Corrigé de la feuille de TD n° 14

Statistique descriptive

Exercice 1

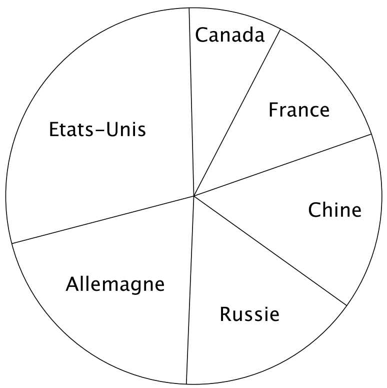
1.

	volume moyen	écart type
2006	8018,3	3125,2
2007	8395,2	3517,5
2008	8385,5	3975,6
2009	8977,8	3413,6
2010	9488,5	3722,5

2. En 2010, le volume total d'importations est égal à 56931 tonnes. On obtient par proportionnalité les angles des secteurs angulaires du diagramme circulaire correspondant :

	Etats-Unis	Allemagne	Russie	Chine	France	Canada
2010	103,2°	73,1°	56,7°	54,8°	43,3°	28,9°

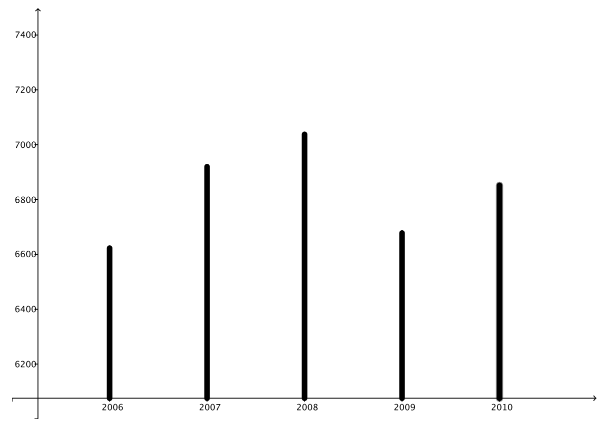
ce qui donne graphiquement :



3.

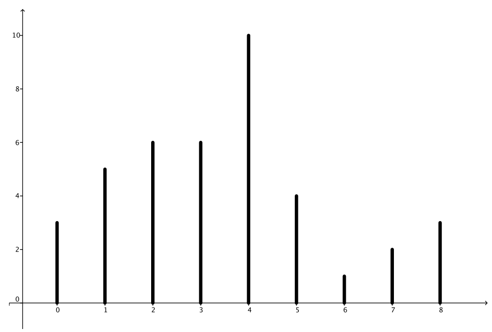
	Etats-Unis	Allemagne	Russie	Chine	France	Canada
volume moyen	15041,6	11079,6	7906,6	6735,0	6822,2	4333,4
écart type	793,6	635,8	717,6	1793,3	153,3	127,7

4.

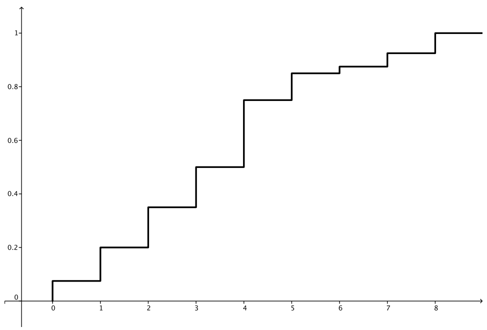


Exercice 2

1.



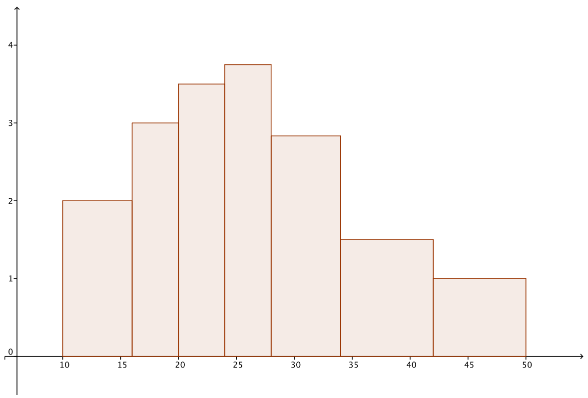
2.



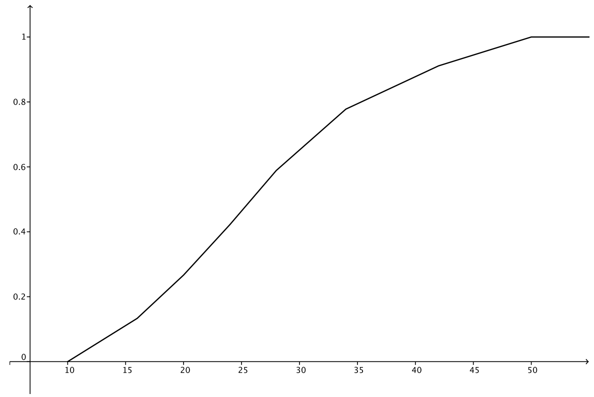
3. On a :
   - la mode : 4
   - la moyenne : $3,475$
   - la médiane : $3,5$
   - l'étendue : 8
   - l'écart type : $\approx 2,156$

Exercice 3

1.



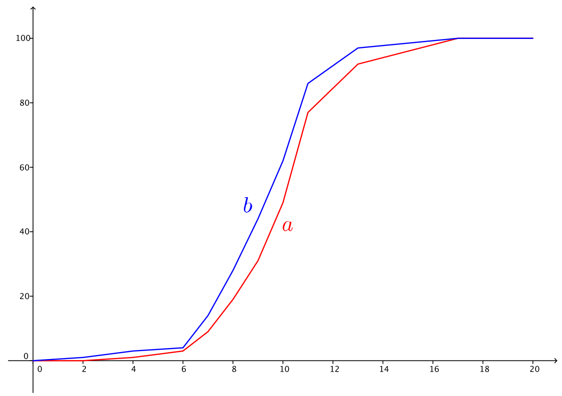
2.



3. On a :
   - la classe modale : $[28,34[$
   - la moyenne : $26,9$
   - la médiane : $\frac{1}{2}\left(\left(24+\frac{28-24}{53-38}(45-38)\right)+\left(24+\frac{28-24}{53-38}(46-38)\right)\right)=26$
   - l'étendue : 40
   - l'écart type : $\approx 9,6$

Exercice 4

1.



2. On a :

$$\overline{a}=9,94 \quad \text{et} \quad s_a \approx 2,31.$$

3. On a :

$$\overline{b}=9,18 \quad \text{et} \quad s_b \approx 2,26.$$

4. (a) On a $c = mb + p$ donc $\overline{c} = \overline{mb+p} = m\overline{b} + p$ par linéarité de la moyenne et d'après la formule de Koenig-Huygens :

$$\begin{aligned} s_c^2 &= \overline{c^2} - \overline{c}^2 = \overline{(mb+p)^2} - \left(\overline{mb+p}\right)^2 = \overline{m^2b^2 + 2mpb + p^2} - \left(m\overline{b} + p\right)^2 \\ &= m^2\overline{b^2} + 2mp\overline{b} + p^2 - m^2\overline{b}^2 - 2mp\overline{b} - p^2 = m^2\left(\overline{b^2} - \overline{b}^2\right) = m^2 s_b^2 \end{aligned}$$

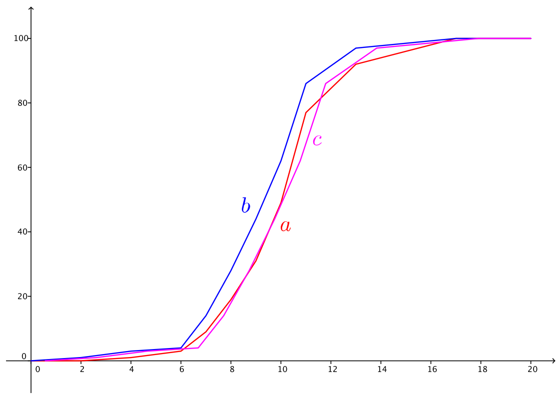
donc $s_c = \sqrt{s_c^2} = \sqrt{m^2 s_b^2} = m s_b$ car $m \geqslant 0$. Puisque $\overline{c} = \overline{a}$ et $s_c = s_a$, on en déduit que :

$$m = \frac{s_a}{s_b} \approx 1,02 \quad \text{et} \quad p = \overline{a} - m\overline{b} \approx 0,57.$$

(b) L'ajustement $c = mb + p$ donne les nouvelles classes de notes suivantes :

classes de notes	$[0,5;2,6[$	$[2,6;4,6[$	$[4,6;6,7[$	$[6,7;7,7[$	$[7,7;8,7[$
nb de copies	0	1	2	6	10

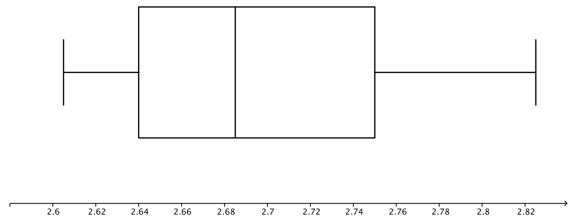
classes de notes	$[8,7;9,8[$	$[9,8;10,8[$	$[10,8;11,8[$	$[11,8;13,8[$	$[13,8;17,9[$
nb de copies	16	18	24	11	3



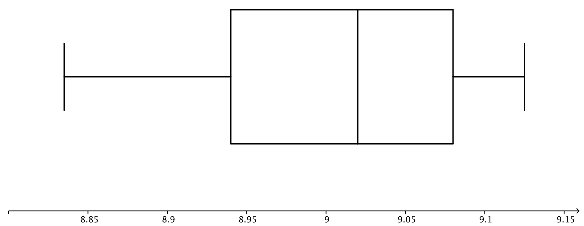
Exercice 5

1) On a pour les épinards :
— la moyenne : $2,70$
— la valeur médiane : $\frac{2,67+2,70}{2} = 2,685$
— l'écart type : $\approx 0,0777$

2) On a pour les épinards :
— le premier décile : $\frac{2,59+2,62}{2} = 2,605$
— le premier quartile : $2,64$
— le troisième quartile : $2,75$
— le neuvième décile : $\frac{2,82+2,83}{2} = 2,825$

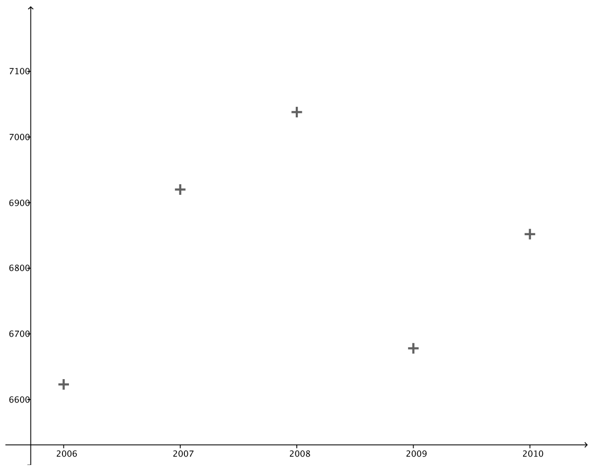


3) On a pour les lentilles :
— la moyenne : $8,997$
— la valeur médiane : $9,02$
— l'écart type : $\approx 0,1$
— le premier décile : $\frac{8,82+8.85}{2} = 8,835$
— le premier quartile : $8,94$
— le troisième quartile : $9,08$
— le neuvième décile : $\frac{9,11+9,14}{2} = 9,125$

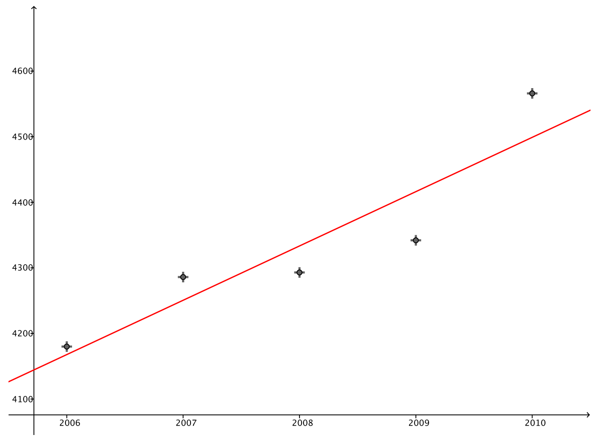


Exercice 6

1.

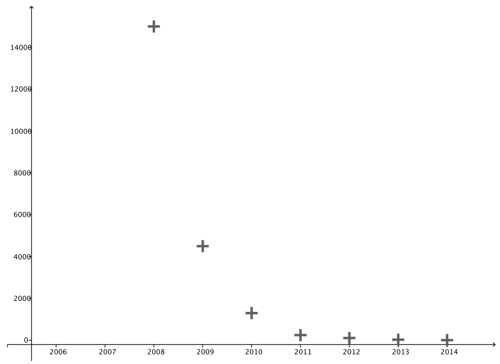


2. On note x le caractère dont les modalités sont les années de 2006 à 2010, et y le caractère correspondant aux importations horticoles en France. On a :
   — les moyennes : $\overline{x} = 2008$ et $\overline{y} = 6822,2$
   — les écarts types : $s_x \approx 1,41$ et $s_y \approx 153,3$
   — la covariance : $s_{x,y} = 43,2$
   D'où l'équation de la droite de régression affine obtenue par la méthode des moindres carrés :
$$y = \frac{s_{x,y}}{s_x^2}(x - \overline{x}) + \overline{y} \approx 21,6x - 36550,6.$$
3. Le coefficient de corrélation affine est égal à :
$$\rho_{x,y} = \frac{s_{x,y}}{s_x s_y} \approx 0,20.$$
   La régression affine n'est donc pas justifiée car la valeur absolue du coefficient de corrélation affine n'est pas proche de 1.
4. On note x le caractère dont les modalités sont les années de 2006 à 2010, et y le caractère correspondant aux importations horticoles au Canada. On a :
   — les moyennes : $\overline{x} = 2008$ et $\overline{y} = 4333,4$
   — les écarts types $s_x \approx 1,41$ et $s_y \approx 127,7$
   — la covariance : $s_{x,y} = 165,6$
   Le coefficient de corrélation affine est égal à :
$$\rho_{x,y} = \frac{s_{x,y}}{s_x s_y} \approx 0,92.$$
   La régression affine est donc justifiée ici car la valeur absolue du coefficient de corrélation affine est proche de 1. L'équation de la droite de régression affine obtenue par la méthode des moindres carrés est donnée par :
$$y = \frac{s_{x,y}}{s_x^2}(x - \overline{x}) + \overline{y} \approx 82,8x - 161929.$$
   Graphiquement :



Exercice 7

1.



On a :
— les moyennes : $\overline{x} = 2011$ et $\overline{y} \approx 3028,3$
— les écarts types $s_x = 2$ et $s_y \approx 5110,5$
— la covariance : $s_{x,y} = -7872,3$

Le coefficient de corrélation affine est égal à :

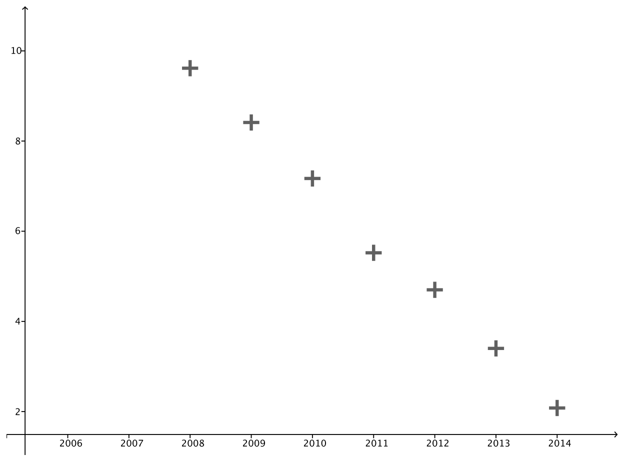
$$\rho_{x,y} = \frac{s_{x,y}}{s_x s_y} \approx -0,77.$$

Donc un ajustement affine n'est pas indiqué ici car $|\rho_{x,y}| = 0,77$ n'est pas assez proche de 1.

2. On pose $z = \ln(y)$.

(a)

x	2008	2009	2010	2011	2012	2013	2014
z	9,62	8,41	7,17	5,52	4,70	3,40	2,08



On a :
- les moyennes : $\overline{x} = 2011$ et $\overline{z} \approx 5,84$
- les écarts types : $s_x = 2$ et $s_z \approx 2,51$
- la covariance : $s_{x,z} \approx -5,01$

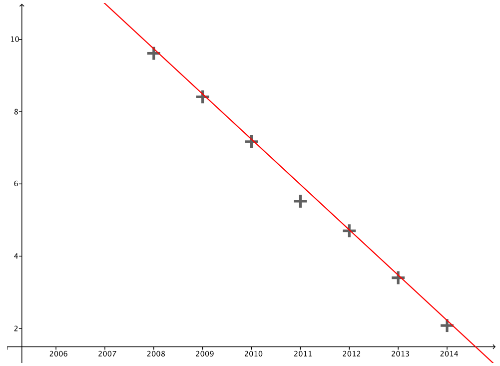
Le coefficient de corrélation affine est égal à :

$$\rho_{x,z} = \frac{s_{x,z}}{s_x s_z} \approx -0,999.$$

Donc un ajustement affine est indiqué ici car $|\rho_{x,z}| = 0,999$ est proche de 1.

b) L'équation de la droite de régression affine de z par rapport à x obtenue par la méthode des moindres carrés est donnée par :

$$z = \frac{s_{x,z}}{s_x^2}(x - \overline{x}) + \overline{z} \approx -1,2536x + 2526,77.$$



c) Ainsi $\ln(y) \approx -1,2536x + 2526,77$ ce qui donne en composant par la fonction exponentielle :

$$y = \exp(\ln(y)) \approx \exp(-1,2536x + 2526,77) = e^{2526,77}\left(e^{-1,2536}\right)^x.$$

En particulier, $\lambda = e^{2526,77}$ et $\mu = e^{-1,2536} \approx 0,29 < 1$ ce qui est cohérent puisque le nombre d'individus y est décroissant.