Série d'exercices corrigés Statistiques

Exercice 1 (corrigé)

Dans un sous-groupe de 40 personnes la taille moyenne est égale à $170\ cm$.
Dans un deuxième sous-groupe de 10 personnes la taille moyenne est égale à $180\ cm$.
Dans un troisième sous-groupe de 50 personnes la taille moyenne est égale à $175\ cm$.
1) Déterminer la taille moyenne du groupe constitué par les trois sous-groupes précédents. 2) Quelle serait la taille moyenne si les trois sous-groupes étaient constitués du même nombre de personnes ?

Exercice 2 (corrigé)

La température est relevée chaque heure pendant 4 jours dans une forêt de Ain Drahem. Les 97 résultats obtenus ont été triés et sont rassemblés dans le tableau suivant :

Température	14,5	15	15,5	16	16,5	17	17,5	18	18,5	19	19,5
Nombre de fois où cette température a été relevée	5	7	10	12	15	10	11	9	7	7	4

Déterminer la médiane M, les quartiles Q_1 et Q_3 de celle série statistique.

Exercice 3 (corrigé)

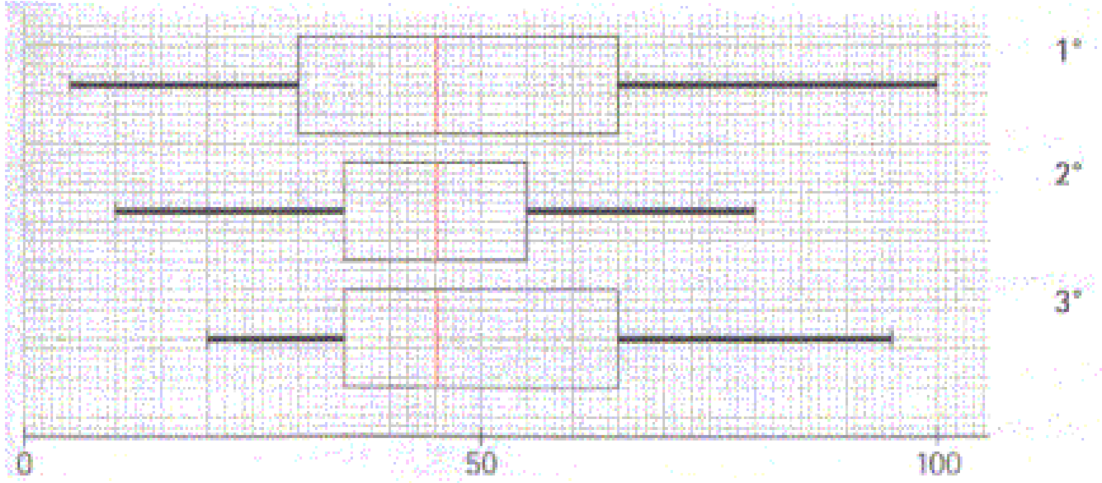
Le tableau ci-dessous donne la répartition des salaires mensuels, des employés d'une entreprise :

Salaire (Dinars)	[800 ; 900[	[900 ; 1000[	[1000 ; 1050[	[1050 ; 1150[	[1150 ; 1300[
Effectif	42	49	74	19	16

1) Calculer le salaire moyen dans cette entreprise. Que penser d'un tel résultat ?
2) Dans cette entreprise, combien d'employés gagnent au plus 1050 Dinars ?
3) Dresser le polygone des effectifs cumulés croissants et lire une valeur approchée de la médiane et de Q_1 et de Q_3.
4) Calculer de manière précise la médiane et les quartiles Q_1 et Q_3.
5) Construire le diagramme en boîte de la série statistique.

Exercice 4 (corrigé)

Sur chacun des diagrammes ci-dessous, lire l'étendue, la médiane, les quartiles et les écarts interquartiles.



Exercice 5 (corrigé)

Le tableau suivant donne les températures moyennes par mois à **Serbie** et à **Kosovo** en degrés Celsius.

Mois	Jan	Fév.	Mars	Avril	Mai	Juin	Juillet	Aout	Sep.	Oct.	Nov.	Déc.
Serbie	–5	– 4	4	15	27	31	31	30	26	20	10	–5
Kosovo	3	4	7	10	14	17	19	18	16	17	7	6

1) Calculer la moyenne, l'étendue, la variance et l'écart-type des températures mensuelles pour chacune de ces deux villes.
2) Comparer et analyser les résultats obtenus.

Exercice 6 (corrigé)

Lors d'un examen écrit, un correcteur a obtenu les notes suivantes (sur 20), sur 80 copies
11 ;11 ;11 ;7 ;6 ;13 ;13 ;7 ;4 ;9 ;5 ;10 ;11 ;8 ;14 ;15 ;8 ;10 ;4 ;9 ;7 ;7 ;9 ;12 ;10 ;14 ;18 ;6 ;9 ;
10 ;13 ;9 ;12 ;8 ;10 ;5 ;7 ;13 ;12 ;12 ;13 ;11 ;9 ;11 ;9 ;8 ;10 ;14 ;10 ;11 ;9 ;7 ;7 ;6 ;10 ;6 ;11 ;10 8 ;8 ;11 ;7 ;6 ;8 ;11 ;12 ;14 ;9 ;12 ;7 ;8 ;8 ;16 ;14 ;9 ;10 ;7 ;10 ;10 ;12.
1) Calculer la moyenne $\bar{x}$ et l'écart type σ de la série.
2) Un échantillon de notes est dit "normal" si environ 30 % des notes sont en dehors de l'intervalle $[\bar{x}-\sigma \ ; \ \bar{x}+\sigma]$ et 5 % en dehors de l'intervalle $[\bar{x}-2\sigma \ ; \ \bar{x}+2\sigma]$. L'échantillon obtenu est-il normal ?

Exercice 7 (corrigé)

Deux classes de 2$^{\text{ème}}$ Economie et services comparent leurs résultats du 1$^{\text{er}}$ trimestre et déclarent : "nos classes ont le même profil puisque dans les deux cas la médiane des résultats est 10". Qu'en pensez-vous ?

notes	5	6	7	8	9	10	11	12	13	14	15	16
effectifs 2$^{\text{ème}}$ Economie et services 1	0	3	4	4	5	7	3	4	2	1	0	0
effectifs 2$^{\text{ème}}$ Economie et services 2	2	4	3	3	3	4	3	2	2	3	1	2

1) Vérifier que les deux médianes valent 10 et déterminer les quartiles de chaque série
2) Tracer côte à côte les diagrammes en boites de ces deux séries.

Exercice 8

Une loterie a été organisée avec des gains en argent liquide. Tous les billets n'ont pas été vendus. Le tableau ci-dessous résume les gains effectivement perçus par les joueurs :

Gain En Dinars	100	200	300	400	500	600	700	800	900	1000
Effectif	2	1	1	3	2	2	3	5	0	1

Partie A : Analyse de la série statistique

1) Combien y a-t-il de gagnants à cette loterie ? (personne n'a gagné plus d'une fois).
2) Quel a été le gain moyen parmi les gagnants ?
3) a) Quelle est la médiane de cette série statistique ? Quels sont les quartiles ?
b) Déterminer l'écart interquartile.
4) Faire un diagramme en boîte à moustaches de la série.
5) Calculer l'écart type de la série

Partie B : Augmentation des gains

L'association qui organise la loterie envisage une augmentation des gains.

6) La première hypothèse envisagée consiste à augmenter tous les gains de 217 Dinars. Dans ce cas, comment varient : a) La moyenne ? b) L'écart type ? c) La médiane ?
7) La deuxième hypothèse envisagée consiste à multiplier tous les gains par 1 puis par 2.
Dans ce cas, comment varient : a) La moyenne ? b) L'écart type ? c) La médiane ?
On donne :

$$2 \times 100 + 200 + 300 + 3 \times 400 + 2 \times 500 + 2 \times 600 + 3 \times 700 + 5 \times 800 + 1000 = 11200$$

$$2 \times 100^2 + 200^2 + 300^2 + 3 \times 400^2 + 2 \times 500^2 + 2 \times 600^2 + 3 \times 700^2 + 5 \times 800^2 + 1000^2 = 752000$$

$560^2 = 313600$ $\sqrt{62400} = 250$.

Exercice 9

On effectue un contrôle de la qualité pendant 100 heures de travail sur deux machines produisant des pièces mécaniques destinées à la fabrication de grues. Certaines pièces présentent un défaut qui les rend inutilisables. On a relevé le nombre de pièces inutilisables durant chaque heure :

Machine A	Nombre de pièces inutilisables	0	1	2	3	4	5	6	7
	Nombres d'heures	13	42	38	2	2	1	1	1

Machine B	Nombre de pièces inutilisables	0	1	2	3	4	5
	Nombres d'heures	35	40	1	1	10	13

1) **a)** Calculer le nombre moyen m_A de pièces inutilisables pendant les 100 heures étudiées pour la machine A. Calculer ensuite la variance V_A.

b) Calculer le nombre moyen m_B de pièces inutilisables pendant les 100 heures étudiées pour la machine B. Calculer ensuite la variance V_B.

2) **a)** Déterminer la médiane, puis l'écart interquartile dans le cas de la machine A. Calculer l'étendue E_A.

b) Déterminer la médiane, puis l'écart interquartile dans le cas de la machine B. Calculer l'étendue E_B.

3) **a)** Parmi la moyenne, l'écart type, la médiane, l'écart interquartile ou l'étendue, quels sont les paramètres qui mesurent la dispersion ?

b) Quel(s) paramètre(s) semble(nt) le(s) plus intéressant(s) à exploiter pour comparer ces deux machines ? Justifier.

On donne :

$$13 \times 0 + 42 \times 1 + 38 \times 2 + 2 \times 3 + 2 \times 4 + 1 \times 5 + 1 \times 6 + 1 \times 7 = 150.$$

$$35 \times 0 + 40 \times 1 + 1 \times 2 + 1 \times 3 + 10 \times 4 + 13 \times 5 = 150.$$

$$13 \times 0^2 + 42 \times 1^2 + 38 \times 2^2 + 2 \times 3^2 + 2 \times 4^2 + 1 \times 5^2 + 1 \times 6^2 + 1 \times 7^2 = 354.$$

$$35 \times 0^2 + 40 \times 1^2 + 1 \times 2^2 + 1 \times 3^2 + 10 \times 4^2 + 13 \times 5^2 = 538.$$

Exercice 10

On mesure les diamètres de troncs d'arbres d'une même espèce. On étudie 400 spécimens. On obtient les résultats suivants :

Diamètre en cm	25	26	27	28	29	30
Pourcentage	10 %	15 %	30 %	35 %	5 %	5 %

1) a) Combien de spécimens ont un diamètre supérieur ou égal à 27 cm ?

b) Parmi les spécimens qui ont un diamètre supérieur ou égal à 26 cm, quel pourcentage présente un diamètre inférieur ou égal à 27 cm ?

2) Quel est le diamètre moyen de ces troncs ?

3) Déterminer la variance, arrondie à 0,01 près, puis l'écart type, arrondi à 0,01 près, de la série statistique résumée dans le tableau ci-dessus.

4) a) Déterminer l'intervalle interquartile et calculer l'écart interquartile de la série statistique.

b) Représenter le diagramme en boîtes de la série en y faisant figurer les valeurs extrêmes et tous les quartiles.

5) Dans un autre pays, une autre étude a recensé les diamètres de 500 troncs d'arbres de la même espèce que précédemment.

Les quartiles obtenus sont : Q_1 = 25,5 ; Q_2 = 27,5 ($Q_2 = mediane$) ; Q_3= 29.

Les spécimens sont-ils plus homogènes (moins de dispersion) ou moins homogènes (plus de dispersion) que lors de la 1ère étude ? Justifier.

On donne : $25 \times 10 + 26 \times 15 + 27 \times 30 + 28 \times 35 + 29 \times 5 + 30 \times 5 = 2725$

$$25^2 \times 10 + 26^2 \times 15 + 27^2 \times 30 + 28^2 \times 35 + 29^2 \times 5 + 30^2 \times 5 = 68425$$

$$25 \times 10^2 + 26 \times 15^2 + 27 \times 30^2 + 28 \times 35^2 + 29 \times 5^2 + 30 \times 5^2 = 74405$$

$27{,}25^2 = 742{,}5625$ $\sqrt{1{,}49} \approx 1{,}22$ $1{,}49^2 \approx 2{,}22$

Correction de la série

Exercice 1

1) Pour calculer la moyenne du groupe constitué par ces trois sous groupes, il faut tenir compte des effectifs de chacun de ces sous-groupes.

La moyenne du groupe des 40+10+50=100 personnes vaut $\frac{40\times170+10\times180+50\times175}{100}$ =173,5cm 2) Si les trois sous-groupes étaient constitués du même nombre de personnes, il suffirait de considérer la moyenne arithmétique des trois valeurs $170cm$,$180cm$ et $175cm$.

En effet, si on note *x* l'effectif commun des trois sous-groupes, alors la moyenne générale vaudra $\frac{x\times170+x\times180+x\times175}{3x}=\frac{170+180+175}{3}=175\ cm$

Exercice 2

Puisque le nombre d'observations est impair (97=2x48+1), la médiane M sera égale à la 49ème mesure de température, c'est-à-dire, en observant le tableau, à 16,5°.

(la 49ème observation fait partie des 15 mesures égales à 16,5°).

Le quartile Q_1 est la plus petite valeur du caractère pour laquelle 25 % des valeurs de la série statistique lui sont inférieures ou égales.

Puisque 25% de l'effectif total représentent $97\times\frac{25}{100}$ **=24,25** ; le quartile Q_1 correspondra à la 25ème mesure, c'est-à-dire 16°.

De même, le quartile Q_3 est la plus petite valeur du caractère pour laquelle 75 % des valeurs de la série statistique lui sont inférieures ou égales.

Puisque 75% de l'effectif total représentent $97\times\frac{75}{100}$ **=72,75** ; le quartile Q_3 correspondra à la 73ème mesure, c'est-à-dire 18°.

Exercice 3

1) Pour calculer le salaire moyen de l'entreprise, il faut considérer le milieu de chaque classe ou encore on l'appelle centre de la classe.

Salaire	[800 ; 900[	[900 ; 1000[	[1000 ; 1050[	[1050 ; 1150[	[1150 ; 1300[
Effectif	42	49	74	19	16
Centre	850	950	1025	1100	1225

Le calcul de la moyenne est donc :

$$\bar{x}=\frac{\overbrace{\sum_{i=1}^{5}n_i\times x_i}^{\text{somme des produits entre les valeurs et leurs effectifs}}}{\underbrace{\sum_{i=1}^{5}n_i}_{\text{somme des effectifs}}}=\frac{n_1\times x_1+n_2\times x_2+\ldots.n_5\times x_5}{\underbrace{n_1+n_2+\ldots n_5}_{\text{effectif total}}}=\frac{42\times850+49\times950+\ldots.16\times1225}{\underbrace{42+49+\ldots.16}_{\text{effectif total}}}=\frac{198600}{200}=993$$

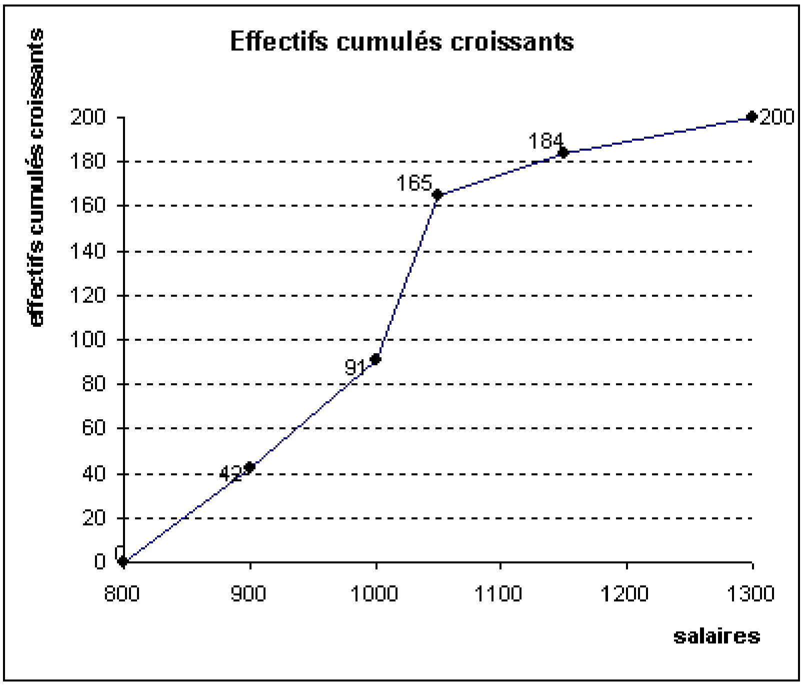
Le salaire moyen dans cette entreprise est donc de 993 Dinars.
Il n'est pas forcément très représentatif de cette entreprise, car plus de la moitié des employés y gagnent plus de 1000 Dinars.

2) Pour répondre à cette question, il faut dresser le tableau des effectifs cumulés croissants :

Salaire	[800 ; 900[	[900 ; 1000[	[1000 ; 1050[	[1050 ; 1150[	[1150 ; 1300[
Effectifs cumulés croissants	42	42+49 =91	91+74=165	165+19=184	184+16=200

Ainsi, 165 employés gagnent au plus 1050 Dinars, au sein de cette entreprise.

A partir de ce tableau, on dresse le polygone des effectifs cumulés croissants.

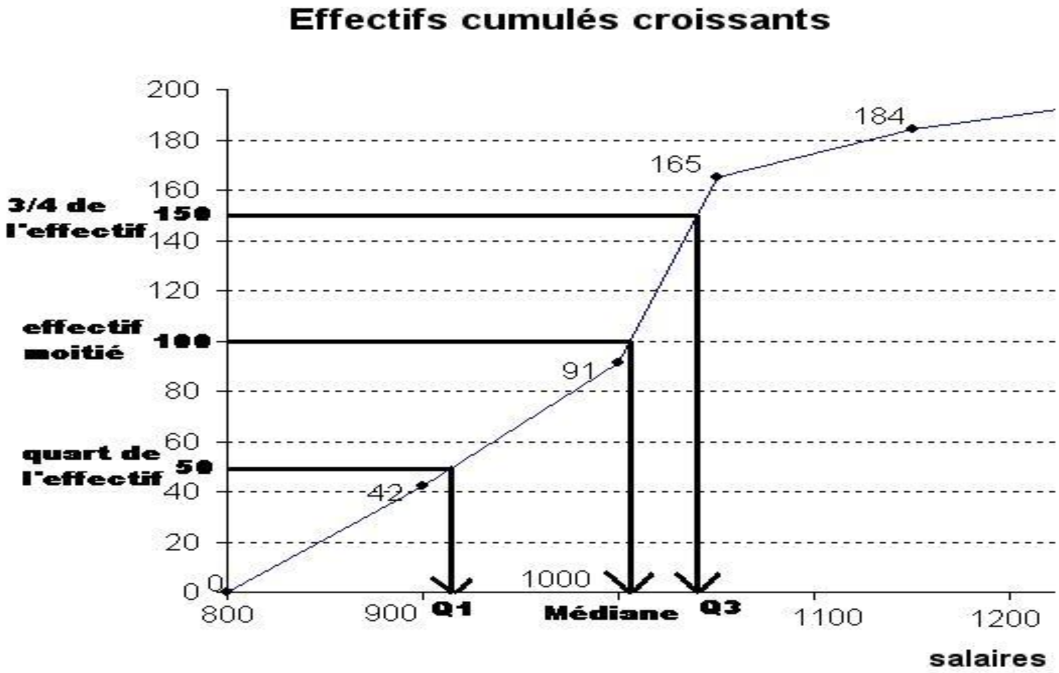


A partir de ce polygone, on cherche le salaire médian, c'est-à-dire celui qui va partager la série statistique en deux parties d'égale amplitude.

Il s'agit donc du salaire correspondant à un effectif cumulé de 100 salariés (moitié de l'effectif). On se place ainsi que l'axe des ordonnées à l'effectif cumulé 100, et on lit l'antécédent de 100. Ce sera la médiane.

On procède de même avec les quartiles Q_1 et Q_3, qui correspondent respectivement à un effectif cumulé de $\frac{1}{4} \times 200 = 50$ et de $\frac{3}{4} \times 200 = 150$.

On lit graphiquement que <u>Médiane</u> ≈ 1010 ; $Q_1 \approx 915$ et $Q_3 \approx 1050$.



3) Calcul précis de la moyenne et des quartiles Q_1 et Q_3

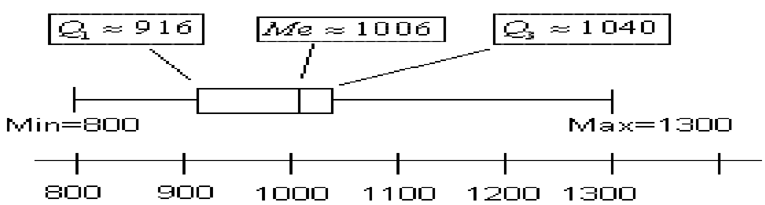
Pour calculer la médiane, on va réaliser une interpolation linéaire entre les points

A (1000 ; 91) et B (1050 ; 165)

L'équation de la droite (AB) est de la forme $y = mx + p$ avec $m = \frac{y_B - y_A}{x_B - x_A} = \frac{165-91}{1050-1000} = 1{,}48$
donc $y = 1{,}48\,x + p$.
Pour trouver la valeur de p , on utilise les coordonnées de A (ou B) : $y_A = 1{,}48\,x_A + p$
donc $p = y_A - 1{,}48\,x_A = 91 - 1{,}48 \times 1000 = -1389$.

<u>L'équation</u> de (AB) est donc $y = 1{,}48x - 1389$.
On trouve la médiane en calculant l'antécédent de la moitié de l'effectif (c'est à dire 200/2=100) par la fonction affine $f : f(x) = 1{,}48x - 1389$, c'est-à-dire en résolvant l'équation :

$1{,}48x - 1389 = 100 \ll\Leftrightarrow\gg x = \frac{1489}{1{,}48} \approx 1006{,}08$. Ainsi Médiane ≈1006.
Puisque le quartile <u>Q_3</u> semble lui aussi appartenir à l'intervalle [1000;1050[, on utilise la même droite, et on résout l'équation : $1{,}48x - 1389 = 150$ *équivaut* à $x = \frac{1539}{1{,}48} \approx 1039{,}86$.
Ainsi $Q_3 \approx 1040$. De la même manière, pour déterminer le quartile <u>Q_1</u>, on doit déterminer l'équation de la droite reliant les points (900 ; 42) et (1000 ; 91).
Cette droite a pour équation $y = 0{,}49x - 399$.
Et la résolution de l'équation $0{,}49x - 399 = 50$ *est équivalente* à $x = \frac{449}{0{,}49} \approx 916{,}33$
ce qui donne : $Q_1 \approx 916$
4) Le diagramme en boîte de la série est donné par :



Exercice 4

Pour le premier diagramme, Max =100, min = 5, donc l'étendue =100−5 = 95, Q_1 = 30,

Médiane = 45, Q_3 = 65. L'écart interquartile vaut donc $Q_3 - Q_1$ = 65 – 30 = 35.
Pour le deuxième diagramme, Max = 80, min = 10, donc l'étendue = 80 – 10 = 70, Q_1 = 35,
Médiane = 45, Q_3 = 55. L'écart interquartile vaut donc $Q_3 - Q_1$ = 55 – 35 = 20.
Pour le troisième diagramme, Max = 95, min = 20, donc étendue = 95 – 20 = 75, Q_1 = 35,
Médiane = 45, Q_3 = 65. L'écart interquartile vaut donc $Q_3 - Q_1$ = 65 - 35 = 30.

Exercice 5

Comparaison de températures :

1) **a)Ville de Serbie :**
L'étendue des températures de la ville de Serbie vaut : Max − min = 31−(−5) = 36°.
La moyenne des températures de la ville de Serbie est égale à :

$$\overline{x}_S = \frac{-5-4+4+15+27+31+30+26+20+10-5}{12} = 15°.$$

La variance des températures vaut :

1ère Méthode de calcul :

$$V_1 = \frac{(-5-15)^2+(-4-15)^2+(4-15)^2+(15-15)^2+(27-15)^2+(31-15)^2+(31-15)^2+\cdots+(-5-15)^2}{12} = \frac{2334}{12}$$

$V_1 = 194{,}5$.

2ème Méthode de calcul (***méthode pratique de calcul de la variance***)

$$V_1 = \frac{1\times(-5)^2+1\times(-4)^2+1\times 4^2+1\times 15^2+1\times 27^2+1\times 31^2+1\times 31^2+\cdots 1\times(-5)^2}{12} - 15^2 = 194{,}5$$

L'écart-type des températures vaut donc $\sigma = \sqrt{V_1} = \sqrt{194{,}5} \approx 13{,}95$.

b) Ville de Kosovo :
L'étendue des températures de la ville de Kosovo vaut : Max − min = 19 − (3) = 16°.
La moyenne des températures de la ville de Kosovo est égale à :

$$\overline{x}_K = \frac{3+4+7+10+14+17+19+18+16+17+7+6}{12} = 11{,}5°.$$

La variance des températures vaut :

$$V_2 = \frac{(3-11{,}5)^2+(4-11{,}5)^2+(7-11{,}5)^2+(10-11{,}5)^2+(14-11{,}5)^2+(17-11{,}5)^2+\cdots+(6-11{,}5)^2}{12} = \frac{387}{12} = 32{,}25$$

$$V_2 = \frac{1\times 3^2+1\times 4^2+1\times 7^2+1\times 10^2+1\times 14^2+1\times 17^2+1\times 19^2+\cdots 1\times 6^2}{12} - 11{,}5^2 = 32{,}25$$

L’écart-type des températures vaut donc : $\sigma = \sqrt{V_2} = \sqrt{32,25} \approx 5,68$
2) Les calculs précédents permettent d’établir quelques remarques :
- En moyenne il fait plus chaud à Serbie qu’à Kosovo.
- L’étendue des températures est plus forte à Serbie qu’à Kosovo.
- Le climat est plus « modéré » à Serbie qu’à Kosovo car les températures sont moins « étirées » autour de la moyenne.

Exercice 6

1) Afin de calculer la moyenne $\bar{x}$ et l'écart type σ de la série, il faut réorganiser cette série en effectifs :

Note	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
Effectif	0	0	0	2	2	5	10	9	10	12	10	7	5	5	1	1	0	1	0	0	80

On calcule alors : $\bar{x} = \frac{2\times4+2\times5+\cdots1\times16+1\times18}{80} = \frac{778}{80} \approx 9,725$ Puis la variance :

$$\overline{V} = \frac{2\times(4-9,725)^2+2\times(5-9,725)^2+\ldots..+1\times(18-9,725)^2}{80} = \frac{619,95}{80} = 7,749375$$

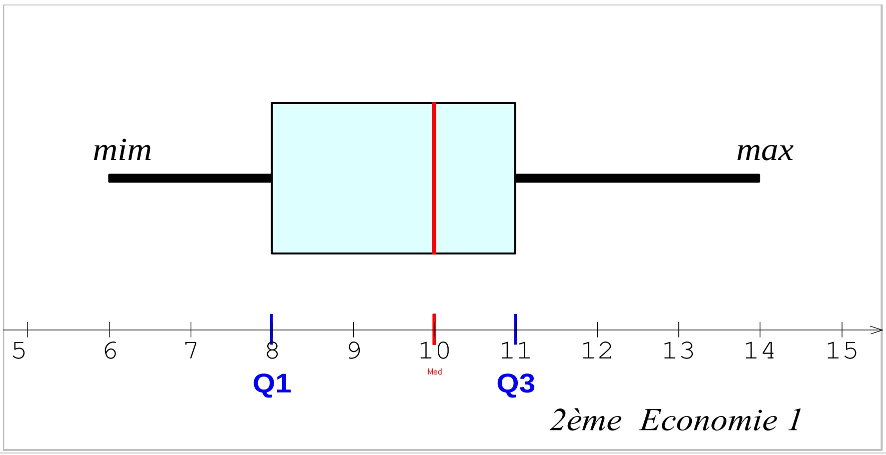
Donc l’écart-type $\sigma = \sqrt{7,749375} \approx 2,78$
2) L’intervalle : $[\bar{x}-\sigma\ ;\ \bar{x}+\sigma]$ = [9,725−2,78 ; 9,725+2,78] = [6,945 ; 12,505] contient 58 notes, soit un pourcentage égal à : $\frac{58}{80}\times 100 = 72,5\%$.
Environ 27,5 % des notes sont donc en dehors de cet intervalle.
L’intervalle $[\bar{x}-2\sigma\ ;\ \bar{x}+2\sigma]$ = [9,725−2 x 2,78 ; 9,725+2 x 2,78] = [4,165 ; 5,285] cet intervalle contient 76 notes, soit un pourcentage égal à $\frac{76}{80}\times 100 = 95\%$. Environ 5 % des notes sont en dehors de cet intervalle. L’échantillon de notes est "normal".

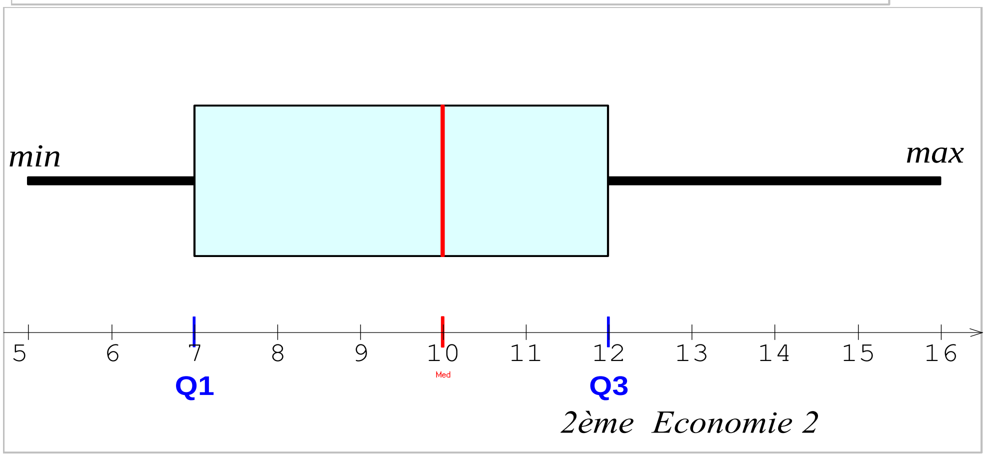
Exercice 7

Pour la 2ème Economie et services 1 :

L'effectif total est : $3+4+4+\cdots+1 = 33$
$\frac{33+1}{2} = 17$ donc la médiane est le 17ème terme de la série : Méd = 10
$\frac{33}{2} = 8,25$ donc le 1er quartile est le 9ème terme de la série : Q1 = 8
$\frac{33\times3}{4}$ = 24,75 donc le 3ème quartile est le 25ème terme de la série : Q3 = 11
Pour la 2ème Economie et services 2 :

L'effectif total est 2+4+3+…+2 = 32
$\frac{32+1}{2}$ = 16,5 donc la médiane est la moyenne des 16ème et 17ème terme de la série :
Méd = 10 +102 = 10
$\frac{32}{4}$ = 8 donc le 1er quartile est le 8ème terme de la série : Q1 = 7
$\frac{32\times3}{4}$ = 24 donc le 3ème quartile est le 24ème terme de la série : Q3 = 12
Diagrammes en boîtes :





Bilan : Les graphiques ci-dessus mettent bien en évidence que l'écart interquartile et l'étendue sont plus resserrés en 2ème Economie 1 qu'en 2ème Economie 2 donc les élèves de 2ème Economie 1 ont globalement un niveau plus homogène que ceux de 2ème Economie 2.