 ; 91 ; 79 ; 74 ; 110 ;
77 ; 90 ; 92 ; 84 ; 88 ; 85 ; 64 ; 71

On regroupe ces données en cinq classes :

$[50 ; 70[$, $[70 ; 80[$, $[80 ; 90[$, $[90 ; 100[$ et $[100 ; 120[$.

Etablir le tableau des effectifs et des fréquences

- Dans une classe 11^{ème} année, les notes obtenues par les élèves en Maths se répartissent ainsi :

Note	2	5	7	9	10	11	13	14	16	18
Effectif	1	2	3	4	4	9	2	3	2	1

- Quel est l'effectif total de cette série ?
- Calculer les fréquences, les fréquences cumulées croissantes et décroissantes
- Quel est le pourcentage des élèves qui ont obtenu une note supérieure ou égale à 10 ?

II. Caractéristiques de position :

A. Mode :

Définition :

On appelle mode d'une série statistique toute modalité d'effectif maximal.

B. Moyenne :

La notion moyenne n'a de sens que pour des caractères quantitatifs.

0 Caractère discret :

Par exemple, pour le caractère « fréquentation » (question 4⁰), la moyenne est

$$\bar{x} = \frac{88 \times 0 + 253 \times 1 + 287 \times 2 + 241 \times 3 + 98 \times 4 + 33 \times 5}{88 + 253 + 287 + 241 + 98 + 33}, \text{ soit } \bar{x} = 2.1$$

Définition :

Pour une série statistique $(x_i; n_i)_{1 \leq i \leq p}$, la moyenne est le nombre, noté $\bar{x}$, défini par :

$$\bar{x} = \frac{\sum_{i=1}^p n_i x_i}{\sum_{i=1}^p n_i} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_p x_p}{n_1 + n_2 + \dots + n_p}$$

0 Caractère continu :

On procède comme un caractère discret, en prenant pour $x_1, x_2, \dots, x_p$ le centre de chaque classe.

Exemple : Temps passé

Si l'on convient que la dernière classe a une largeur de 60, on obtient :

Modalité	0 ; 15	15 ; 30	30 ; 45	45 ; 60	60 ; 120
Centre	7.5	22.5	37.5	52.5	90
Effectif	62	363	465	87	23

On en déduit la moyenne :

$$\bar{x} = \frac{1}{1000} (62 \times 7.5 + 363 \times 22.5 + 465 \times 37.5 + 87 \times 52.5 + 23 \times 90)$$

Soit environ 33.

Ce qui signifie qu'un client passe en moyenne 33 minutes dans le magasin.

C. Médiane :

Définition :

On appelle médiane d'une série statistique groupée, tout nombre M_e tel que :

- La moitié au moins des individus ont une modalité inférieure ou égale à M_e .
- La moitié au moins des individus ont une modalité supérieure ou égale à M_e .

C'est-à-dire M_e est donc tel la fréquence cumulée croissante est supérieure ou égale à 50 % et sa fréquence cumulée décroissante est aussi supérieure ou égale à 50 %.

Exemple : Temps passé.

Nombre de fois : x_i	0	1	2	3	4	5
Effectifs : n_i	88	253	287	241	98	33
Valeurs	$x_i \leq 0$	$x_i \leq 1$	$x_i \leq 2$	$x_i \leq 3$	$x_i \leq 4$	$x_i \leq 5$
Effectifs cumulés croissants	88	341	628	869	967	1000
Valeurs	$x_i \geq 0$	$x_i \geq 1$	$x_i \geq 2$	$x_i \geq 3$	$x_i \geq 4$	$x_i \geq 5$
Effectifs cumulés décroissants	1000	912	659	372	131	33

Effectif total N étant 1000 alors sa moitié 500. Donc la médiane est la modalité 2.

III. Caractéristiques de dispersion :

a) Ecart moyen :

Définition :

Soit une série statistique $(x_i; n_i)_{1 \leq i \leq p}$ de moyenne $\bar{x}$. L'écart moyen est le nombre réel e_m défini par :

$$e_m = \frac{\sum_{i=1}^p n_i |x_i - \bar{x}|}{\sum_{i=1}^p n_i} = \frac{n_1 |x_1 - \bar{x}| + n_2 |x_2 - \bar{x}| + \dots + n_p |x_p - \bar{x}|}{n_1 + n_2 + \dots + n_p}$$

Exemple :

Les notes en Mathématiques de deux élèves A et B sont consignées dans les tableaux suivants :

Elève A

Note	8	9	12	13
Effectif	2	3	2	1

Elève B

Note	8	9	10	11	12
Effectif	1	1	4	1	1

Ces deux élèves ont la **moyenne 10**. Mais calculons les écarts moyens :

Pour l'élève A : $e_A = \frac{2|8-10| + 3|9-10| + 2|12-10| + |13-10|}{8} = 1.75$

Pour l'élève B : $e_B = \frac{|8-10| + |9-10| + 4|12-10| + |11-10| + |12-10|}{8} = 0.625$

Ainsi $e_B < e_A$. Les notes de l'élève B sont moins dispersées que celles de l'élève A : l'élève B est plus régulier que l'élève A.

b) Variance et Ecart type :

Définition :

Soit une série statistique $(x_i; n_i)_{1 \leq i \leq p}$ de moyenne $\bar{x}$.

- La variance est le nombre réel V défini par :

$$V = \frac{\sum_{i=1}^p n_i (x_i - \bar{x})^2}{\sum_{i=1}^p n_i} = \frac{n_1 (x_1 - \bar{x})^2 + n_2 (x_2 - \bar{x})^2 + \dots + n_p (x_p - \bar{x})^2}{n_1 + n_2 + \dots + n_p}$$

- L'écart type est le nombre réel positif σ défini par : $\sigma = \sqrt{V}$

Evaluation corrigée :

Les notes en Mathématiques de NGolo ont été les suivantes : 12-8-10-14-9-12-7-13.

- Calcule la moyenne des notes :
- Calcule la variance et l'écart type de la série.
- Le professeur décide de supprimer la note plus haute et la note la basse.
 - Calcule la nouvelle moyenne ;

- b) Calcule les nouvelles valeurs de la variance et écart type, et les compare aux valeurs obtenues à la deuxième question. Interprète la variation constatée.
- 4) On reprend la série statistique initiale. Quelle note minimale doit obtenir NGolo au prochain devoir pour que sa moyenne dépasse 11.

Solution :

- 1) $\bar{x}=10.635$
- 2) $V=5.48$ et $\sigma=2.34$
- 3) La nouvelle série est : 12-8-10-12-13
 - a) $\bar{x}=10.67$
 - b) $V'=3.24$ et $\sigma'=1.80$. ces valeurs sont inférieures à celles trouvées à la deuxième question puisque la nouvelle série de notes est moins dispersée que l'ancienne.
- 4) Soit n la note :

$$\frac{1}{9}(12+8+10+14+9+12+7+13+n) \geq 11$$

C'est-à-dire $85+n \geq 9 \times 11$ alors $n \geq 14$

NGolo doit obtenir au moins 14 au prochain devoir pour que sa moyenne dépasse 11.