



Math93.com

TD 1 - Seconde

Statistiques

Partie I. Moyenne, Médiane, quartiles et représentation

Exercice 1. Une entreprise de produits alimentaires (c)

Une entreprise de produits alimentaires conditionne des pâtes dans des sachets de 500 grammes.

On suppose que le poids du sachet vide est négligeable.

Pour vérifier le réglage de la machine utilisée pour remplir les sachets, un échantillon aléatoire de 30 sachets est prélevé dans la production ; on mesure la masse de chaque sachet et on calcule la moyenne \bar{m} des masses des sachets de l'échantillon ainsi que l'écart-type σ .

Au cours de la production, l'échantillon suivant a été prélevé.

501	500	500	502	499	497	497	499	498	498
498	500	502	503	500	500	501	500	495	504
501	503	505	502	501	501	497	498	499	496

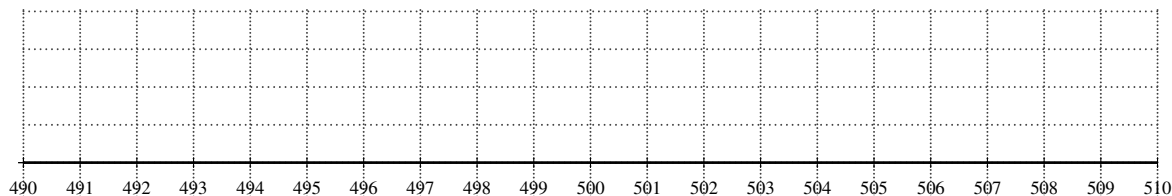
1. Compléter le tableau suivant :

Masses en grammes	495	496	497	498	499	500	501	502	503	504	505
Effectifs											
Effectifs Cumulés croissants											30

2. Calculer la médiane et les quartiles en expliquant votre raisonnement.

3. Vérifier vos résultats à l'aide de la calculatrice.

4. Représenter la dispersion de cette série à l'aide d'un diagramme en boîte.



5. Donner la moyenne \bar{m} de l'échantillon.

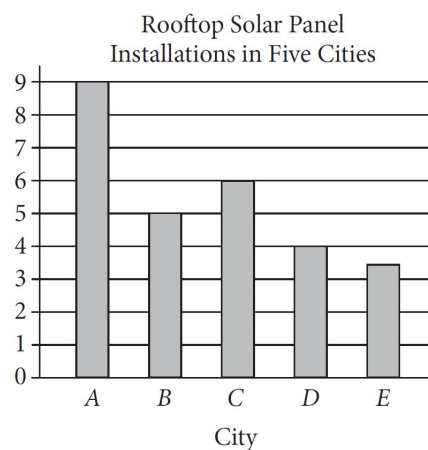
6. Un réglage classique de la machine est nécessaire si l'un des critères suivants n'est pas vérifié :

- les sachets ont une masse en grammes comprise dans l'intervalle $[495; 505]$;
- au moins la moitié des sachets ont une masse en grammes comprise dans l'intervalle $[498; 502]$.

Faut-il effectuer un réglage de la machine ? (Justifier)

Exercice 2. SAT, c'est aussi ma passion! (c)

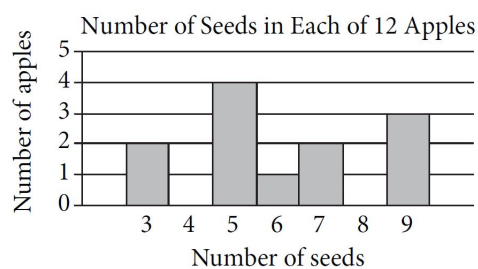
Question 1 (SAT - Practice test 1 - section 4 - Q7)



The number of rooftops with solar panel installations in 5 cities is shown in the graph above. If the total number of installations is 27,500, what is an appropriate label for the vertical axis of the graph?

- a.** Number of installations (in tens)
b. Number of installations (in hundreds)
c. Number of installations (in thousands)
d. Number of installations (in tens of thousands)

Question 2 (SAT - Practice test 1 - section 4 - Q12)



Based on the histogram above, of the following, which is closest to the average (arithmetic mean) number of seeds per apple?

- a.** 4
b. 5
c. 6
d. 7

Question 3 (SAT - Practice test 1 - section 4 - Q13)

		Course			Total
		Algebra I	Geometry	Algebra II	
Gender	Female	35	53	62	150
	Male	44	59	57	160
	Total	79	112	119	310

A group of tenth-grade students responded to a survey that asked which math course they were currently enrolled in. The survey data were broken down as shown in the table above. Which of the following categories accounts for approximately 19 percent of all the survey respondents ?

- a. Females taking Geometry
- b. Females taking Algebra II
- c. Males taking Geometry
- d. Males taking Algebra I

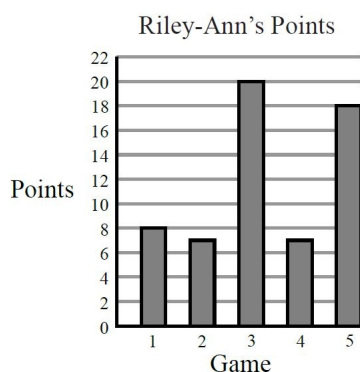
Question 4 (SAT - Practice test 1 - section 4 - Q14)

Lengths of Fish (in inches)						
8	9	9	9	10	10	11
11	12	12	12	12	13	13
13	14	14	15	15	16	24

The table above lists the lengths, to the nearest inch, of a random sample of 21 brown bullhead fish. The outlier measurement of 24 inches is an error. Of the mean, median, and range of the values listed, which will change the most if the 24-inch measurement is removed from the data ?

- a. Mean
- b. Median
- c. Range
- d. They will all change by the same amount.

Question 5 (Gauss contest - 2014 - Q16 (link))



The graph shows points scored by Riley-Ann in her first five basketball games. The difference between the mean and the median of the number of points that she scored is :

- a. 1
- b. 2
- c. 3
- d. 4
- e. 5

Partie II. Écart-type

Exercice 3. Une entreprise de produits alimentaires ... (c)

On reprend l'exercice 1.

- Donner l'arrondi au centième près de l'écart-type σ de la série.
- Un réglage de précision de la machine est nécessaire si l'un des critères suivants n'est pas vérifié :
 - les sachets ont une masse en grammes comprise dans l'intervalle $[495; 505]$;
 - au moins la moitié des sachets ont une masse en grammes comprise dans l'intervalle $[498; 502]$;
 - 95% au moins des sachets ont une masse en grammes comprise dans l'intervalle $[\bar{m} - 2\sigma; \bar{m} + 2\sigma]$.

Faut-il effectuer un réglage de précision de la machine ? (Justifier)

Exercice 4. Des cartouches d'encre (c)

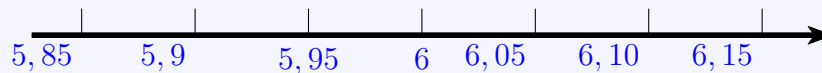
On a relevé le prix d'une même cartouche d'encre dans 350 magasins d'une enseigne de la grande distribution. Les valeurs données sont résumées dans le tableau suivant :

Prix	5,85 €	5,9 €	5,95 €	6 €	6,05 €	6,10 €	6,15 €
Nombre de magasins	27	80	98	56	52	29	8
ECC	27	107	205	261	313	342	350

Si besoin, les résultats seront arrondis au centième d'euro.

- Déterminer :
 - la moyenne,
 - la médiane
 - et les quartiles Q1 et Q3 de la série.
- Préciser l'étendue et l'écart interquartile.
- Construire sur cette feuille le diagramme en boîte (à moustache) de cette série.

A compléter sur cette feuille



- Est-il vrai qu'au moins 50% des valeurs de la série sont entre Q1 et Q3 ?
- Calculer l'écart-type σ de la série ainsi que le pourcentage de valeurs comprises dans l'intervalle $[\bar{m} - 2\sigma; \bar{m} + 2\sigma]$.

Exercice 5. Statistiques ... à la main! (c)

Dans tout l'exercice, les résultats seront si nécessaire arrondis au centième.

On donne la répartition des salaires des 400 employés d'une entreprise (en milliers d'euros). Compléter le tableau ci-dessous et calculer la moyenne des salaires ainsi que la variance et l'écart-type et expliciter clairement la formule utilisée.

Salaire (en milliers d'euros)	[1 ; 2[[2 ; 3[[3 ; 4[Total
Centre des classes x_i				X
Effectifs n_i	150	200	50	400
$n_i \times x_i^2$				

Exercice 6. Statistiques ... à la calculatrice! (c)

Dans tout l'exercice, les résultats seront si nécessaire arrondis au centième.

On lance 600 fois un dé cubique (à six faces), chaque face étant numérotée de 1 à 6. On note les sorties de chacune des six faces.

Face	1	2	3	4	5	6	Total
Effectifs	120	80	122	88	120	70	600
Effectifs cumulés croissants	120	X
Rangs des valeurs	$1^e \rightarrow 120^e$	X

- Déterminer la moyenne \bar{x} , la médiane, et les quartiles Q_1 et Q_3 de la série en expliquant rapidement la méthode utilisée.
- Déterminer à l'aide de la calculatrice et sans détailler les calculs l'écart-type σ de cette série statistique.
- Que pensez-vous de l'affirmation suivante :

Affirmation 1

Au moins 65% des valeurs de cette série statistique appartiennent à l'intervalle $[\bar{x} - \sigma ; \bar{x} + \sigma]$.

Exercice 7. Les 25 plus grandes entreprises mondiales (2010)

Le tableau ci-dessous, répertorie les 25 plus grandes entreprises mondiales (en termes de chiffre d'affaires) selon le classement 2010 Fortune Global 500. Le montant du chiffre d'affaires est en milliards de dollars américain.

RANG	ENTREPRISE	RÉGION	CHIFFRE D'AFFAIRES	BRANCHE D'ACTIVITÉ
1	Wal-Mart Stores	États-Unis	408,2	Commerce de détail
2	Royal Dutch Shell	Europe	285,1	Pétrole
3	Exxon Mobil	États-Unis	284,7	Pétrole
4	BP	Europe	246,1	Pétrole
5	Toyota Motor	Asie	204,1	Automobile
6	Japan Post Holdings	Asie	202,2	Services
7	Sinopec	Asie	187,5	Pétrole
8	State Grid	Asie	184,5	Electricité
9	AXA	Europe	175,3	Assurances
10	China National Petroleum	Asie	165,5	Pétrole
11	Chevron	États-Unis	163,5	Pétrole
12	ING Group	Europe	163,2	Services financiers
13	General Electric	États-Unis	156,8	Société mixte
14	Total	Europe	155,9	Pétrole
15	Bank of America Corp.	États-Unis	150,5	Banque
16	Volkswagen	Europe	146,2	Automobile
17	ConocoPhillips	États-Unis	139,5	Pétrole
18	BNP Paribas	Europe	130,7	Banque
19	Assicurazioni Generali	Europe	126,0	Assurances
20	Allianz	Europe	125,9	Assurances
21	AT&T	États-Unis	123,0	Télécommunications
22	Carrefour	Europe	121,5	Commerce de détail
23	Ford Motor	États-Unis	118,3	Automobile
24	ENI	Europe	117,2	Pétrole
25	J.P. Morgan Chase & Co.	États-Unis	115,6	Banque

On considère la série statistique constituée du chiffre d'affaires des 25 plus grandes entreprises :

- Déterminer le chiffre d'affaires médian ainsi que les premier et troisième quartiles.
- À l'aide de la calculatrice, donner le chiffre d'affaires moyen, la variance et l'écart-type σ .
- Quel pourcentage d'entreprises ont un chiffre d'affaires inférieur au chiffre d'affaires moyen \bar{m} ?
- Que pensez-vous de l'affirmation suivantes :

Affirmation 2

95% au moins des entreprises ont un chiffre d'affaires compris dans l'intervalle $[\bar{m} - 2\sigma; \bar{m} + 2\sigma]$.

Exercice 8. Dans le secteur du cuir

Le tableau suivant donne le salaire brut horaire, par catégorie socioprofessionnelle simplifiée dans le secteur de «l'habillement, cuir» :

	Ouvriers non qualifiés	Ouvriers qualifiés	Employés	Professions intermédiaires	Cadres
Salaire brut en €	9,4	10,6	11,8	16,2	32,0
Nombre de milliers d'heures	33 832	55 920	16 872	23 356	11 759

Source INSEE.

1.

1. a. Calculer l'arrondi à 10^{-1} près du salaire moyen brut \bar{m} . Donner l'écart type σ , arrondi à 10^{-2} près.

1. b. Calculer à 1% près, le pourcentage du nombre de milliers d'heures travaillées pour un salaire brut compris dans l'intervalle $\left[\bar{m} - \frac{2\sigma}{3} ; \bar{m} + \frac{2\sigma}{3} \right]$.

2.

2. a. Quel est le montant du salaire médian ? Déterminer les montants du premier et du troisième quartile.

2. b. Calculer à 1% près, le pourcentage du nombre de milliers d'heures travaillées pour un salaire brut compris dans l'intervalle interquartile.

3. Dans ce cas, quel est le critère le plus pertinent pour le salaire brut, la moyenne ou la médiane ?

Corrections

Correction de l'exercice 1

Une entreprise de produits alimentaires conditionne des pâtes dans des sachets de 500 grammes.

On suppose que le poids du sachet vide est négligeable.

Pour vérifier le réglage de la machine utilisée pour remplir les sachets, un échantillon aléatoire de 30 sachets est prélevé dans la production ; on mesure la masse de chaque sachet et on calcule la moyenne \bar{m} des masses des sachets de l'échantillon ainsi que l'écart-type σ .

Au cours de la production, l'échantillon suivant a été prélevé.

501	500	500	502	499	497	497	499	498	498
498	500	502	503	500	500	501	500	495	504
501	503	505	502	501	501	497	498	499	496

1. Compléter le tableau suivant :



Corrigé

Masses en grammes	495	496	497	498	499	500	501	502	503	504	505
Effectifs	1	1	3	4	3	6	5	3	2	1	1
Effectifs Cumulés croissants	1	2	5	9	12	18	23	26	28	29	30

2. Calculer la médiane et les quartiles en expliquant votre raisonnement.



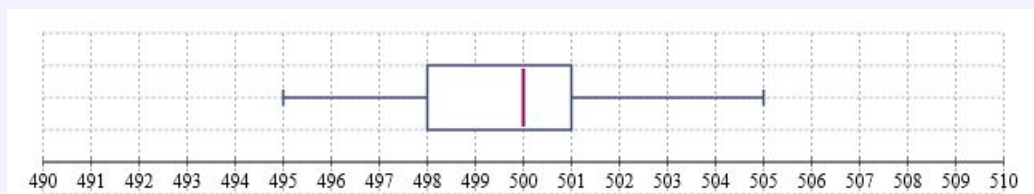
Corrigé

- **Médiane :**
Il y a 30 valeurs, donc la médiane sera entre la 15e (soit 500g) et la 16e valeur (soit 500g).
La médiane de la série est Me=500 grammes ;
- **Quartile 1 :**
Il y a 30 valeurs, et $30/4 = 7,5$ donc le premier quartile sera la 8e valeur soit 498g.
Le premier quartile est Q1=498 grammes ;
- **Quartile 3 :**
Il y a 30 valeurs, et $\frac{30}{4} \times 3 = 22,5$ donc le troisième quartile sera la 23e valeur soit 501g.
Le troisième quartile est Q3=501 grammes.

3. Vérifier vos résultats à l'aide de la calculatrice.
4. Représenter la dispersion de cette série à l'aide d'un diagramme en boîte.



Corrigé



5. Donner la moyenne \bar{m} de l'échantillon.



Corrigé

$$\bar{m} = \frac{495 + 496 + 497 \times 3 + \dots + 505}{30} = 499,9$$

La moyenne des masses des sachets de l'échantillon est $\bar{m} = 499,9$.

6. Un réglage classique de la machine est nécessaire si l'un des critères suivants n'est pas vérifié :

- les sachets ont une masse en grammes comprise dans l'intervalle $[495; 505]$;
- au moins la moitié des sachets ont une masse en grammes comprise dans l'intervalle $[498; 502]$.

Faut-il effectuer un réglage de la machine ? (Justifier)



Corrigé

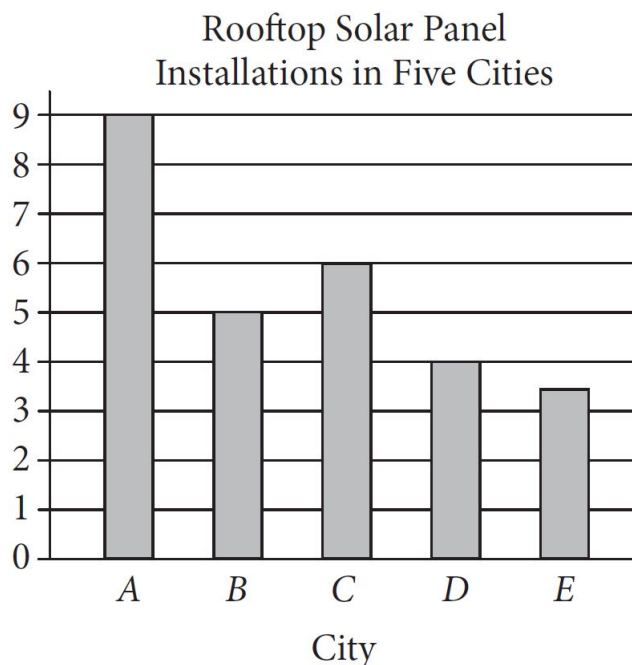
- Tous les sachets ont une masse en grammes comprise dans l'intervalle $[495; 505]$.
- 21 sachets sur 30 ont une masse en grammes comprise dans l'intervalle $[498; 502]$ soit plus de la moitié des sachets ont une masse comprise dans l'intervalle $[498; 502]$. Plus précisément :

$$\frac{21}{30} = 0,7 = \frac{70}{100} > 50\%$$

- Conclusion : aucun réglage n'est nécessaire.

Correction de l'exercice 2 : SAT

Question 6 (SAT - Practice test 1 - section 4 - Q7)



The number of rooftops with solar panel installations in 5 cities is shown in the graph above. If the total number of installations is 27,500, what is an appropriate label for the vertical axis of the graph ?

- a.** Number of installations (in tens)
b. Number of installations (in hundreds)
c. Number of installations (in thousands)
d. Number of installations (in tens of thousands)



Corrigé

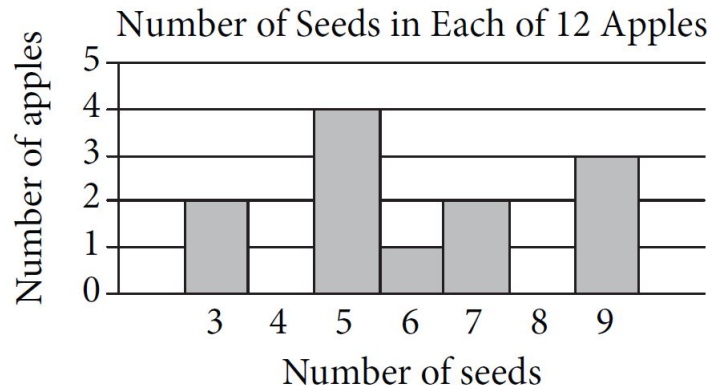
Quite simple by proceeding by elimination, the only possible solution is c.

With a) we have for the first city $9 \times 10 = 90$ panels, it is impossible to have a total of 27,500.

With b) we have for the first city $9 \times 100 = 900$ panels, it is impossible to have a total of 27,500 because $900 \times 5 = 4500$.

With d) we have for the first city $9 \times 10000 = 90000$ panels, which is greater than 27,500

Question 7 (SAT - Practice test 1 - section 4 - Q12)



Based on the histogram above, of the following, which is closest to the average (arithmetic mean) number of seeds per apple?

- a. 4 b. 5 **c. 6** d. 7



Corrigé

The mean is :

$$\bar{m} = \frac{3 \times 2 + 4 \times 2 + 5 \times 4 + 6 \times 1 + 7 \times 2 + 9 \times 3}{2 + 2 + 4 + 1 + 2 + 3} = \frac{73}{12} \approx 6,1$$

Question 8 (SAT - Practice test 1 - section 4 - Q13)

		Course			Total
		Algebra I	Geometry	Algebra II	
Gender	Female	35	53	62	150
	Male	44	59	57	160
	Total	79	112	119	310

A group of tenth-grade students responded to a survey that asked which math course they were currently enrolled in. The survey data were broken down as shown in the table above. Which of the following categories accounts for approximately 19 percent of all the survey respondents?

- a. Females taking Geometry b. Females taking Algebra II **c. Males taking Geometry** d. Males taking Algebra I



Corrigé

19 percent of all the 310 survey respondents is :

$$\frac{19}{100} \times 310 = 58,9$$

This result fit with the 59 males taking geometry.

Question 9 (SAT - Practice test 1 - section 4 - Q14)

Lengths of Fish (in inches)						
8	9	9	9	10	10	11
11	12	12	12	12	13	13
13	14	14	15	15	16	24

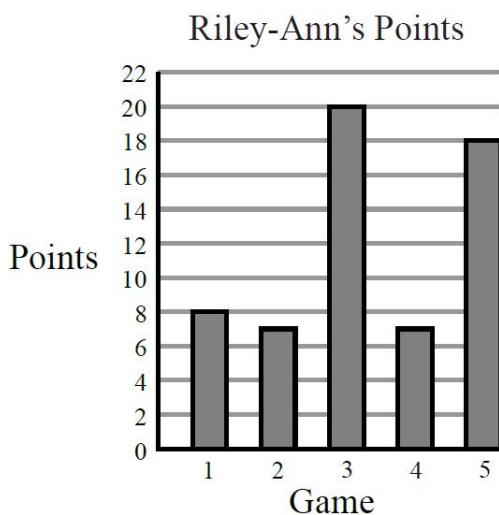
The table above lists the lengths, to the nearest inch, of a random sample of 21 brown bullhead fish. The outlier measurement of 24 inches is an error. Of the mean, median, and range of the values listed, which will change the most if the 24-inch measurement is removed from the data ?

- a. Mean b. Median c. Range d. They will all change by the same amount.



Corrigé

Question 10 (Gauss contest - 2014 - Q16 (link))



The graph shows points scored by Riley-Ann in her first five basketball games. The difference between the mean and the median of the number of points that she scored is :

- a. 1 b. 2 c. 3 **d. 4** e. 5



Corrigé

- The mean is :

$$m \frac{8 + 7 + 20 + 7 + 18}{5} = 12$$

- There is 5 values and so the median is the 3th value which is 8.

- The difference between the mean and the median of the number of points that she scored is $12 - 8 = 4$.

Correction de l'exercice 3

1. Donner l'arrondi au centième près de l'écart-type σ de la série.



Corrigé

L'écart-type obtenu à la calculatrice est $\sigma = 2,33$.

2. Un réglage de précision de la machine est nécessaire si l'un des critères suivants n'est pas vérifié :

- les sachets ont une masse en grammes comprise dans l'intervalle $[495; 505]$;
- au moins la moitié des sachets ont une masse en grammes comprise dans l'intervalle $[498; 502]$;
- 95% au moins des sachets ont une masse en grammes comprise dans l'intervalle $[\bar{m} - 2\sigma; \bar{m} + 2\sigma]$.

Faut-il effectuer un réglage de précision de la machine ? (Justifier)



Corrigé

- On a vu que les deux premiers critères étaient vérifiés dans l'exercice 1.
- L'intervalle

$$[\bar{m} - 2\sigma; \bar{m} + 2\sigma] = [499,9 - 2 \times 2,33; 499,9 + 2 \times 2,33] = [495,24; 504,56]$$

Or 28 sachets sur 30 ont une masse en grammes comprise dans l'intervalle $[495,24; 504,56]$ soit :

$$\frac{28}{30} \approx 93\% < 95\%$$

- Conclusion : Moins de 95% des sachets de l'échantillon ont une masse comprise dans l'intervalle $[\bar{m} - 2\sigma; \bar{m} + 2\sigma]$ par conséquent, il faut effectuer un réglage de la machine.

Correction de l'exercice 4

Prix	5,85 €	5,9 €	5,95 €	6 €	6,05 €	6,10 €	6,15 €
Nombre de magasins	27	80	98	56	52	29	8
ECC	27	107	205	261	313	342	350

Si besoin, les résultats seront arrondis au centième d'euro.

1. Déterminer la moyenne, la médiane et les quartiles Q1 et Q3 de la série.



Corrigé

- Moyenne :

$$\bar{m} = \frac{5,85 \times 27 + \dots + 6,15 \times 8}{350} = \frac{2089,75}{350} \approx \underline{5,97}$$

- Médiane : il y a 350 valeurs, donc la médiane est toute valeur comprise entre la 175^e et 176^e soit $Me = \underline{5,95 \text{ €}}$.
- Q1 : il y a 350 valeurs, et $350/4 = 87,5$ donc Q1 sera la 88^e valeur soit $Q1 = \underline{5,9 \text{ €}}$.
- Q3 : il y a 350 valeurs, et $3 \times 350/4 = 262,5$ donc Q3 sera la 263^e valeur soit $Q3 = \underline{6,05 \text{ €}}$.

2. Préciser l'étendue et l'écart interquartile.



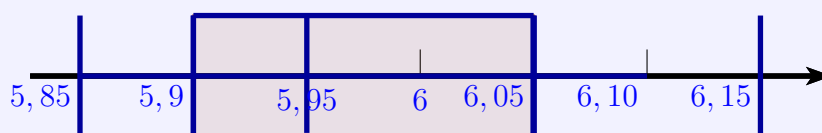
Corrigé

- L'étendue est la différence entre les valeurs extrêmes soit $e = 6,15 - 5,85 = \underline{0,30 \text{ €}}$.
- L'écart interquartile est la différence entre Q3 et Q1 soit $Q3 - Q1 = \underline{0,15 \text{ €}}$.

3. Construire le diagramme en boîte de cette série.



Corrigé



4. Est-il vrai qu'au moins 50% des valeurs de la série sont entre Q1 et Q3?



Corrigé

Il y a $80 + 98 + 56 + 52 = 286$ valeurs entre Q1 et Q3 ce qui représente $286/350 \approx 82\%$ des valeurs. L'affirmation est donc vraie.

Correction de l'exercice 5

Dans tout l'exercice, les résultats seront si nécessaire arrondis au centième.

On donne la répartition des salaires des 400 employés d'une entreprise (en milliers d'euros). Compléter le tableau ci-dessous et calculer la moyenne des salaires ainsi que la variance et l'écart-type et explicitant clairement la formule utilisée.



Corrigé

Salaires (en milliers d'euros)	[1 ; 2[[2 ; 3[[3 ; 4[Total
Centre des classes x_i	1.5	2.5	3.5	X
Effectifs n_i	150	200	50	400
$n_i \times x_i^2$	337.5	1250	612.5	2 200

- La moyenne :

$$\bar{x} = \frac{900}{400} = \underline{2,25}$$

- Variance :

$$V = \frac{337.5 + 1205 + 612.5}{400} - (\bar{x})^2 = \frac{2\,200}{400} - 2,25^2 = \underline{0,4375}$$

- Écart-type :

$$\sigma = \sqrt{V} \approx \underline{0,66}$$

Correction de l'exercice 6

Dans tout l'exercice, les résultats seront si nécessaire arrondis au centième.

1. Déterminer la moyenne, la médiane, et les quartiles Q_1 et Q_3 de la série en expliquant la méthode utilisée.



Corrigé

Face	1	2	3	4	5	6	Total
Effectifs	120	80	122	88	120	70	600
ECC	120	200	322	410	530	600	X
Rangs	1 ^e → 120 ^e	121 ^e → 200 ^e	201 ^e → 322 ^e	323 ^e → 410 ^e	410 ^e → 530 ^e	531 ^e → 600 ^e	X

- La moyenne est :

$$m = \frac{120 \times 1 + 80 \times 2 + \dots + 70 \times 6}{600} = \frac{2\,018}{600} \approx \underline{3,36}$$

- Il y a 600 valeurs, donc on prendra comme médiane une valeur comprise entre les 300^e et 301^e valeur. La 300^e valeur est 3 et la 301^e aussi donc la médiane est 3.
- Il y a 600 valeurs et $0,25 \times 600 = 150$.
Le premier quartile sera donc la 150^e valeur soit $\boxed{Q_1 = 2}$.
- Il y a 600 valeurs et $0,75 \times 600 = 450$.
Le troisième quartile sera donc la 450^e valeur soit $\boxed{Q_3 = 5}$.

2. Déterminer à l'aide de la calculatrice et sans détailler les calculs l'écart-type σ de cette série statistique.



Corrigé

$$\sigma \approx \underline{1,67}$$

3. Que pensez-vous de l'affirmation suivante :

Affirmation 3

Au moins 65% des valeurs de cette série statistique appartiennent à l'intervalle $[\bar{x} - \sigma ; \bar{x} + \sigma]$.

**Corrigé**

- On a :

$$[\bar{x} - \sigma ; \bar{x} + \sigma] \approx [1,69 ; 5,04]$$

- Il y a alors 410 valeurs sur les 600 comprises entre 2 et 5 et donc la fréquence cherchée est :

$$f = \frac{410}{600} \approx 68,33\% > 65\%$$

L'affirmation est donc vraie.