12

Le championnat de WNBA fait partie de ceux qui font l'objet des statistiques les plus poussées.

Statistiques descriptives

Je dois être capable de…	Proposition de parcours	
Calculer une moyenne.	1 p. 294	1 2 p. 294 18 20 22 p. 296
Utiliser la linéarité de la moyenne.		23 24 26 p. 296-297
Utiliser et interpréter un écart-type.	2 p. 294 TP 1 p. 304	3 4 p. 294 27 29 31 p. 297
Utiliser et interpréter un écart interquartile.	3 p. 295	5 6 p. 295 33 35 p. 297
Décrire et différencier des séries.		61 62 p. 301
Algo & Prog Lire et comprendre une fonction PYTHON renvoyant la moyenne m, l'écart-type s et la proportion d'éléments appartenant à $[m-2s\,;\,m+2s]$.	TP 3 p. 306	

1 exercices résolus — 16 exercices corrigés — 14 exercices non corrigés — TP travaux pratiques

ZOOM SUR...

Logique & Démonstration
p. 291, 298, 299, 303

Algo & Prog
p. 303, 306

TICE
p. 288, 289, 305

Les autres disciplines
p. 294, 298, 300, 302, 305

Problème ouvert
p. 301

Pour prendre un bon départ

Exo — Parcours différenciés
Lienmini.fr/maths2-23

1. Travailler avec une série en ligne

On considère la série statistique 2 ; 7 ; 9 ; 11 ; 13.

1. Déterminer sa moyenne, sa médiane et son étendue.

2. Même question avec 12 ; 1 ; 44 ; 35 ; 5 ; 18 ; 7 ; 102.

2. Travailler avec des effectifs

Tous les week-ends, Kirushika va pêcher et elle note dans un carnet le nombre de poissons qu'elle a pris. Les résultats après deux mois sont donnés dans le tableau ci-contre.

Valeur	0	1	2	3
Effectif	2	3	1	2

1. Recopier et compléter la phrase suivante afin d'interpréter la dernière colonne.
Il y a eu … week-ends durant lesquels Kirushika a pris … poissons.

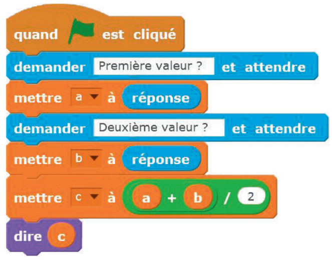
2. Quel est l'effectif de la valeur 1 ?

3. Calculer la fréquence de la valeur 1 exprimée en pourcentage.

4. Calculer l'étendue, la médiane et la moyenne de cette série statistique.

3. Écrire un « programme statistique »

On considère le programme Scratch ci-contre.



1. Que dit le lutin si l'utilisateur saisit successivement 3 et 11 comme valeurs demandées ?

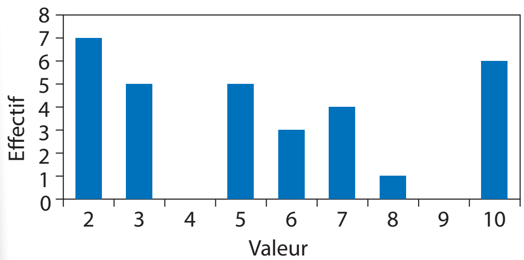
2. Décrire ce que fait ce programme.

3. Comment le modifier pour qu'il calcule la moyenne de trois valeurs ?

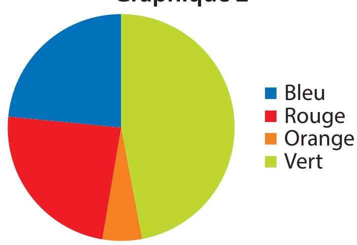
4. Représenter une série statistique

1. Comment s'appellent ces deux représentations graphiques de séries statistiques ?

Graphique 1



Graphique 2



2. Compléter le tableau suivant à l'aide du graphique 1.

Valeur							
Effectif							

3. Compléter le tableau suivant à l'aide du graphique 2 sachant que l'effectif total de la série est 17 (donner l'angle au degré près et les effectifs sont entiers).

Couleur				
Angle				
Effectif				

Doc — Corrigés
Lienmini.fr/maths2-27

Orientation 20 min

1 Calculer une moyenne pondérée

1. Établir une liste de cinq métiers qui paraissent intéressants en menant éventuellement des recherches au CDI.

2. Par groupes de deux ou trois, établir quatre critères qui semblent importants pour le choix d'un métier.

3. Attribuer des coefficients à chacun de ces critères selon son importance aux yeux du groupe.

4. Recopier le tableau ci-contre en complétant :
- la 1re ligne avec les noms des quatre critères choisis ;
- la 1re colonne avec les noms des cinq métiers choisis ;
- l'intérieur du tableau avec une note sur 10 pour chaque critère selon le métier.

Critères / Métiers				

5. Pour chaque métier, calculer sa moyenne pondérée, c'est-à-dire la moyenne des notes sur 10 en considérant que le coefficient donne le nombre de fois où la note du critère « compte » dans le calcul.

6. Établir un classement de ces cinq métiers à partir des résultats de la question précédente.

7. a) Recommencer avec les mêmes métiers mais les critères et coefficients d'un autre groupe.
b) Le classement est-il identique ?

Cours 1 p. 290

TICE 35 min

2 Découvrir la linéarité de la moyenne

1. Préparer une feuille de calcul avec les en-têtes ci-contre.

	A	B	C	D	E	F	G	H
1	Valeurs	Valeurs*k	Valeurs+k	Valeurs/k	Valeurs–k		k=	2

2. Dans la cellule A2, générer un entier aléatoire entre 1 et 100 avec la commande = `ALEA.ENTRE.BORNES(1;100)` puis recopier vers le bas jusqu'à la cellule A21.

3. a) Saisir une formule dans la cellule B2 permettant de compléter la colonne B dans laquelle sont inscrites les valeurs de la colonne A multipliées par la valeur présente dans la cellule H1. Attention, le contenu de la cellule B2 doit s'adapter si celui de H1est modifié.

Coup de pouce Quand on utilise l'adresse d'une cellule dans une formule et que l'on souhaite qu'un élément de cette adresse (colonne et/ou ligne) ne soit pas modifié quand la formule est étendue, on écrit le symbole $, devant la lettre donnant la colonne et/ou le nombre donnant la ligne, qui permet de figer cette lettre ou ce nombre.

b) Recopier vers le bas jusqu'à la cellule B21.

4. Faire de même pour les colonnes C, D et E (adapter au titre de la colonne !).

5. Saisir = `MOYENNE(A2:A21)` dans la cellule A22. À quoi la valeur obtenue correspond-elle ?

6. Quelle va être la moyenne des valeurs de la colonne B ? Vérifier en tirant la formule en A22 d'**une (seule)** cellule vers la droite.

7. Reprendre cette question avec chacune des colonnes.

8. a) Modifier la valeur de *k* en H1 et observer les différentes moyennes dans la ligne 22.
b) Énoncer des règles que l'on peut conjecturer sur la moyenne d'une série pour laquelle on multiplie (respectivement ajoute, divise, etc.) toutes les valeurs par un même nombre.

Cours 1 p. 291

TICE 30 min

3 Découvrir et comprendre l'écart-type

On donne le tableau des scores de certaines joueuses de bowling durant les qualifications du Championnat du monde de 2017 qui s'est joué en six manches successives (source : www.bowling-wm.de).

	A	B	C	D	E	F	G	H	I
1		Manche 1	Manche 2	Manche 3	Manche 4	Manche 5	Manche 6	Score Moyen	Écart-type
2	Robert, Nadine	148	137	166	165	209	200		
3	Goron, Nadia	201	153	160	179	162	172		
4	Tamminen, Tuula	171	163	164	181	176	174		
5	Nordenson, Kristina	239	157	205	232	205	171		
6	Yoshida, Yumiko	194	212	205	198	211	197		

1. Saisir `= MOYENNE(B2:G2)` dans la cellule H2 puis recopier vers le bas afin de remplir la colonne H.

2. a) Peut-on considérer que les trois premières joueuses sont du même niveau ?

b) Classer ces trois joueuses de la moins régulière à la plus régulière.

c) L'écart-type est un indicateur permettant de mesurer si les valeurs d'une série statistique sont plus ou moins proches de sa moyenne : plus il est petit, plus elles en sont proches.
Saisir `= ECARTYPE(B2:G2)` dans la cellule I2 et recopier vers le bas jusqu'à I4 (**et pas plus loin**).

d) Les résultats affichés confirment-ils la réponse à la question **2. b)** ?

3. Laquelle des deux séries de six valeurs des scores de Kristina Nordensen et de Yumiko Yoshida paraît avoir le plus grand écart-type ? Expliquer pourquoi. Vérifier avec le tableur.

↪ Cours 2 p. 292

TICE 45 min

4 Étudier une série avec les quartiles

1. À votre avis, si l'on prend n'importe quel pays appartenant au quart des pays mondiaux les moins peuplés et n'importe quel pays appartenant au quart des pays mondiaux les plus peuplés, par combien au moins faut-il multiplier la population du premier pour obtenir celle du deuxième ?

2. Récupérer le fichier `chapitre12_activite4.ods` sur le site compagnon. Celui-ci contient les populations de la plupart des pays du monde en 2017 (source : banque mondiale).

Doc — Fichier TICE
Lienmini.fr/maths2-37

3. Quelle plage de données contient les populations de tous ces pays ?

4. Dans la cellule E3, saisir `= MIN(B3:B219)`. À quoi la valeur affichée correspond-elle ?

5. Saisir une formule dans la cellule E7 permettant d'obtenir la population la plus élevée en 2017.

6. Sélectionner la plage A3 : B219, puis sélectionner l'onglet Données puis Trier, faire un tri croissant selon la colonne B puis donner les noms des pays ayant les populations minimale et maximale.

8. Saisir `= MEDIANE(B3:B219)` dans la cellule E5. Calculer le pourcentage de valeurs comprises entre le minimum (inclus) et la médiane (incluse) de la série. Pouvait-on s'attendre à un tel résultat ?

9. a) Saisir `= QUARTILE(B3:B219;1)` dans la cellule E4. Cette valeur est appelée 1er quartile, notée Q_1.

b) Calculer le pourcentage de valeurs comprises entre le minimum (inclus) et Q_1 (inclus).

c) Proposer une définition du 1er quartile d'une série statistique.

10. Reprendre et adapter la question **9.** pour le 3e quartile Q_3 dans la cellule E6.

11. Que penser de la réponse à la question **1.** ?

↪ Cours 3 p. 293

Cours

1 Moyenne

a Moyenne et moyenne pondérée

Propriété Moyenne avec effectifs

Soit une série statistique de p valeurs distinctes $x_1 ; x_2 ; \dots ; x_p$ d'effectifs respectifs $n_1 ; n_2 ; \dots ; n_p$ donnée dans le tableau ci-contre.

Valeur	x_1	x_2	…	x_p
Effectif	n_1	n_2	…	n_p

La moyenne de cette série est $m = \dfrac{n_1 \times x_1 + n_2 \times x_2 + \dots + n_p \times x_p}{n_1 + n_2 + \dots + n_p}$.

Exemple

Audrey prend souvent le train venant de Rouen en direction de Paris. En rentrant dans le wagon, elle compte le nombre de places assises disponibles.

Valeur	0	1	2	5	6	7	10
Effectif	5	1	3	1	5	4	1

Après 20 trajets, elle obtient les résultats ci-contre.

Le nombre moyen de places assises disponibles sur ces 20 trajets est :

$$\frac{5\times 0 + 1\times 1 + 3\times 2 + 1\times 5 + 5\times 6 + 4\times 7 + 1\times 10}{5+1+3+1+5+4+1} = 4$$

Remarque Dans la propriété précédente, en posant $n = n_1 + n_2 + \dots + n_p$, on obtient la fréquence de x_1 :
$f_1 = \dfrac{\text{effectif de } x_1}{\text{effectif total}} = \dfrac{n_1}{n}$; celle de x_2 : $f_2 = \dfrac{n_2}{n}$, etc.

On déduit donc la formule $m = \dfrac{n_1}{n} \times x_1 + \dfrac{n_2}{n} \times x_2 + \dots + \dfrac{n_p}{n} \times x_p = f_1 \times x_1 + f_2 \times x_2 + \dots + f_p \times x_p$.

Propriété Moyenne pondérée

On considère une série statistique constituée de p valeurs $x_1 ; x_2 ; \dots ; x_p$ affectées de p coefficients (ou poids) $c_1 ; c_2 ; \dots ; c_p$.
La moyenne pondérée de cette série est $m = \dfrac{c_1 \times x_1 + c_2 \times x_2 + \dots + c_p \times x_p}{c_1 + c_2 + \dots + c_p}$.

Remarque La formule est la même que pour la moyenne d'une série donnée sous forme de tableau d'effectifs, mais il est important de comprendre que les séries sont de natures différentes :

– dans le 1er cas (avec un tableau d'effectifs) la série est constituée de $n_1 + n_2 + \dots + n_p$ valeurs, précisément :
$\overbrace{x_1 ; \dots ; x_1}^{n_1 \text{ fois}} ; \overbrace{x_2 ; \dots ; x_2}^{n_2 \text{ fois}} ; \dots ; \overbrace{x_p ; \dots ; x_p}^{n_p \text{ fois}}$;

– dans le 2e cas (avec les coefficients ou poids) la série est constituée de p valeurs $x_1 ; x_2 ; \dots ; x_p$ (éventuellement identiques) auxquelles on attribue un coefficient (ou poids) $c_1 ; c_2 ; \dots ; c_p$ qui correspond à l'importance de la valeur.

Exemple

Ce trimestre, Émilia a eu quatre contrôles de mathématiques (notés sur 20) de coefficients 1 ; 1,5 ; 4 et 0,5 auxquels elle a obtenu respectivement les notes 8 ; 9 ; 20 et 5.

Sa moyenne en mathématiques est donc $m = \dfrac{1\times 8 + 1{,}5 \times 9 + 4 \times 20 + 0{,}5 \times 5}{1 + 1{,}5 + 4 + 0{,}5} \approx 14{,}9$.

Remarque On constate que, bien qu'elle ait eu 3 notes sur 4 en dessous de 10, sa moyenne est bonne : ceci est dû au fait qu'elle ait eu une très bonne note (20) à un devoir ayant un grand « poids » (donc une plus grande importance) par rapport aux autres.

Exercice résolu 1 p. 294

b Linéarité de la moyenne

Propriété Linéarité de la moyenne

Soit a et b deux nombres réels et x_1 ; x_2 ; … ; x_n une série statistique de moyenne m.

- **Si on multiplie par a toutes les valeurs de la série, on obtient la moyenne de la nouvelle série en multipliant par a la moyenne de la série de départ.**
  Autrement dit, la moyenne de la série ax_1 ; ax_2 ; … ; ax_n est am.
- **Si on ajoute b à toutes les valeurs de la série, on obtient la moyenne de la nouvelle série en ajoutant b à la moyenne de la série de départ.**
  Autrement dit, la moyenne de la série $x_1 + b$; $x_2 + b$; … ; $x_n + b$ est $m + b$.
- **Les deux points précédents assurent également que la moyenne de la série $ax_1 + b$; $ax_2 + b$; … ; $ax_n + b$ est $am + b$.**

Démonstration

- Démontrons directement le dernier point.

On appelle $m = \dfrac{x_1 + x_2 + \ldots + x_n}{n}$ la moyenne de la série statistique constituée des n valeurs x_1 ; x_2 ; … ; x_n.

La moyenne de la série statistique constituée des n valeurs $ax_1 + b$; $ax_2 + b$; … ; $ax_n + b$ est :

$$\begin{aligned}\frac{ax_1 + b + ax_2 + b + \ldots + ax_n + b}{n} &= \frac{ax_1 + ax_2 + \ldots + ax_n + b + b + \ldots + b}{n}\\ &= \frac{a(x_1 + x_2 + \ldots + x_n) + n \times b}{n} = \frac{a(x_1 + x_2 + \ldots + x_n)}{n} + \frac{n \times b}{n}\\ &= a\frac{x_1 + x_2 + \ldots + x_n}{n} + b = am + b.\end{aligned}$$

- Pour $b = 0$, le point précédent justifie bien que la moyenne de la série ax_1 ; ax_2 ; … ; ax_n est $am + 0 = am$.
- Pour $a = 1$, le premier point justifie bien que la moyenne de la série $x_1 + b$; $x_2 + b$; … ; $x_n + b$ est $1 \times m + b = m + b$.

Exemple

La semaine dernière, pour se préparer le matin, Juan a mis 20 minutes en moyenne :

– s'il avait mis deux minutes de plus chaque jour, il aurait mis en moyenne $20 + 2 = 22$ minutes pour se préparer ;

– s'il avait mis 5 % de temps en plus chaque jour, c'est-à-dire si son temps de préparation avait été multiplié par $1 + \dfrac{5}{100} = 1{,}05$ chaque jour, alors il aurait mis en moyenne $20 \times 1{,}05 = 21$ minutes pour se préparer.

Remarque La propriété de linéarité de la moyenne reste vraie lorsque :

– on soustrait un même nombre à toutes les valeurs de la série, puisque soustraire un nombre est équivalent à ajouter son opposé ;

– on divise toutes les valeurs de la série par un même nombre, puisque diviser par un nombre est équivalent à le multiplier par son inverse.

Exemples

1. Si l'on soustrait 5 à toutes les valeurs d'une série statistique de moyenne m_1 alors cela revient à ajouter -5 à toutes ses valeurs. Ainsi, la moyenne de la nouvelle série est $m_1 + (-5) = m_1 - 5$.
2. Si l'on divise par 4 toutes les valeurs d'une série statistique de moyenne m_2 alors cela revient à multiplier par $\dfrac{1}{4}$ toutes ses valeurs. Ainsi, la moyenne de la nouvelle série est $m_2 \times \dfrac{1}{4} = m_2 \div 4$.

2 Écart-type

Définition Écart-type

L'écart-type s **d'une série statistique est un indicateur de dispersion de cette série statistique autour de la moyenne. Concrètement il donne une certaine mesure de l'écart entre les valeurs de la série et la moyenne de celle-ci :**

– plus l'écart-type s d'une série est petit, plus les valeurs de la série sont concentrées autour de la moyenne, donc plus la série est homogène ;

– plus l'écart-type s d'une série est grand, plus les valeurs de la série sont éloignées de la moyenne, donc moins la série est homogène.

Remarque Nous utiliserons la calculatrice pour déterminer l'écart-type (voir le **TP 1**) mais il existe des formules pour le calculer.

- **Pour** la série x_1 ; x_2 ; … ; x_n de moyenne m, l'écart-type est :

$$s=\sqrt{\frac{(x_1-m)^2+(x_2-m)^2+\ldots+(x_n-m)^2}{n}}.$$

- Pour la série donnée par le tableau ci-contre, de moyenne m et d'effectif total $n_1+n_2+\ldots+n_p=n$, l'écart-type est :

Valeur	x_1	x_2	…	x_p
Effectif	n_1	n_2	…	n_p

$$s=\sqrt{\frac{n_1(x_1-m)^2+n_2(x_2-m)^2+\ldots+n_p(x_p-m)^2}{n}}.$$

Exemple

On considère deux entreprises de 10 employés :

– l'entreprise 1 dans laquelle 5 employés gagnent 2 500 € et 5 employés gagnent 3 500 € par mois ;

– l'entreprise 2 dans laquelle 9 employés gagnent 1 200 € et 1 employé gagne 19 200 € par mois.

Le salaire moyen dans l'entreprise 1 est de $\dfrac{5\times 2500+5\times 3500}{10}=3000$ € et le salaire moyen dans l'entreprise 2 est de $\dfrac{9\times 1200+1\times 19\,200}{10}=3\,000$ € également.

On comprend pourtant bien que la répartition des salaires dans les deux entreprises est extrêmement différente : la moyenne seule ne fournit pas une information suffisante pour résumer la série de manière satisfaisante. On utilise alors un indicateur permettant de mesurer l'homogénéité des salaires dans les deux entreprises : l'écart-type.

Avec la calculatrice, on obtient :

– l'écart-type de la série des salaires de l'entreprise 1, qui est de 500 € ;

– l'écart-type de la série des salaires de l'entreprise 2, qui est de 5 400 €.

On constate donc que, bien que le salaire moyen soit le même dans les deux entreprises, 3 000 € :

– dans l'entreprise 1, les employés ont globalement des salaires proches de cette moyenne ;

– dans l'entreprise 2, les employés ont globalement des salaires éloignés de cette moyenne.

Remarques

- On peut utiliser le couple (moyenne ; écart-type) pour résumer une série et en comparer plusieurs.
- Les valeurs éloignées de la moyenne ont de l'influence sur l'écart-type, plus précisément elles le font augmenter.
- Pour les séries dont le diagramme en bâtons (ou l'histogramme) est en forme de cloche, on peut s'attendre à trouver l'essentiel des valeurs de la série (environ 95 %) dans l'intervalle $[m-2s\,;m+2s]$ où m désigne la moyenne et s l'écart-type de la série (voir le **TP 2**).

↪ Exercice résolu 2 p. 294

3 Quartiles et écart interquartile

Définition Quartiles

Le 1^er quartile Q_1 (resp. le 3^e quartile Q_3) d'une série statistique est la plus petite valeur de la série telle qu'au moins 25 % (resp. 75 %) des valeurs de la série lui soient inférieures ou égales.

Exemple

On considère la série ordonnée de 9 valeurs 1 ; 3 ; 7 ; 8 ; 10 ; 11 ; 12 ; 12 ; 58. On a alors :

$\overbrace{\underbrace{1;3}_{\text{Moins de 25 \% des valeurs}};7}^{\text{Plus de 25 \% des valeurs}};8;10;11;12;12;58$ donc $Q_1 = 7$

$\overbrace{\underbrace{1;3;7;8;10;11}_{\text{Moins de 75 \% des valeurs}};12}^{\text{Plus de 75 \% des valeurs}};12;58$ donc $Q_3 = 12$

Propriété Rang des quartiles

Pour une série ordonnée d'effectif n, Q_1 (resp. Q_3) est la k-ième valeur où k est le plus petit entier supérieur ou égal à $\frac{n}{4}$ (resp. $\frac{3n}{4}$).

Exemple

On reprend la série du nombre de places disponibles dans un train sur 20 trajets du paragraphe **1. a.**

Valeur	0	1	2	5	6	7	10
Effectif	5	1	3	1	5	4	1

- Pour trouver Q_1, on calcule $\frac{n}{4} = \frac{20}{4} = 5$ donc Q_1 est la 5^e valeur, c'est-à-dire $Q_1 = 0$ (car la série est 0 ; 0 ; 0 ; 0 ; 0 ; 1 ; 2 ; 2 ; 2 ; etc.).
- Pour trouver Q_3, on calcule $\frac{3n}{4} = \frac{3 \times 20}{4} = 15$ donc Q_3 est la 15^e valeur, c'est-à-dire $Q_3 = 6$.

Propriété Médiane

Pour une série ordonnée d'effectif n, la médiane est :

- **la valeur de rang $\frac{n}{2} + 0{,}5$ si n est impair ;**
- **la moyenne des valeurs de rang $\frac{n}{2}$ et $\frac{n}{2} + 1$ si n est pair.**

Exemple

Dans l'exemple précédent, la série a pour effectif $n = 20$ qui est pair. $\frac{20}{2} = 10$ et $\frac{20}{2} + 1 = 11$ donc la médiane est la moyenne des 10^e et 11^e valeurs, c'est-à-dire $\frac{5+6}{2} = 5{,}5$.

Remarque Pour une série statistique de valeur minimale x_{min} et de valeur maximale x_{max}, chacun des intervalles $[x_{min} ; Q_1]$, $[Q_1$; médiane], [médiane ; $Q_3]$ et $[Q_3 ; x_{max}]$ contient au moins 25 % des valeurs de la série (et environ 25 % si la série est de grand effectif et constituée essentiellement de valeurs différentes).

Définition Écart interquartile

L'écart interquartile d'une série statistique est $Q_3 - Q_1$. Il s'agit d'un indicateur de dispersion.

Exercice résolu 3 p. 295

Remarques

- Plus l'écart interquartile est petit, plus les valeurs « centrales » de la série (celles dans l'intervalle $[Q_1 ; Q_3]$) sont proches les unes des autres. Les valeurs supérieures à Q_3 ou inférieures à Q_1 n'ont pas d'influence sur l'écart interquartile.
- On peut utiliser le couple (médiane ; écart interquartile) pour résumer une série et en comparer plusieurs.