

Exercices supplémentaires – Statistiques

Partie A : Médiane et quartiles

Exercice 1

Le tableau ci-dessous indique la répartition des élèves d'un lycée suivant leur âge.

Age	14	15	16	17	18	19	20
Effectif	130	204	271	316	198	77	14

- 1) Compléter le tableau avec les effectifs cumulés croissants.
- 2) Déterminer la médiane et les quartiles de cette série.
- 3) Exprimer par une phrase les informations données par ces indicateurs.

Exercice 2

Dans un lycée, on étudie les moyennes trimestrielles du premier trimestre des classes Jaune et Rouge.

- 1) Les 25 élèves de la classe Jaune ont obtenus les moyennes suivantes : 3 ; 4 ; 5 ; 7 ; 7 ; 10 ; 10 ; 10 ; 10 ; 11 ; 11 ; 12 ; 12 ; 12 ; 12 ; 13 ; 13 ; 13 ; 14 ; 15 ; 15 ; 16 ; 18.
 - a. Déterminer la médiane et les quartiles de cette série.
 - b. Représenter le diagramme en boîte correspondant à cette série.
- 2) Dans la classe Rouge, les indicateurs des moyennes sont : $\text{Min} = 3$; $Q_1 = 8$; $\text{Med} = 10$; $Q_3 = 12$; $\text{Max} = 17$.
 - a. Représenter le diagramme en boîte correspondant à cette série.
 - b. Parmi les affirmations suivantes, lesquelles sont vraies, fausses ou indécidables ? Justifier.
 - 50% des élèves de la classe Rouge ont une note comprise entre 10 et 12.
 - 75% des élèves de la classe Rouge ont une note inférieure ou égale à 12.
 - Au moins 50% des élèves de la classe Rouge ont une note inférieure ou égale à la note médiane de la classe Jaune.

Exercice 3

Une machine A déverse du caoutchouc de façon continue dans un moule pour fabriquer des joints d'étanchéité pour l'industrie automobile. Pour contrôler la régularité de cet écoulement, on a effectué 40 mesures de la masse de caoutchouc écoulée sur des durées de 10 secondes.

69,7	63,6	64,4	62,4	63,4	59,7	60,7	65	67	65,6
68,8	60,3	63,4	67,6	64,1	72,9	64,5	66,2	65,3	66,4
65,9	55,8	67,1	65,5	64,5	62,2	71	64,4	69,8	66,1
68,7	61,2	63,1	64,6	58,7	62,3	61,2	62,1	61,4	64,8

- 1) Déterminer la médiane et les quartiles de cette série.
- 2) Si on considère comme aberrantes les mesures qui n'appartiennent pas à l'intervalle $[Q_1 - 1,5I; Q_3 + 1,5I]$ où I désigne l'écart interquartile, quel est le pourcentage de valeurs aberrantes ?
- 3) Deux autres machines B et C, du même type, ont aussi subi un contrôle mais portant sur 1200 mesures. On a obtenu :

	Min	Q_1	Med	Q_3	Max
B	64,4	65,4	65,8	66,2	72
C	65,2	65,8	65,9	66,1	66,3

Construire les diagrammes en boîtes relatifs aux trois machines.

- 4) En quoi le diagramme de la machine A est-il très différent des deux autres ? Comment l'expliquer ?
- 5) Entre les machines B et C, laquelle paraît la plus régulière ? Argumenter.

Partie B : Moyenne et écart-type

Exercice 1

Dans un tournoi de scrabble, le nombre de points totalisé par les huit participants a été : 298 ; 407 ; 336 ; 425 ; 512 ; 321 ; 543 ; 396.

Calculer le score moyen et l'écart type de ces scores.

Exercice 2

Dans un jury, un professeur a corrigé 100 copies avec une moyenne de 10,8 et une médiane de 10 ; un autre professeur a corrigé 80 copies avec une moyenne de 9,8 et une médiane de 10,2. Quel(s) indicateur(s) peut-on calculer sur l'ensemble des 180 copies ? Le(s) calculer.

Exercice 3

La pesée automatique d'un lot de 20 barquettes d'un produit alimentaire a donné les résultats suivants (arrondis au gramme) :

300 ; 311 ; 315 ; 308 ; 311 ; 317 ; 308 ; 309 ; 311 ; 312 ; 309 ; 318 ; 307 ; 308 ; 303 ; 310 ; 314 ; 313 ; 310 ; 319

Un lot est accepté si les trois conditions sont remplies :

- le poids moyen m d'une barquette est de 310g à 1g près.
- l'écart-type s des poids est inférieur à 5g
- 80% au moins des poids sont dans l'intervalle $[m - s; m + s]$.

Qu'en est-il pour ce lot ?

Exercice 4

Une machine fabrique des pièces cylindriques de diamètre théorique 25 mm. Le service qualité a prélevé un échantillon de 100 pièces au hasard dans la fabrication. Voici les mesures obtenues :

Diamètre	[24,2;24,4[[24,4 ;24,6[[24,6;24,8[[24,8;25[[25;25,2[
Effectif	5	13	24	19	14

Diamètre	[25,2;25,4[[25,4;25,6[[25,6;25,8[[25,8 ;26[
Effectif	10	8	5	2

- 1) Calculer la moyenne \bar{x} et l'écart-type s de la série de mesure.
 - 2) La production de la machine est jugée bonne si la série des mesures vérifie les trois conditions :
- $24,9 < \bar{x} < 25,1$
- $s < 0,4$
- 90% au moins de l'effectif figure dans $[\bar{x} - 2s; \bar{x} + 2s]$.

La production est-elle bonne ?

Partie C : Bilan

Exercice 1

Dans tout l'exercice les tailles sont exprimées en centimètre.

- 1) L'équipe de soins de la maternité « Beaux jours » a relevé la taille des nouveau-nés. Pendant la troisième semaine du mois de janvier 2003, il y a eu 9 naissances. Les tailles sont données dans le tableau ci-dessous :

48	50,5	51,5	50	52,5	50	49	53	50
----	------	------	----	------	----	----	----	----

- a) Calculer la moyenne des tailles de ces 9 nouveau-nés.
- b) Déterminer la médiane des tailles de ces 9 nouveau-nés.
- 2) Sur la totalité du mois de janvier 2003, il y a eu 57 naissances à la maternité « Beaux jours ». Les 57 tailles sont données dans le tableau ci-dessous :

Taille	46	47,5	48	48,5	49	49,5	50	50,5	51	51,5	52	52,5	53
Effectif	1	2	3	5	5	7	9	8	7	5	2	2	1

- 1) Calculer les fréquences puis les fréquences cumulées croissantes.
 - 2) Calculer la moyenne des tailles de ces 57 nouveau-nés.
 - 3) Déterminer la médiane des tailles de ces 57 nouveau-nés en précisant la démarche.
 - 4) Calculer le pourcentage de nouveau-nés ayant une taille inférieure ou égale à 49 cm. Donner la réponse arrondie à 0,1 %.
 - 5) Déterminer le premier et le troisième quartile de cette série.
- 3) L'étude statistique de la taille, en centimètre, des 64 nouveau-nés durant le même mois de janvier 2003 à la maternité « Bon accueil » a donné les résultats suivants :

Minimum	Maximum	Moyenne	Médiane	1 ^{er} quartile	3 ^{ème} quartile
46	53	49,3	49	48	50,5

- a) Parmi les deux maternités « Beaux jours » et « Bon accueil », une seule possède un service pour les naissances prématurées. Peut-on trouver laquelle ? Justifier votre réponse.
- b) Les deux maternités « Beaux jours » et « Bon accueil » sont les seules maternités de la même ville. Calculer la moyenne des tailles des nouveau-nés en janvier 2003 dans les maternités de la ville.
- c) Les données de l'énoncé permettent-elles de déterminer la médiane des tailles de ces nouveau-nés ? Si oui, la déterminer ; sinon, expliquer pourquoi.

Exercice 2

Au cours d'une fabrication de fromages de chèvres, on a relevé les masses suivantes des fromages :

masse en grammes	[80;85[[85;90[[90;95[[95;100[[100;105[[105;110[[110;115[
effectifs	5	9	14	18	25	16	7

- 1) Calculer le poids moyen d'un fromage de chèvre pour cette production.
- 2) Calculer les fréquences et les fréquences cumulées.
- 3) Tracer la courbe des fréquences cumulées croissantes et déterminer graphiquement la médiane et le 1^{er} et le 3^{ème} quartile.
- 4) Déterminer le pourcentage de fromages de chèvres ayant un poids compris entre 92 et 107 g.

Exercice 3

Après étude, les autorités d'une île isolée ont décidé d'installer une éolienne pour répondre aux besoins énergétiques de leur communauté. L'éolienne choisie fonctionne lorsque le vent atteint au moins 8 nœuds et il faut l'arrêter lorsque le vent atteint ou dépasse 48 nœuds.

Les autorités décident de mesurer pendant un mois la vitesse du vent, à l'aide d'un anémomètre, sur le site *M* au sommet d'une montagne. Une mesure est effectuée chaque jour. Voici les résultats obtenus (le mois comporte 30 jours et on peut y lire que la vitesse de 22 nœuds a été mesurée 5 jours) :

Vitesse du vent (en nœuds)	7	14	16	18	20	22	24	26	27	30	44	50
Effectif (en jours)	1	2	1	1	4	5	3	4	4	2	1	2

- 1) Calculer les effectifs cumulés croissants.
- 2) Calculer le pourcentage des jours du mois où l'éolienne ne produirait pas d'électricité.
- 3) Déterminer la moyenne et l'étendue de cette série statistique.

- 4) Déterminer la médiane et les quartiles de cette série statistique en expliquant la démarche. Calculer l'écart interquartile.
- 5) Le même mois, les autorités ont fait mesurer les vitesses des vents sur un autre site F au bord d'une falaise. Les résultats obtenus donnent une médiane de 22 nœuds, un premier quartile de 14 nœuds et un troisième quartile de 34 nœuds, un minimum de 7 nœuds et un maximum de 50 nœuds.
Le rendement maximal de l'éolienne se situe à environ 23 nœuds. Quel site vous paraît le plus intéressant pour l'installation de l'éolienne ? Justifier la réponse.

Exercice 4

Un trufficulteur décide de tester l'influence de l'arrosage de ses truffières sur la masse des truffes récoltées. Il décide donc de répartir ses récoltes en deux lots de 100 truffes :

- le premier, appelé lot A , provient de truffières ne recevant aucun arrosage ;
- le second, appelé lot B , provient de truffières arrosées.

- 1) Au moment de la récolte, il pèse ses truffes et obtient, pour le lot B , les résultats suivants :

masse (en g)	15	15,5	16	16,5	17	17,5	18	18,5	19	19,5	20	20,5	21	21,5	22
Nombre de truffes	16	4	20	14	22	4	8	3	2	1	2	0	1	0	3

Déterminer le minimum, le premier quartile, la médiane, le troisième quartile, le maximum, la moyenne et l'étendue de cette série statistique.

- 2) Pour le lot A , on obtient :

Les phrases suivantes sont elles vraies ou fausses.

	minimum	Q_1	médiane	Q_3	maximum	Moyenne
	12	14	15,5	16,5	20	15,75

Justifier :

- a) « environ la moitié du lot B est constitué de truffes d'un poids égal ou supérieur aux trois-quarts des truffes du lot A . »
- b) « en arrosant, on réduit l'écart interquartile des masses entre les truffes récoltées. »

Correction exercices supplémentaires – Statistiques

Partie A : Médiane et quartiles

Exercice 1

Age	14	15	16	17	18	19	20
Effectif	130	204	271	316	198	77	14
Eff. Cum. croissant	130	234	504	820	1018	1095	1109

1) Voir ci-dessus.

2) L'effectif total est 1109 qui est impair donc la médiane est la 555^{ème} valeur, soit 17.

Pour $Q_1 : \frac{1109}{4} = 277,25$ donc le 1^{er} quartile est la 278^{ème} valeur, soit 16.

Pour $Q_3 : 3 \times \frac{1109}{4} = 831,75$ donc le 3^{ème} quartile est la 832^{ème} valeur, soit 18.

3) La moitié des élèves ont 17 ans ou moins. Au moins un quart des élèves a 16 ans et au moins 75% des élèves ont 18 ans.

Exercice 2

1)

a. Il y a 25 notes donc la médiane est la 13^{ème} valeur, soit 12.

Pour $Q_1 : \frac{25}{4} = 6,25$ donc le 1^{er} quartile est la 7^{ème} valeur, soit 10.

Pour $Q_3 : 3 \times \frac{25}{4} = 18,75$ donc le 3^{ème} quartile est la 19^{ème} valeur, soit 13.

b. Voir ci-dessous

2)

a.

b. La 1^{ère} proposition est indécidable : en effet, il y a au moins 25% des valeurs entre 10 (qui est la médiane) et 12 (qui est le 3^{ème} quartile). Mais nous ne savons pas combien il y a de 10 et de 12 dans la classe. Par exemple, avec 8 notes, on peut avoir :

3 ; 8 ; 10 ; 10 ; 12 ; 13 ; 17 qui vérifie toutes les conditions...

La 2^{ème} proposition est vraie : c'est la définition du 3^{ème} quartile.

La 3^{ème} proposition est vraie car la médiane de la classe Jaune est égale à 12 et comme cela correspond au 3^{ème} quartile de la classe Rouge, cela signifie qu'au moins 75% des élèves de la classe Rouge ont une note inférieure ou égale à 12.

Exercice 3

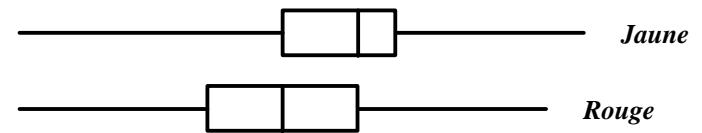
1) On commence par classer les valeurs dans l'ordre croissant. On obtient (rangement par colonne):

55,8	60,7	62,1	63,1	64,1	64,5	65,3	66,1	67,1	69,7
58,7	61,2	62,2	63,4	64,4	64,6	65,5	66,2	67,6	69,8
59,7	61,2	62,3	63,4	64,4	64,8	65,6	66,4	68,7	71
60,3	61,4	62,4	63,6	64,5	65	65,9	67	68,8	72,9

Il y a 40 valeurs donc la médiane est entre la 20^{ème} et la 21^{ème} valeur qui sont égales toutes les deux à 64,5 donc la médiane est $\boxed{64,5}$.

Pour le 1^{er} quartile, c'est la 10^{ème} valeur, soit $\boxed{Q_1 = 62,2}$

Le 3^{ème} quartile est la 30^{ème} valeur, soit $\boxed{Q_3 = 66,2}$



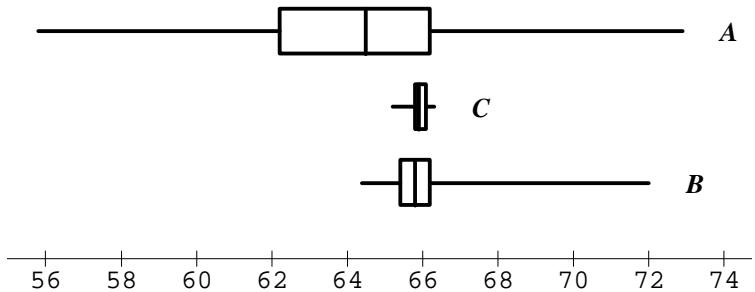
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

2) L'écart interquartile est égal à 4 donc $Q_1 - 0,5I = 62,2 - 1,5 \times 4 = 56,2$ et $Q_3 + 1,5I = 66,2 + 1,5 \times 4 = 72,2$. Nous allons donc exclure les valeurs qui n'appartiennent pas à $[56,2; 72,2]$. Il ya donc 2 valeurs à exclure. Cela représente 5% des valeurs car $2 \times \frac{100}{40} = 5$.

3)

4) Le diagramme de la machine A est beaucoup plus étendu que ceux des machines B et C. Cela vient du fait qu'il n'y a que 40 valeurs pour la machine A contre 1200 pour B et C. Plus il y a de valeurs et plus les résultats sont fiables. Il y a proportionnellement moins de valeurs aberrantes.

5) Il semble que la machine C soit plus régulière que la machine B. En effet, l'écart interquartile pour la machine C, qui est de 0,3, est plus faible que celui de la machine B, qui est de 0,8. De même pour l'étendue de la machine C (1,1) qui est inférieure à celle de la machine B (7,6).



Partie B : Moyenne et écart-type

Exercice 1

$$\bar{x} = \frac{298 + 407 + \dots + 396}{8} = [404,75]$$

$$V = \frac{(298 - 404,75)^2 + (407 - 404,75)^2 + \dots + (396 - 404,75)^2}{8} = 6780,4375$$

$$\sigma = \sqrt{6780,4375} \approx [82,34]$$

La moyenne est donc de 404,75 points pour un écart-type de 82,34 points environ.

Exercice 2

Pour la moyenne : on peut estimer que les 100 premières copies ont toutes des notes égales à 10,8 et que les 80 copies suivantes ont toutes une note égale à 9,8. Nous avons donc une moyenne globale de 10,36 environ car $\frac{100 \times 10,8 + 80 \times 9,8}{180} \approx 10,36$.

Pour la médiane, nous ne pouvons pas déterminer la médiane des 180 copies car on ne sait pas comment sont répartis précisément les deux paquets.

Exercice 3

Calcul du poids moyen :

$$m = \frac{300+311+\dots+319}{20} = 310,65$$

Le poids moyen est égal à 310,65g donc la 1^{ère} condition est vérifiée.

Calcul de l'écart-type :

$$V = \frac{300^2 + 311^2 + \dots + 319^2}{20} - 310,65^2 = \frac{1930483}{20} - 96503,4225 = 20,7275$$

$$\sigma = \sqrt{V} \approx 4,55$$

L'écart-type est égal à 4,55g environ donc la 2^{ème} condition est vérifiée.

Pour la 3^{ème} condition, on détermine l'intervalle $[m - \sigma; m + \sigma]$: $m - \sigma \approx 306,1$ et $m + \sigma \approx 315,2$.

Il y a 15 valeurs qui appartiennent à $[306,1; 315,2]$ ce qui représente 75% des valeurs.

La 3^{ème} condition n'est donc pas vérifiée.

Le lot est donc rejeté.

Exercice 4

$$1) \bar{x} = \frac{24,3 \times 5 + 24,5 \times 13 + \dots + 25,9 \times 2}{100} = [24,946]$$

$$V = \frac{24,3^2 \times 5 + 24,5^2 \times 13 + \dots + 25,9^2 \times 2}{100} - 24,946^2 = 0,148684$$

$$s = \sqrt{V} \approx 0,386$$

La moyenne de cette série est égale à 24,946 et l'écart-type est de 0,386 environ.

2) La 1^{ère} condition est bien vérifiée, tout comme la seconde. Pour la troisième, on commence par déterminer l'intervalle indiqué. $\bar{x} - 2s \approx 24,174$ et $\bar{x} + 2s \approx 25,718$. Or sept valeurs (ce qui représente 7%) n'appartiennent pas à l'intervalle [24,2; 25,6[. Il y a donc au moins 93% des valeurs dans l'intervalle indiqué. La 3^{ème} condition est donc bien vérifiée et la production de la machine est bonne.

Partie C : Bilan

Exercice 1

1)

a. $\frac{48+50,5+\dots+50}{9} = 50,5$ donc la moyenne des neuf tailles est 50,5 cm.

b. Il y a 9 valeurs donc la médiane est la 5^{ème} valeur, après les avoir rangés dans l'ordre croissant, autrement dit, la médiane est 50 cm.

2)

Taille	46	47,5	48	48,5	49	49,5	50	50,5	51	51,5	52	52,5	53
Effectif	1	2	3	5	5	7	9	8	7	5	2	2	1
Fréq	0,018	0,035	0,053	0,088	0,088	0,123	0,158	0,14	0,123	0,088	0,035	0,035	0,018
Fréq cum	0,018	0,053	0,106	0,194	0,282	0,405	0,563	0,703	0,826	0,914	0,949	0,984	1

3) $\bar{x}_2 = \frac{46 \times 1 + 47,5 \times 2 + \dots + 53 \times 1}{57} \approx 49,99$

La moyenne des nouveau-nés est de 49,99 cm dans la maternité « Beaux jours ».

4) D'après les fréquences cumulées croissantes, on passe les 50% pour une taille de 50 cm donc la médiane est de 50 cm.

5) Il y a 16 bébés dont la taille est inférieure ou égale à 49 cm. Cela représente 28,1% environ car $16 \times \frac{100}{57} \approx 28,07$

6) D'après les fréquences cumulées croissantes, on dépasse les 25% des valeurs pour une taille de 49 cm et 75% des valeurs pour une taille de 51 cm donc $Q_1 = 49$ et $Q_3 = 51$.

7)

a. La moyenne, tout comme la médiane et les quartiles, de la maternité « Bon accueil » sont inférieurs aux valeurs de la maternité « Beaux jours ». Il est donc probable que le service de prématurés soit dans la maternité « Bon accueil ».

b. $\frac{49,99 \times 57 + 49,3 \times 64}{57 + 64} \approx 49,63$ donc la moyenne des tailles des naissances sur la ville est de 49,63 cm.

c. Comme toutes les données de la maternité « Bon accueil » ne sont pas connues, on ne peut pas déterminer la médiane globale de la ville.

Exercice 2

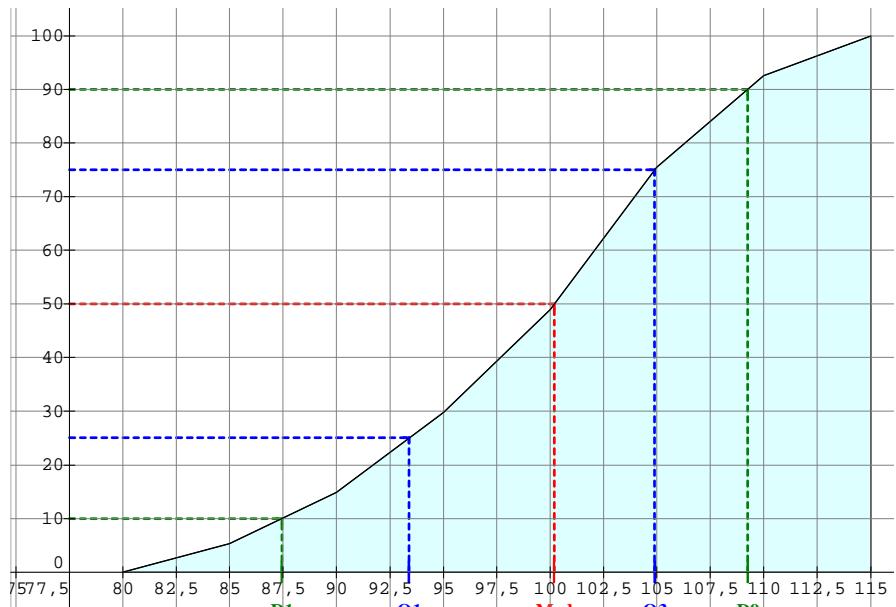
1) $\bar{x} = \frac{82,5 \times 5 + 87,5 \times 9 + \dots + 112,5 \times 7}{5 + 9 + \dots + 7} \approx 99,15$ donc le poids moyen des fromages de chèvres est 99,15 g

2)

masse en grammes	[80;85[[85;90[[90;95[[95;100[[100;105[[105;110[[110;115[
effectifs	5	9	14	18	25	16	7
fréquences	0,053	0,096	0,149	0,191	0,266	0,170	0,074
Fréquences cumulées	0,053	0,149	0,298	0,489	0,755	0,925	1

3) La médiane est environ 100g, le 1^{er} quartile est environ 93g et le 3^{ème} quartile est environ 105g.

4) D'après le diagramme, environ 22% des fromages de chèvres ont un poids inférieur à 92g et environ 84% des fromages ont un poids inférieur à 107g. Il y a donc environ 62% des fromages de chèvres dont le poids est compris entre 92 et 107g.



Exercice 3

1)

Vitesse du vent (en nœuds)	7	14	16	18	20	22	24	26	27	30	44	50
Effectif (en jours)	1	2	1	1	4	5	3	4	4	2	1	2
Effectifs cumulés	1	3	4	5	9	14	17	21	25	27	28	30

2) Il y a trois jours dans le mois où l'éolienne ne fournit pas d'électricité (un où le vent est insuffisant et deux où le vent est trop fort). Cela représente donc 10% de jours car $3 \times \frac{100}{30} = 10$.

$$3) \bar{x} = \frac{7 \times 1 + 14 \times 2 + \dots + 50 \times 2}{30} = 24,9 \quad e = 50 - 7 = 43$$

La vitesse moyenne est donc de 24,9 nœuds et l'étendue est égale à 43 nœuds.

4) Pour la médiane : il y a un effectif total de 30 donc la médiane est comprise entre la 15^{ème} et la 16^{ème} valeur, soit 24.

Pour le 1^{er} quartile : $\frac{30}{4} = 7,5$ donc le 1^{er} quartile est la 8^{ème} valeur et $Q_1 = 20$

Pour le 3^{ème} quartile : $\frac{30}{4} \times 3 = 22,5$ donc le 3^{ème} quartile est la 23^{ème} valeur et $Q_3 = 27$

L'écart interquartile est alors de 7 nœuds.

5) Le site le plus adapté semble être le site M. En effet, les médianes pour les deux sites sont proches de la valeur souhaitée, à savoir 23 nœuds, les valeurs extrêmes sont les mêmes. Cependant, l'écart interquartile pour le site M est de 7 nœuds contre 20 nœuds pour le site F. Concrètement, les vitesses des vents pour le site M sont plus régulières et donc on est plus souvent proche du rendement maximal.

Exercice 4

1) Commençons par calculer les effectifs cumulés croissants :

Masse (en g)	15	15,5	16	16,5	17	17,5	18	18,5	19	19,5	20	20,5	21	21,5	22
Nombre de truffes	16	4	20	14	22	4	8	3	2	1	2	0	1	0	3
Eff. Cum.	16	20	40	54	76	80	88	91	93	94	96	96	97	97	100

L'effectif total est 100 donc la médiane est comprise entre la 50^{ème} et la 51^{ème} valeur, soit $Med = 16,5g$

$\frac{100}{4} = 25$ donc le 1^{er} quartile est la 25^{ème} valeur, soit $Q_1 = 16g$.

$100 \times \frac{3}{4} = 75$ donc le 3^{ème} quartile est la 75^{ème} valeur, soit $Q_3 = 17g$

Le minimum est bien sûr 15g et le maximum 22g.

L'étendue est alors de 7g.

$\bar{x} = \frac{15 \times 16 + 15,5 \times 4 + \dots + 22 \times 3}{100} = 16,81$ donc la moyenne est de 16,81g.

2)

- a. **VRAI** : en effet, la médiane du lot B, à savoir 16,5g correspond au 3^{ème} quartile du lot A.
- b. **VRAI** : en effet, pour le lot A (sans arrosage), l'écart interquartile est égal à 2,5g alors que celui du lot B (avec arrosage) est de 1g.