

CH IV Statistique

I) Rappels de seconde :

1) Le caractère d'une série statistique :

J Le caractère d'une série statistique peut être ou
Lorsqu'il est, il sera ou

Caractère Caractère

```
graph TD; A[Caractère] --> B[.....]; A --> C[.....];
```

Ex1 : L'étude porte sur le nombre de voitures rouges, vertes, bleues et noires qui se trouvent sur un parking. Le caractère étudié sera ici

Ex2 : L'étude porte sur le nombre de frères et sœurs qu'un élève de lycée peut avoir. Un élève pourra avoir 0 ; 1 ; 2 ; 3 ; 4 ... frères et sœurs. Le caractère étudié sera

Ex3 : L'étude porte sur l'âge des élèves qui passent le bac. On pourra classer le nombre d'élèves dans des intervalles : [16 ; 17[; [17 ; 18[; [18 ; 19[... Le caractère étudié ici sera

2) La fréquence :

J La fréquence f_i d'une valeur ou d'une classe de valeurs d'un caractère est le rapport $\frac{n_i}{N}$, où n_i est l'effectif et N l'effectif total.

$$f_i = \frac{n_i}{N}$$

3) Les représentations graphiques :

Exemple 1 : La répartition selon la surface en m^2 , des 67 magasins d'un centre commercial est représenté par le diagramme en bâtons de la page suivante :

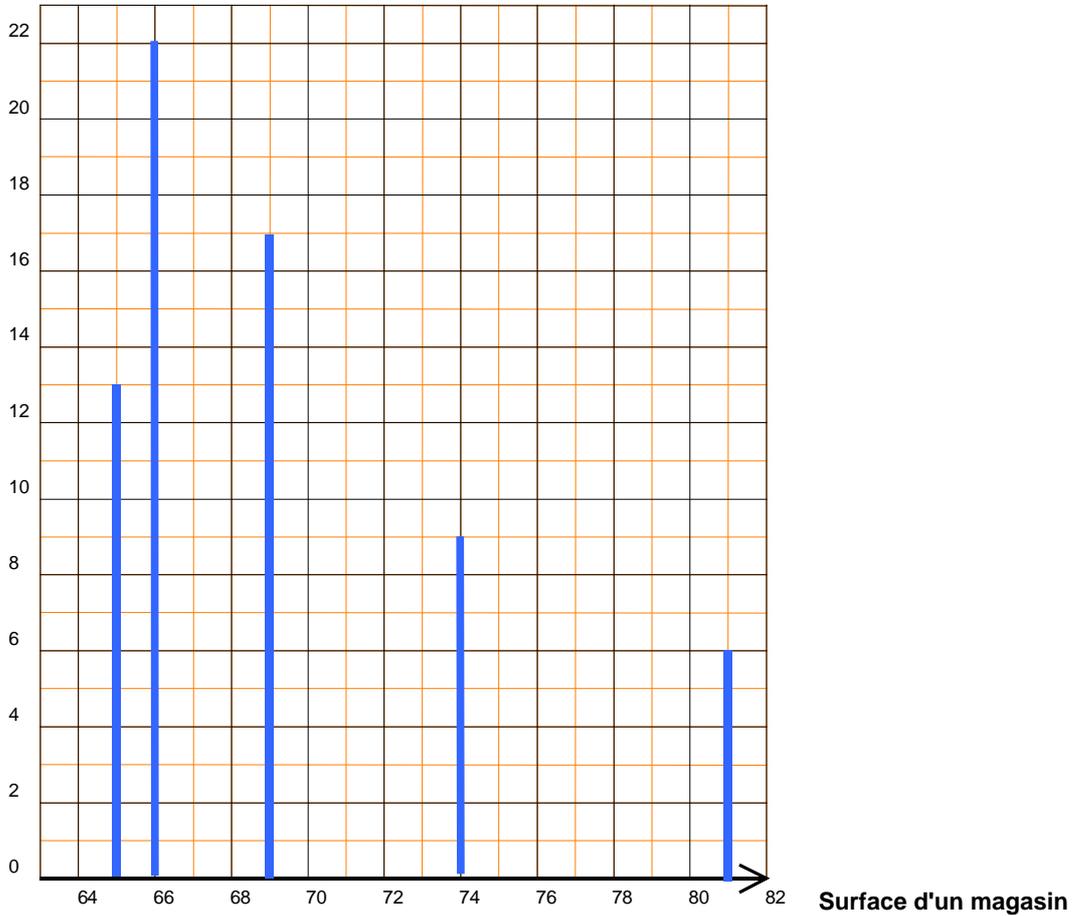
a) Quel est la nature de ce caractère ?

.....

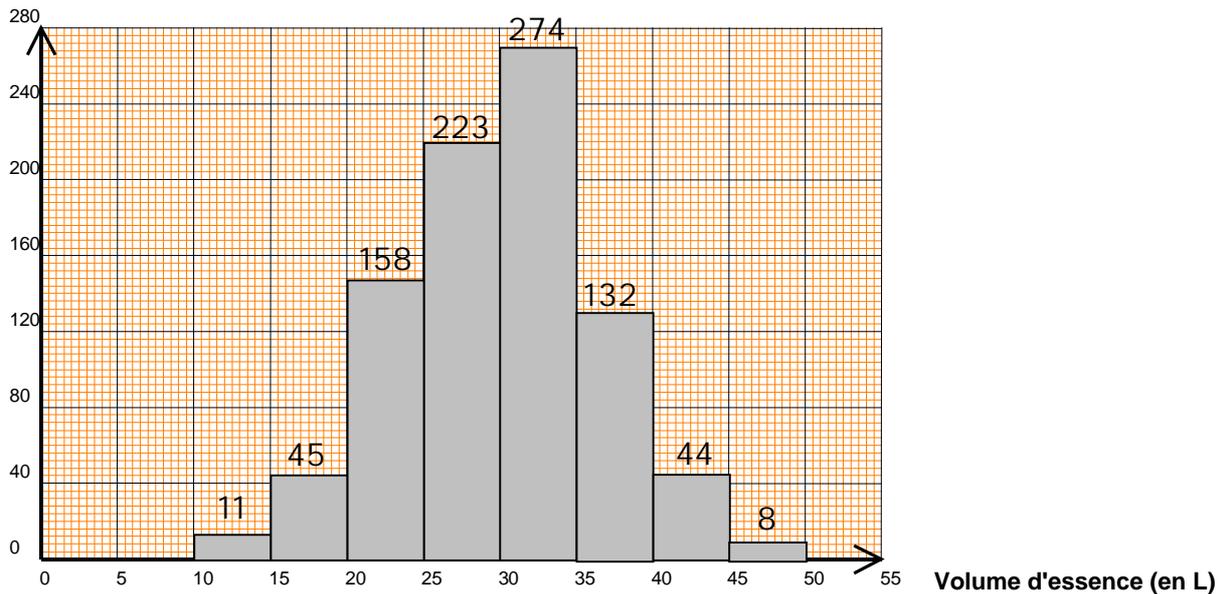
b) A l'aide du graphique, compléter le tableau :

Surface d'un magasin (en m^2) x_i						Total
Nombre de magasins n_i						
Fréquence f_i (arrondie à 0,001)						

Nombre de magasins



Exemple 2 : On relève dans une station service, une semaine donnée, le volume d'essence, acheté par 895 clients. Les résultats sont représentés par l'histogramme suivant :



a) Quelle est la nature de ce caractère ?

b) A l'aide du graphique, compléter le tableau suivant :

Volume d'essence (en L)	Nombre de clients	Fréquence f_i (en %, arrondie au centième)
Total		

II) Les indicateurs de tendance centrale :

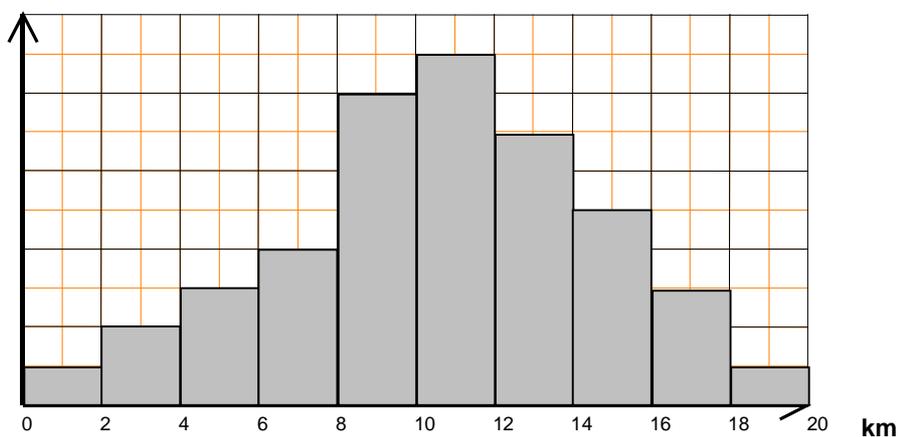
1) Le mode, la classe modale :

J Le mode est la valeur du caractère ayant Une distribution avec un seul mode est dite, avec deux modes, elle est dite Dans le cas de répartition en classes d'égalles amplitudes, la classe modale désigne celle qui a Le mode est de cette classe.

Exemples :

Deux séries de distances entre le domicile et le lycée (en km) rangées en classes.

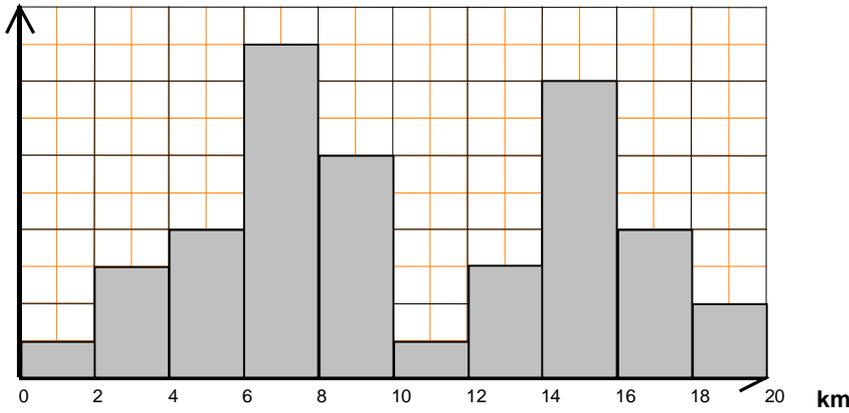
effectif



La classe modale est :

Le mode $M_0 =$

effectif



Les classes modales sont :

Le mode M_{01} =

Le mode M_{02} =

2) La moyenne :

On a relevé les âges des participants à une manifestation sportive :

Âges x_i	14	17	19	20	25	38	43	50
Effectifs n_i	2	3	1	5	3	1	1	1

J La moyenne \bar{x} de cette série est obtenue en divisant la somme des valeurs $n_i \cdot x_i$ par l'effectif total N.

$$\bar{x} = \frac{\sum x_i \cdot n_i}{N} = \frac{\quad}{\quad} = \quad .$$

Lorsque les valeurs sont regroupées en classes, la somme des valeurs est obtenue en multipliant chaque valeur centrale par l'effectif de la classe.

On peut obtenir la moyenne directement à l'aide de la calculatrice : Voir Fiche calcul des paramètres statistiques .

On peut également utiliser un tableur pour obtenir cette moyenne.

J L'interprétation de la moyenne est la suivante :

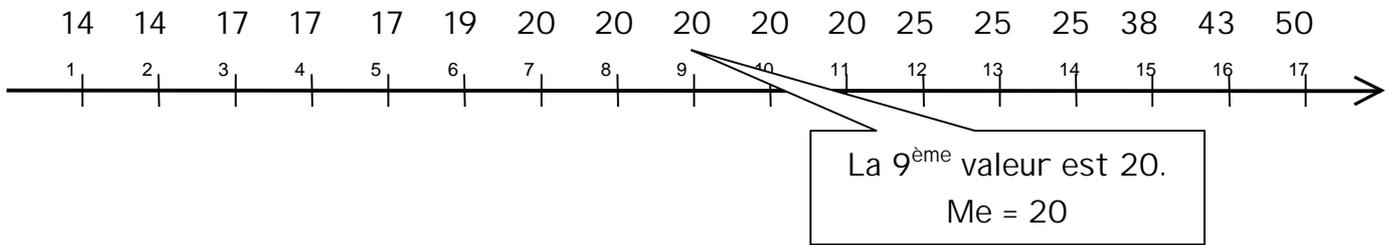
3) La médiane d'une série statistique :

J La médiane Me de la série est un nombre qui découpe la liste des âges, rangée dans l'ordre croissant, en deux listes. Si l'effectif total est impair, on obtient la médiane en prenant la valeur correspondant à l'effectif + 1 divisé par 2.

On reprend le relevé les âges des participants à une manifestation sportive précédent

La médiane correspond ici à la $\frac{17+1}{2} = 9^{\text{ème}}$ valeur. $Me = 20$.

On place les âges des 17 participants à la manifestation sportive dans l'ordre croissant :



Lorsque l'effectif est pair, la médiane sera la moyenne des deux valeurs correspondantes aux rangs $\frac{\text{Effectif}}{2}$ et $\frac{\text{Effectif}}{2} + 1$

On peut obtenir la médiane directement à l'aide de la calculatrice : Voir Fiche calcul des paramètres statistiques .

On peut également utiliser un tableur pour obtenir cette médiane.

J L'interprétation de la médiane est la suivante : « »

Exercice N°1 : Pour chaque série statistique :

- 1) Déterminer le ou les modes :
- 2) Déterminer, avec une calculatrice, la moyenne et la médiane.

Valeurs x_i	2	8	10	14	21	25	30
Effectif n_i	3	5	7	10	8	4	3

Valeurs x_i	10	15	30	40	45	50	55	65
Effectif n_i	1	1	2	4	3	2	1	1

Exercice N°2 : Dans une classe de première, on demande aux élèves le nombre d'activités sportives qu'ils pratiquent régulièrement. On obtient les réponses suivantes :

2 élèves ne font aucun sport ; 10 élèves pratiquent un sport ; 8 élèves pratiquent deux sports ; 6 élèves pratiquent trois sports ; 4 élèves pratiquent quatre sports.

a) Déterminer la population et le caractère étudié. Quelle est la nature de ce caractère ?

La population correspond aux

Le caractère étudié correspond au

La nature de ce caractère est

b) Regrouper ces résultats dans un tableau.

c) Quel est le nombre d'activités sportives le plus fréquent dans cette classe.

Quel est l'indicateur de tendance central associé ?

d) Déterminer avec une calculatrice la moyenne de cette série statistique. Interpréter le résultat par une phrase.

e) Déterminer avec une calculatrice la médiane de cette série statistique. Interpréter le résultat par une phrase.

III) Les indicateurs de dispersion : Étendue, écart type, quartiles, écart interquartile :

1) L'étendue :

J L'étendue est

2) Évaluer la dispersion d'une série avec l'écart type :

Voici le relevé de 3 notes de mathématiques de Théa, Hector et Basile.

Calculer à l'aide de la calculatrice pour chaque élève la moyenne \bar{x} et l'écart type σ .

Théa	11	9	10	$\bar{x} =$	$\sigma =$
Hector	7	15	8	$\bar{x} =$	$\sigma =$
Basile	11	3	16	$\bar{x} =$	$\sigma =$

J L'écart type indique la dispersion des valeurs de la série autour de leur moyenne.
Sur la calculatrice Casio, l'écart type correspond à la valeur $x\sigma n$ ou σx .
Sur la calculatrice TI, l'écart type correspond à la valeur σx .

Dans ce cas les moyennes sont les mêmes, les écarts types indiquent une plus grande régularité des notes pour Théa que pour Basile.

J Le couple (moyenne ; écart type) rend compte d'un certain point de vue de la série.

3) Évaluer la dispersion d'une série avec l'écart interquartile :

J Le premier quartile Q_1 partage les valeurs du caractère, rangées en ordre croissant, en deux groupes : le 1^{er} groupe représente à peu près de l'effectif total et le 2^{ème} groupe à peu près

J Le troisième quartile Q_3 partage les valeurs du caractère, rangées en ordre croissant, en deux groupes : le 1^{er} groupe représente à peu près de l'effectif total et le 2^{ème} groupe à peu près

J L'écart interquartile $Q_3 - Q_1$ représente de l'effectif total.

J Le couple (médiane ; écart interquartile) rend compte d'un certain point de vue de la série.

La répartition des années d'ancienneté de 24 employés d'une entreprise est donnée par le tableau suivant :

Années d'ancienneté x_i	1	2	3	4	5
Nombre d'employés n_i	5	5	7	4	3

1) Déterminer l'étendue de cette série. Interpréter ce résultat.

2) Déterminer, à l'aide de la calculatrice, la moyenne et l'écart type arrondis à 0,1.

3) Déterminer, à l'aide de la calculatrice, le premier et le troisième quartile. Interpréter ces résultats.

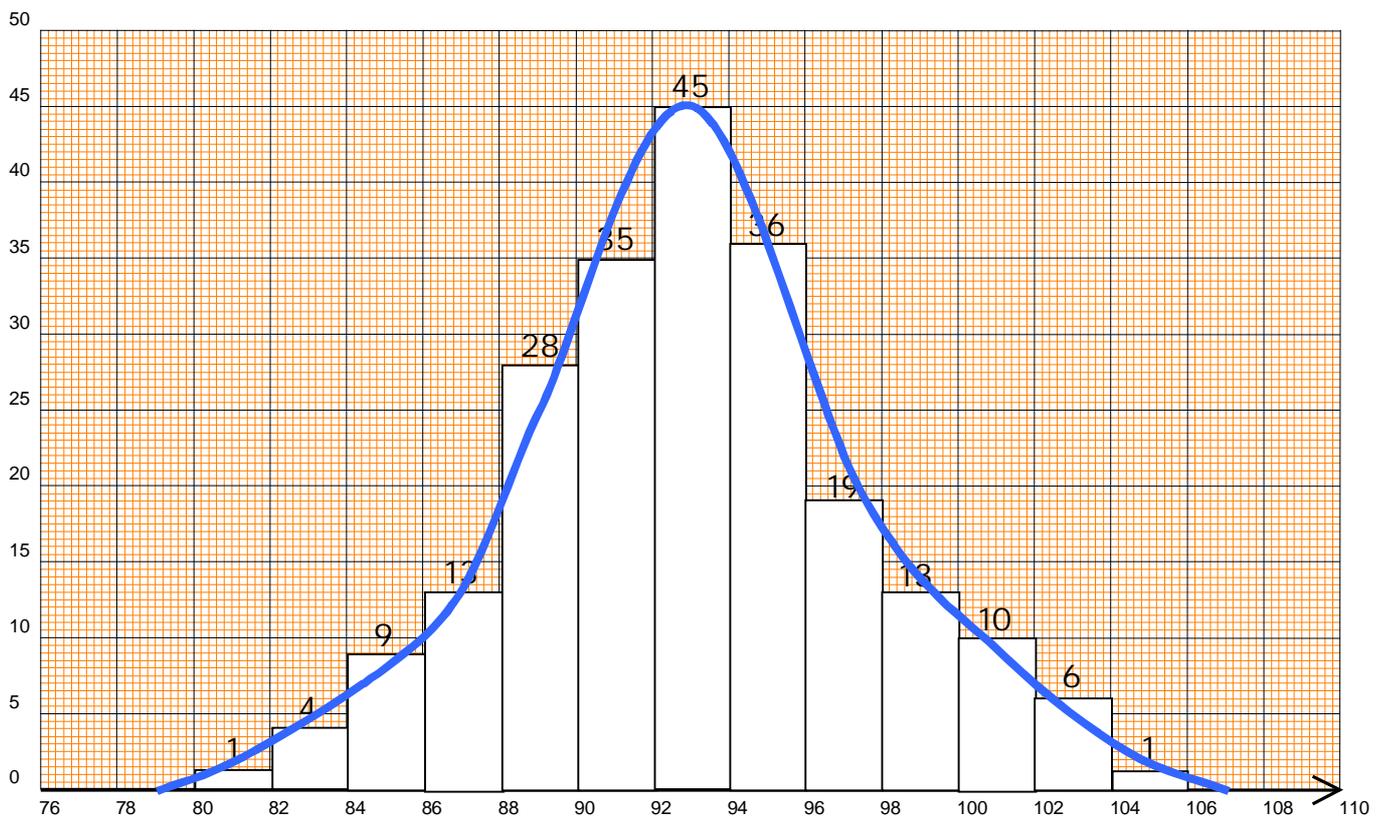
4) Calculer l'écart interquartile $Q_3 - Q_1$. Interpréter ce résultat.

IV) La courbe de Gauss et la boîte à moustaches :

1) La courbe de Gauss :

Un laboratoire d'analyse biologiques étudie les taux de calcium (en mg/L) de 220 personnes. Les résultats sont représentés par un histogramme, sur lequel la courbe tracée visualise l'allure de la série. Cette courbe, en forme de cloche, est appelée courbe de Gauss.

J Pour une série statistique dont la représentation graphique a l'allure de la courbe de Gauss, les valeurs sont réparties à peu près symétriquement autour de leur moyenne. Environ 95 % d'entre elles appartiennent à l'intervalle $[\bar{x} - 2\sigma ; \bar{x} + 2\sigma]$.



Calculer, à l'aide de la calculatrice, les moyenne et écart type de la série précédente.

$\bar{x} = \dots\dots\dots$ $\sigma = \dots\dots\dots$

Calculer $\bar{x} - 2\sigma = \dots\dots\dots$ $\bar{x} + 2\sigma = \dots\dots\dots$

Calculer le nombre de personnes dont le taux, en mg/L, est compris entre 84 et 102.

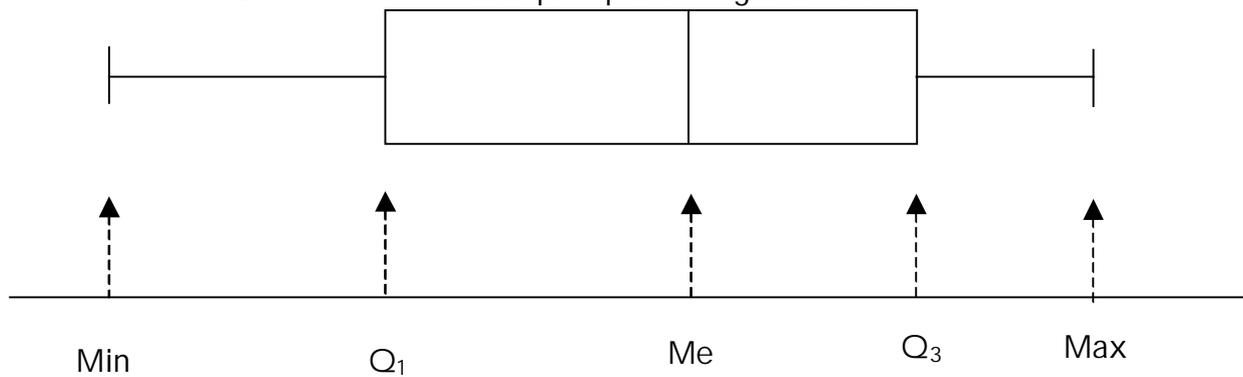
Nombre = $\dots\dots\dots$

Exprimer ce nombre en pourcentage par rapport au nombre total de personnes. Cela vous semble-t-il cohérent par rapport à l'affirmation de l'encadré précédent.

Pourcentage =

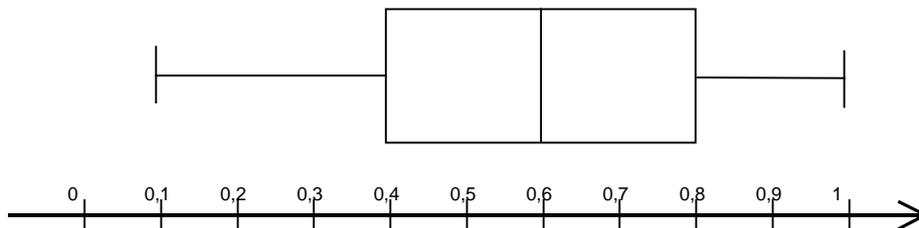
2) La boîte à moustaches :

J On représente sur un graphique que l'on appelle « boîte à moustaches » un certain nombre d'indicateurs statistiques. Chaque « boîte » est délimitée par les premiers et troisièmes quartiles, et les « moustaches » par les valeurs minimales et maximales de la série associée. La médiane est marquée par le segment vertical à l'intérieur de la boîte.



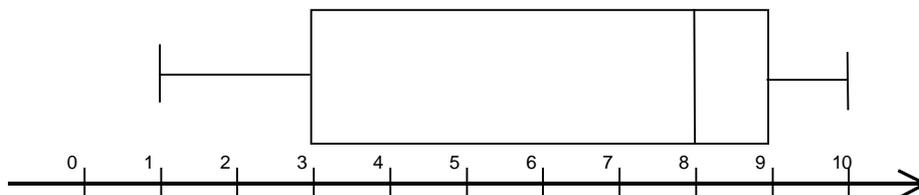
Exercice N°1 : Compléter le tableau en y reportant la valeur des indicateurs statistiques donnés par le diagramme en boîte à moustaches.

a)



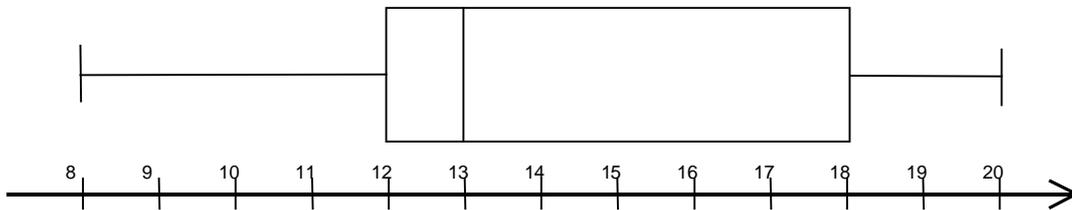
Minimum	Premier quartile	Médiane	Troisième quartile	Maximum

b)



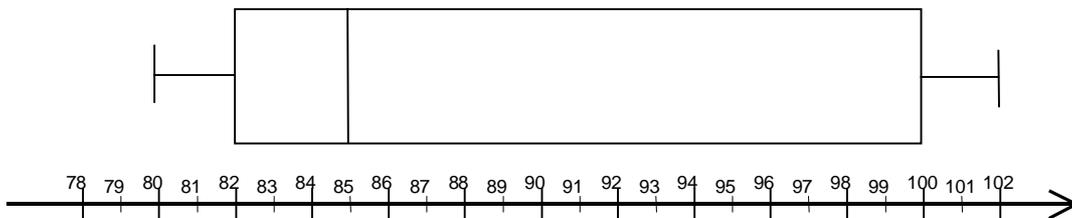
Minimum	Premier quartile	Médiane	Troisième quartile	Maximum

c)



Minimum	Premier quartile	Médiane	Troisième quartile	Maximum

d)



Minimum	Premier quartile	Médiane	Troisième quartile	Maximum

Exercice N°2 : La glycémie à jeun est un test sanguin effectué lorsque le patient a passé 12 heures sans s'alimenter ni boire (sauf de l'eau). Pour une personne non diabétique, le taux normal de glycémie est compris entre 0,70 mg/L et 1,10 mg/L.

Un laboratoire a réalisé des analyses de glycémie à jeun sur 50 personnes. Les résultats sont regroupés dans le tableau suivant :

Glycémie x_i (en mg/L)	Nombre de personnes n_i
0,65	1
0,68	1
0,74	3
0,81	5
0,83	10
0,87	11
0,92	8
0,96	6
1,00	1
1,02	1
1,11	2
1,13	1

a) Quel est le taux de glycémie le plus fréquent de ces personnes ?

b) Quel est l'écart maximal des taux de glycémie de ces personnes ?

c) Quelle est la glycémie moyenne de ces personnes ? (Arrondir au millième)

d) Quel le taux médian de glycémie de ces personnes ?

e) Quel est le pourcentage de personnes dont le taux de glycémie est normal ?

f) Représenter le diagramme à moustache de cette série.



V) Problèmes :

Problème N°1 : Afin de proposer un service de bus, une enquête a été réalisée sur la durée du trajet, en minutes, des 400 employés de l'entreprise « Fabric » pour se rendre sur leur lieu de travail. Ce service sera mis en place si la durée moyenne du trajet est supérieure à 20 minutes.

a) Compléter le tableau statistique suivant :

Durée du trajet (en min)	Centre de la classe x_i	Nombre d'employés n_i	Fréquence f_i (en %)
[0 ; 10[40	
[10 ; 20[80	
[20 ; 30[90	
[30 ; 40[120	
[40 ; 50[50	
[50 ; 60[20	

b) Quelle est la classe modale ?

c) Déterminer le pourcentage d'employés dont la durée du trajet est supérieure ou égale à 20 min.

d) Déterminer, à l'aide d'une calculatrice, la moyenne de cette série statistique.

e) Indiquer si le service de bus sera mis en place par l'entreprise. Justifier la réponse.

Problème N°2 : La vente du bois de chauffage ne peut pas se faire au « poids » car la masse varie de manière importante suivant les essences et le taux d'humidité du bois. On mesure le taux d'humidité en pourcentage des bûches livrées à l'aide d'un appareil.

Les mesures sont rassemblées dans le tableau suivant :

Indications de l'appareil	Centre de classes	Nombre de bûches testées
[10 ; 14[410
[14 ; 18[820
[18 ; 22[1 100
[22 ; 26[670

a) Déterminer l'indication moyenne de l'appareil ainsi que son écart type.

b) On estime que le bois est prêt à l'emploi si l'indication moyenne de l'appareil est inférieure à 20 et l'écart type inférieur à 3. Indiquer si le bois livré est prêt à l'emploi.

Problème N°3 : Une machine produit des pièces usinées pour des accessoires de couture. On prélève un échantillon aléatoire de 120 pièces et on mesure leur rayon. Les résultats obtenus sont les suivants :

Rayon x_i (en mm)	Nombre de pièces n_i
316	2
317	10
318	30
319	50
320	22
321	6

a) Déterminer le mode et l'étendue de cette série. Interpréter les résultats obtenus.

b) Déterminer, à l'aide d'une calculatrice, le rayon moyen m de cette série de pièces et l'écart type (arrondir à l'unité).

c) Le rayon annoncé par l'entreprise est 319 mm : celui-ci correspond au rayon \bar{x} programmé par l'entreprise sur ses machines-outils. Une étude statistique sur les performances des machines-outils achetées par cette entreprise a montré que, pour une dimension programmée \bar{x} , les dimensions effectivement obtenues ont graphiquement la forme d'une courbe de Gauss de moyenne \bar{x} et d'écart type $s = 1$ mm.

à Calculer $\bar{x} - 2s =$

$\bar{x} + 2s =$

à Calculer, parmi les valeurs observées du tableau, le pourcentage de celles qui appartiennent à l'intervalle $[\bar{x} - 2s ; \bar{x} + 2s]$ (arrondir au centième).

à On rejette les pièces dont le rayon se situe en dehors de l'intervalle $[\bar{x} - 2s ; \bar{x} + 2s]$. Déterminer le nombre, puis le pourcentage de pièces rejetées dans le lot considéré (arrondir au centième).

VI) Quelques fonctions d'OpenOffice utilisables en Statistique :

On suppose une série de valeurs contenues dans le tableau (A1:H3). Cette série correspond aux relevés des prix du pain (en €) dans 24 boulangeries.

	A	B	C	D	E	F	G	H
1	1,10 €	0,70 €	0,80 €	1,00 €	0,90 €	0,75 €	0,90 €	1,00 €
2	1,00 €	0,90 €	0,95 €	0,80 €	1,00 €	0,90 €	0,80 €	0,90 €
3	0,70 €	0,90 €	1,20 €	1,10 €	0,90 €	0,80 €	0,90 €	1,00 €

J Organiser des données		
=SOMME(A1:H3)	21,90 €	Somme des valeurs.
=SI (A1=1,1;1;0)	1	1 si la cellule A1 contient 1,1 et 0 dans le cas contraire.
=NB.SI (A1:H3;0,9)	8	Nombre de cellule où 0,9 apparaît.
=NB.SI (A1:H3;">0,9")	9	Nombre de cellules où la valeur est supérieure à 0,9. Attention 0,9 est mis entre guillemets " " sans espace.
= NB.SI (A1:H3;"<1,1")-NB.SI (A1:H3;"<0,9")	14	Nombre de cellules où la valeur appartient à [0,9 ; 1,1[
J Calculer des indicateurs de tendance centrale		
=MODE(A1:H3)	0,9	Mode Mo
=MOYENNE(A1:H3)	0,91 €	Moyenne \bar{x}
=MEDI ANE(A1:H3)	0,9	Médiane Me
J Calculer des indicateurs de dispersion		
=MI N(A1:H3)	0,70 €	Valeur minimale x_{min}
=MAX(A1:H3)	1,20 €	Valeur maximale x_{max}
=MAX(A1:H3)-MI N(A1:H3)	0,50 €	Étendue e
=QUARTILE(A1:H3;1)	0,8	Premier quartile Q_1
=QUARTILE(A1:H3;2)	0,9	Deuxième quartile Q_2 correspond à la médiane
=QUARTILE(A1:H3;3)	1	Troisième quartile Q_3
=ECARTYPEP(A1:H3)	0,12	Écart type s Attention il faut écrire écart type avec un P au bout ECARTYPEP.