Série brute.

I Représentation et tableur.

Exercice 1.

Pour l'exercice suivant il faut utiliser le fichier de l'INSEE des populations légales. Les réponses pourront être écrites sur une nouvelle feuille du fichier. L'exercice s'intéresse à la feuille *Communes*.

1. Observer le fichier fourni par l'INSEE, comment les communes sont-elles triées ?
2. Les communes de moins de 10 000 habitants sont recensées par roulement tous les 5 ans, alors que pour les communes dont la population est supérieure ou égale à 10 000 habitants, une enquête par sondage est réalisée chaque année. Quel est le pourcentage des communes concernées par les enquêtes par sondage ?
3. Choisir l'un des diagrammes proposés par le tableur pour représenter les données. Cette représentation vous semble-t-elle satisfaisante ?
4. Enregistrez vos fichiers en précisant vos nom et prénoms, les dates et l'exercice traité. Puis envoyez ces fichiers par mail à aubry.levavasseur@gmail.com

II Indicateurs statistiques de position.

Minimum et maximum.

Le *minimum* (respectivement le *maximum*) d'une série statistique est le plus petit (resp. plus grand) nombre de la série.

Moyenne.

La *moyenne* d'une série est le quotient de la somme des données par le nombre de données.

Médiane.

Définition 1

Lorsque la série statistique ordonnée comporte un nombre impair de données la *médiane* est la donnée centrale.
Lorsque la série statistique ordonnée comporte un nombre pair de données la *médiane* est la moyenne entre les deux données en position centrale.

La démarche pour trouver manuellement une médiane est donc la suivante :

Étape 1. *Ranger dans l'ordre croissant la série des données*.

Étape 2. *Trouver la position centrale (moitié de l'effectif total)*.

Étape 3. *Lire dans la liste ordonnée la donnée qui est en position centrale*.

Exercice 2.

Modélisez la situation suivante puis répondez à la question.

Un pêcheur ne gardera que la moitié des poissons qu'il a pêché en ne gardant que les plus grands.

Voici les tailles des poissons attrapés mesurées en centimètres : 25 - 22 - 31 - 27 - 30 - 17 - 13 - 24 - 23.

Pour savoir à partir de quelle taille il doit garder ou relâcher un poisson nous devons

Déterminer la médiane.

Étape 1. Ranger les poissons dans l'ordre croissant de leur taille : $13 < 17 < 22 < 23 < 24 < 25 < 27 < 30 < 31$.

Étape 2. Trouver la position centrale : il y a 9 tailles différentes et $\frac{9}{2} = 4{,}5$ donc c'est la cinquième taille qui est au centre (4 avant et autant après).

Étape 3. On lit dans les tailles rangées par ordre croissant que la cinquième taille est 24.

La taille médiane est de 24 cm.

Le pêcheur gardera tous les poissons de plus 24 cm.

Exercice 3.

Modélisez la situation suivante puis répondez à la question.

Le gérant d'un club d'échecs a obtenu des offres promotionnelles sur l'achat de logiciels d'échecs. Il n'en n'a pas suffisamment pour tous les membres. Il choisit de n'offrir des promotions qu'aux 50% de plus récents inscrits. D'après ses bases de données les membres du club sont adhérents depuis (en mois) : 12 - 36 - 25 - 12 - 3 - 6 - 45 - 27 . Quels membres bénéficieront de la promotion ?

Il faut rechercher la médiane de la liste des durées d'abonnements.

Déterminons la médiane.

Étape 1. Ranger dans l'ordre croissant la série des données : $3 < 6 < 12 < 12 < 25 < 27 < 36 < 45$

Étape 2. Trouver la position centrale (moitié de l'effectif total) : il y a 8 données donc la position centrale est entre la quatrième et la cinquième valeur $\left(\frac{8}{2} = 4\right)$.

Étape 3. Lire dans la liste ordonnée la donnée qui est en position centrale : il n'y a pas une mais deux valeurs en position centrale alors, par convention, on fait la moyenne entre les deux valeurs en position centrale : $Me = \frac{12+25}{2} = 18{,}5$.

Il offrira les promotions aux membres dont l'ancienneté n'excède pas 18 mois.

Quartiles.

Après avoir séparé les séries de nombres en deux avec la médiane on sépare parfois en 4 quarts avec les quartiles :

Définition 2

Le *premier quartile*, noté Q_1, d'une série statistique numérique est la plus petite valeur prise par le caractère telle qu'au moins 25% des valeurs lui soient inférieures ou égales.
Le *troisième quartile*, noté Q_3, d'une série statistique numérique est la plus petite valeur prise par le caractère telle qu'au moins 75% des valeurs lui soient inférieures ou égales.

Remarques.

1. On dit parfois que la médiane est le deuxième quartile. Elle est donc parfois notée Q_2.
2. La locution « au moins » implique que si une donnée ne correspond pas exactement à 25%, le premier quartile est la valeur supérieure.
3. La détermination d'un quartile suit les trois mêmes étapes que celle de la médiane : ordonner la série, puis trouver la position (un quart ou trois quarts) et, enfin, lecture de la valeur.

Exercice 4.

Dans un laboratoire d'astrophysique, un détecteur a relevé les durées d'attente en heures entre les réceptions successives des particules captées.
75 - 265 - 225 - 402 - 35 - 105 - 411 - 346 - 159 - 229 - 62 - 256 - 431 - 177 - 56 - 144 - 354 - 178 - 386 -294.
Déterminez Q_1 et Q_3. Interprétez ces résultats par une phrase.

Déterminons Q_1.

On réalise les trois mêmes étapes que pour la médiane :

Étape 1. Ordonner la série. $35 < 56 < 62 < 75 < 105 < 144 < 159 < 177 < 178 < 225 < 229 < 256 < 265 < 294 < 346 < 354 < 386 < 402 < 411 < 431$

Étape 2. Déterminez la position de Q_1. Il y a au total 20 données. $\frac{20}{4} = 5$. Donc Q_1 est la cinquième valeur.

Étape 3. Lire la valeur du premier quartile. On lit dans la liste ordonnée que la cinquième valeur est $Q_1 = 105$.

Ce résultat s'interprète en disant que dans 25% des cas le temps d'attente entre deux réceptions est inférieur à 105 heures.

Déterminons Q_3.

En procédant de même pour le troisième quartile :

Étape 1. La série est déjà ordonnée ci-dessus.

Étape 2. Déterminons la position de Q_3. Il y a au total 20 données. $\frac{3\times 20}{4} = 15$. Donc Q_3 est la quinzième valeur.

Étape 3. Lire la valeur du troisième quartile. On lit dans la liste ordonnée que la quinzième valeur est $Q_3 = 265$.

Ce résultat s'interprète en disant que dans 75% des cas le temps d'attente entre deux réceptions est inférieur à 265 heures.

III Indicateurs statistiques de dispersion.

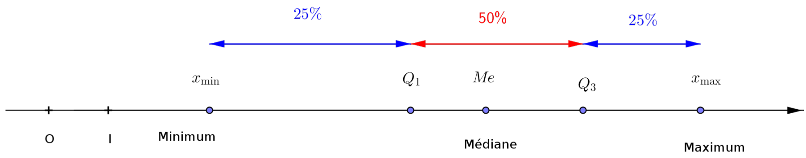
Étendue.

L'*étendue* est l'écart, la distance séparant la plus petite et la plus grande donnée de la série. Autrement dit c'est la différence du maximum et du minimum de la série.

Écart interquartile.

L'*écart interquartile* est l'écart, la distance séparant le premier quartile et le troisième quartile de la série. Autrement dit c'est la différence de Q_3 et de Q_1.

Schématisation.



IV Exercices.

Exercice 5.

Un prof de maths annonce à ses élèves de seconde que seule la moitié des élèves ayant les meilleures notes pourra choisir la spécialité mathématiques. Que seront proposés des spécialités littéraire ou technique au quart des élèves de la classe dont les notes de maths sont les moins bonnes.
Les moyennes obtenues en mathématiques sont les suivantes : 13 - 13 - 12 - 10 - 13 - 11 - 13 - 14 - 11 - 12 - 12 - 13 - 10 - 11 - 12 - 9 - 12 - 14 - 10 - 8 - 8 - 10 - 14 - 6 - 13 - 8 - 10 - 14 - 13 - 13 - 11 - 10.
Interprétez ce problème avec du vocabulaire de statistique descriptive et donnez les notes limites qui détermineront les orientations.

6 8 8 8 9 10 10 10 10 10 10 11 11 11 11 12 12 12 12 12 13 13 13 13 13 13 13 13 14 14 14 14
32 valeurs
médiane 12
$Q_1 = 10$
$Q_3 = 13$

Exercice 6.

Des moines décident de tout partager. Ils mettent en commun les sous dont chacun dispose : 1 523, 2 345, 600, 857, 1 765, 1 365, 1 432, 389, 1 059, 1 391, 1 432, 1 415, 1 399, 1 381, 1 390, 1 408 et 1 370.

1. Puis il redistribuent équitablement les sous afin que chacun aille de son côté utiliser ces ressources pour des œuvres caritatives. Quelle est la part de sous que recevront chacun d'entre eux ? Comment appelle-t-on en statistique descriptive cette valeur ?
2. Comme la parabole de l'ouvrier de la onzième heure affirme que : « Ainsi les derniers seront les premiers, et les premiers seront les derniers » , les moines décident que les premiers à recevoir leur part seront formés du quart qui a le moins donné. Déterminez à partir de quelle somme un moine se trouve dans les premiers à recevoir ?

Correction exercice 6

1. Déterminons la moyenne.

   $\overline{x} \approx 1324{,}7647$.

2. Déterminons le premier quartile.

   (a) Ordonnons la série : $389 \leq 600 \leq 857 \leq 1059 \leq 1365 \leq 1370 \leq 1381 \leq 1390 \leq 1391 \leq 1399 \leq 1408 \leq 1415 \leq 1432 \leq 1523 \leq 1765 \leq 2345$

   (b) $\frac{1}{4} \times N = \frac{1}{4} \times 17 = 4{,}25$. Q_1 est la cinquième valeur.

   (c) $\boxed{Q_1 = 1365 \text{ €}.}$

Exercice 7.

Téléchargez le fichier des populations des communes.

1. (a) Trouvez (ctrl+F) une commune dont la population est de 9 053.
   (b) Colorez en bleu le fond de la cellule contenant le nom de cette commune.
2. Calculez la population moyenne des communes françaises (vous pourrez utiliser la formule « =somme(»).
3. Déterminez la population médiane des communes françaises.
4. Déterminez les premier et troisième quartiles de cette série statistique.
5. Complétez chacune des affirmations suivantes :
   (a) Environ 25% des communes françaises ont une population inférieure à ...
   (b) Environ ... des communes françaises ont une population comprise entre 191 et 1041 habitants.
   (c) Environ 50% des communes françaises ont une population inférieure à ...
   (d) Environ 25% des communes françaises ont une population supérieure à ....
6. Illustrez les résultats obtenus sur une droite graduée.
7. Enregistrez vos fichiers en précisant vos nom et prénoms, les dates et l'exercice traité. Puis envoyez ces fichiers par mail à aubry.levavasseur@gmail.com

Exercice 8.

Quelle note doit-on ajouter à la liste 8 ; 12 ; 15 ; 8 ; 9 ; 14 pour avoir une moyenne égale à 12 ?

Exercice 9.

Téléchargez le fichier des populations des communes des départements de l'Ardèche et de la Drôme.

1. Comparez les populations des villes de deux départements voisins à l'aide des médianes, quartiles, moyennes.
2. Une étude est menée pour un éventuel regroupement de ces deux départements.
   (a) Quelle serait la nouvelle population moyenne ? La nouvelle population médiane ?
   (b) Comment calculer la population moyenne du regroupement à partir des moyennes ?
   (c) Peut on trouver la nouvelle médiane à partir des populations médianes des deux départements ?
3. Enregistrez vos fichiers en précisant vos nom et prénoms, les dates et l'exercice traité. Puis envoyez ces fichiers par mail à aubry.levavasseur@gmail.com

V Ce qu'il faut retenir.

1. Les noms et rôles de chaque indicateur.
2. La façon de trouver les indicateurs.