MATHS 2^{nd}	Plan de travail - Ch 12: Statistiques descriptives	CORRECTION

[BLOC 1] I – Série statistique

❖ Dresser et interpréter un tableau

50 **1.**

Nombre de spams	0	1	2	3	4	5	6	7	8	9	10
Effectif	3	0	1	4	4	2	1	5	3	1	1

2. $3+0+1+4+4+2+1+5+3+1+1=25$
L'effectif total de la classe est 25.
3. a) « Plus de 7 spams » correspond à 8 ou 9 ou 10 spams. Donc 5 élèves ont reçu plus de 7 spams.
b) « Au plus 5 spams » correspond à 0 ou 1 ou 2 ou 3 ou 4 ou 5 spams. Donc 14 élèves ont reçu au plus 5 spams.

[BLOC 1] II – Indicateurs de tendance centrale

❖ Calculer une moyenne pondérée

3 **a)** $m=\dfrac{3\times3+7\times6+4\times7+6\times8{,}5}{3+7+4+6}=6{,}5$
Le prix moyen de mugs dans ce magasin est 6,50 €.
b) Grâce à la linéarité de la moyenne, le prix moyen des mugs après cette réduction s'obtient en effectuant une baisse de 10 % du prix moyen initial : $0{,}9\times6{,}5=5{,}85$.
Le prix moyen de mugs après réduction est 5,85 €.

15

Centre	98,25	98,75	99,25	99,75	100,25	100,75	Total
Effectif	9	17	31	26	13	4	100

$m\approx\dfrac{9\times98{,}25+17\times98{,}75+\ +4\times100{,}75}{100}$ soit

$m\approx\dfrac{9\,939{,}5}{100}$

La moyenne de cette série est environ 99,40.

13 La moyenne de cette série est :
$0{,}17\times54+0{,}19\times59+0{,}48\times68+0{,}16\times88=67{,}11$

14 $0{,}2\times1+0{,}4\times2+0{,}25\times3+0{,}1\times4+0{,}05\times5=2{,}4$
Donc il y a 2,4 enfants par famille en moyenne.

❖ Déterminer la médiane et les quartiles d'une série

27 Série ordonnée : 0 – 1 – 1 – 2 – 3 – 3 – 4 – 4 – 5 – 5 – 5 – 6 – 7 – 10

a) $\text{Me} = 4$, $Q_1 = 2$ et $Q_3 = 5$.

b) Cela n'entraîne pas de modifications car $7 > Q_3$.

4 **a)** L'effectif total de la série est 20, donc la médiane est la demi-somme des 10^e^ et 11^e^ valeurs. Donc $\text{Me} = \frac{6\text{ €} + 7\text{ €}}{2} = 6{,}50$ €. Ainsi, au moins 50 % des mugs ont un prix inférieur ou égal à 6,50 € et au moins 50 % des mugs ont un prix supérieur ou égal à 6,50 €.

b) L'effectif total est maintenant égal à 25. La médiane de la nouvelle série est la valeur de rang 13, soit 7 €. Ainsi, au moins 50 % des mugs ont un prix inférieur ou égal à 7 € et au moins 50 % des mugs ont un prix supérieur ou égal à 7 €.

7 **a)** $\frac{1}{4} \times 30 = 7{,}5$ donc le 1^er^ quartile est la 8^e^ valeur de la série. Ainsi $Q_1 = 19$.

$\frac{3}{4} \times 30 = 22{,}5$ donc le 3^e^ quartile est la 23^e^ valeur de la série. Ainsi $Q_3 = 20$.

b) Au moins 25 % des sachets contiennent un nombre de vis inférieur ou égal à 19. Au moins 75 % des sachets contiennent un nombre de vis inférieur ou égal à 20.

51 **a)**

Valeur	35	37	38	39	41	42	43	45
Effectif	5	5	7	10	10	2	9	5
Effectif cumulé croissant	**5**	**10**	**17**	**27**	**37**	**39**	**48**	**53**

L'effectif total de la série est $N = 53$.

b) $N = 53$ donc la médiane est la 27^e^ valeur : $\text{Me} = 39$.

c) $\frac{N}{4} = 13{,}25$ donc Q_1 est la 14^e^ valeur de la série : $Q_1 = 38$.

$\frac{3N}{4} = 39{,}75$ donc Q_3 est la 40^e^ valeur de la série : $Q_3 = 43$.

26

Valeur	5	15	20	25	30	35
Effectif	3	1	2	4	4	2
Effectif cumulé croissant	3	4	6	10	14	16

$\text{Me} = 25$; $Q_1 = 15$; $Q_3 = 30$.

28 **1.** L'effectif total est N=45 000. $\frac{N}{4} = 11250$ et $\frac{3N}{4} = 33750$.

Ainsi $Q_1 = 2$, $\text{Me} = 2$ et $Q_3 = 3$.

2. a) « Au moins **75 %** des foyers possèdent trois véhicules au plus » ;

b) « Au moins 50 % des foyers possèdent **deux** véhicules ou moins ».

[BLOC 2] III – Indicateurs de dispersion

❖ **Utiliser des indicateurs de dispersion (étendue, écart-interquartile, écart-type)**

33 L'étendue de la série est égale à :

$\max - \min = 53 - 27 = 26.$

L'écart interquartile de la série est égal à :

$Q_3 - Q_1 = 42 - 31 = 11.$

34 L'étendue de la série est égale à $234 - 78 = 156$.

8 **a)** La moyenne m de cette série est égale à 1 678 €.

x_i	1 100	1 200	1 500	1 800	2 200	2 800
$x_i - m$	– 578	– 478	– 178	122	522	1 122
$(x_i - m)^2$	334 084	228 484	31 684	14 884	272 484	1 258 884

$$V = \frac{6 \times 334\,084 + 10 \times 228\,484 + \ldots + 4 \times 1\,258\,884}{6 + 10 + 9 + 14 + 7 + 4}$$

donc $V = 234\,516$ et $s \approx 484{,}27$ €.

b) On détermine les bornes de l'intervalle : $m - 2s \approx 709{,}46$ € et $m + 2s \approx 2\,646{,}54$ €.

$6 + 10 + 9 + 14 + 7 = 46$ donc 46 employés ont un salaire dans l'intervalle $[m - 2s\,;\, m + 2s]$.

39 **a)** $m = \frac{1 \times 3 + 2 \times 4 + \ldots + 2 \times 10}{1 + 2 + \ldots + 2} = 6{,}65$

donc $m \approx 6{,}7$.

b)

Valeur x_i	3	4	5	6	7	8	9	10
$^*(x_i - m)^2$	13,32	7,02	2,72	0,42	0,12	1,82	5,52	11,22
Effectif	1	2	7	10	8	6	4	2

* Les carrés des écarts sont arrondis au centième.

$$V \approx \frac{1 \times 13{,}32 + 2 \times 7{,}02 + \ldots + 2 \times 11{,}22}{40}$$

$$V \approx \frac{107}{40} \text{ et } V \approx 2{,}7$$

Donc $s = \sqrt{V}$ et $s \approx 1{,}6$

44 **a)** Par lecture graphique, la série de gauche semble avoir des valeurs plus dispersées autour de sa moyenne que la série de droite. Son écart-type est donc plus grand que celui de la série de droite.

b) Série de gauche

L1	L2
20	3
40	2
60	1
80	5
100	3

$m \approx 64{,}3$
$s \approx 29{,}5$

Stats 1 var
$\bar{x}$=64.28571429
Σx=900
Σx²=70000
Sx=30.56249228
σx=29.45075447
n=14
minX=20

Série de droite

L1	L2
20	1
40	4
60	5
80	3
100	1

$m \approx 58{,}6$
$s \approx 20{,}7$

Stats 1 var
$\bar{x}$=58.57142857
Σx=820
Σx²=54000
Sx=21.43223412
σx=20.65261756
n=14
minX=20

40 **a)**

L1	L2
8.7	1
9.9	3
10.4	4
10.7	6
12.1	5
12.7	2

b)

Stats 1 var
$\bar{x}$=10.95714286
Σx=230.1
Σx²=2543.93
Sx=1.06516263
σx=1.039492288
n=21
minX=8.7

$m \approx 10{,}96$ et $s \approx 1{,}04$

c) On détermine les bornes de l'intervalle $[m - s\,; m + s]$:

$m - s \approx 9{,}92$ et $m + s \approx 12$.

$\dfrac{4+6}{21} \approx 0{,}48$ donc environ 48 % des données se trouvent dans l'intervalle $[m - s\,; m + s]$.

On détermine les bornes de l'intervalle $[m - 2s\,; m + 2s]$:

$m - 2s \approx 8{,}88$ et $m + 2s \approx 13{,}04$.

$\dfrac{3+4+6+5+2}{21} \approx 0{,}95$ donc environ 95 % des données se trouvent dans l'intervalle $[m - 2s\,; m + 2s]$.

53 **a)** Pour la série des filles :

$Q_3 - Q_1 = 168 \text{ cm} - 159 \text{ cm} = 9 \text{ cm}$

Pour la série des garçons :

$Q_3 - Q_1 = 181 \text{ cm} - 165 \text{ cm} = 16 \text{ cm}$

b) D'après **a)** les tailles des garçons sont plus dispersées autour de la médiane Me que celles des filles.

c) L'écart-type de la série des garçons est supérieur à celui de la série des filles. Donc les tailles des garçons sont plus dispersées autour de la moyenne m que celles des filles.

64 **a)** Avec la calculatrice, on obtient : $m \approx 25{,}55$ et $s \approx 3{,}89$.

b) On détermine les bornes de l'intervalle :

$m - 2s \approx 17{,}77$ et $m + 2s \approx 33{,}33$.

$\dfrac{451}{471} \approx 0{,}958$ donc 95,8 % des projets ont une durée de réalisation dans l'intervalle $[m - 2s\,; m + 2s]$.

85 La moyenne est égale à 45 et l'écart-type à 7,5.
En soustrayant 45 puis en divisant par 7,5 chaque valeur de la série, on obtient une nouvelle série de moyenne égale à 0 et d'écart-type égal à 1.
Il ne reste plus qu'à multiplier chaque valeur par 3 puis ajouter 10 afin de répondre au problème.
Si x est une des valeurs initiales, la nouvelle valeur est alors :

$$\frac{x-45}{7,5}\times 3+10=0,4x-8$$

Donc $a=0,4$ et $b=-8$.
La nouvelle série est 4 ; 10 ; 10 ; 10 ; 13 ;13.

79 On obtient des valeurs approchées de m et s en prenant les milieux des classes.

Répartition initiale :

L1	L2
1350	201
1600	213
1850	350
2250	500
2750	233
3250	51
5250	9

Stats 1 var
$\bar{x}$=2079.897238
Σx=3238400
Σx²=7189540000
Sx=540.1613486
σx=539.9878585
n=1557
minX=1350

Méthode 1 :

L1	L2
1409.5	201
1652	213
1894.5	350
2282.5	500
2767.5	233
3252.5	51
5192.5	9

Stats 1 var
$\bar{x}$=2117.500321
Σx=3296948
Σx²=7408457786
Sx=523.9565081
σx=523.7882227
n=1557
minX=1409.5

Méthode 2 :

L1	L2
1350	191
1600	213
1850	350
2250	500
2750	233
3250	51
5250	19

Stats 1 var
$\bar{x}$=2104.945408
Σx=3277400
Σx²=7446940000
Sx=593.5557532
σx=593.3651138
n=1557
minX=1350

La méthode 1 augmente le salaire moyen et réduit les disparités puisque l'écart-type diminue.
La méthode 2 augmente le salaire moyen mais accentue les disparités puisque l'écart-type augmente.
La méthode 1 est donc la plus appropriée.