

UN EXEMPLE « DE REFERENCE »

D'EXERCICE DE STATISTIQUES A UNE VARIABLE

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

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

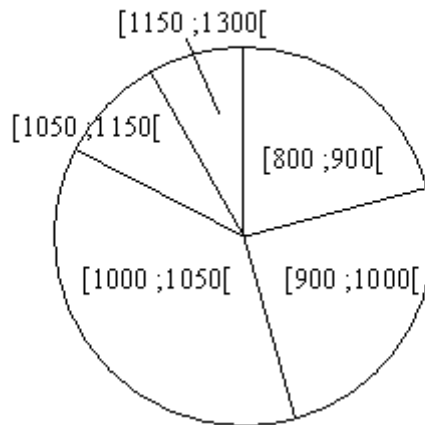
- 1) Représenter cette série par un diagramme circulaire
 - 2) Calculer le salaire moyen dans cette entreprise. Que penser d'un tel résultat ?
 - 3) Dans cette entreprise, combien d'employés gagnent au plus 1050 euros ?
- Dresser le polygone des effectifs cumulés croissants et lire une valeur approchée de la médiane et de Q_1 et Q_3
- 4) Calculer de manière précise la médiane et les quartiles Q_1 et Q_3
 - 5) Calculer l'écart type de cette série statistique
 - 6) Dans cette série statistiques se rajoute une sixième catégorie d'employés dont les salaires appartiennent à la classe [1300 ;1500[. Quel est l'effectif de cette classe sachant que le salaire moyen au sein de cette entreprise est alors de 1200

CORRECTION

- 1) Les angles des secteurs angulaires du diagramme circulaire ont des mesures proportionnelles aux effectifs :

Effectif	Angle
200	360°
42	$\frac{360 \times 42}{200} = 75,6^\circ$
49	$\frac{360 \times 49}{200} = 88,2^\circ$
74	$\frac{360 \times 74}{200} = 133,2^\circ$
19	$\frac{360 \times 19}{200} = 34,2^\circ$
16	$\frac{360 \times 16}{200} = 28,8^\circ$

Répartition des salaires au sein d'une entreprise



- 2) Pour calculer le salaire moyen de l'entreprise, il faut considérer le milieu de chaque classe :

Salaire	850	950	1025	1100	1225
Effectif	42	49	74	19	16

Le calcul de la moyenne est donc :

$$\bar{x} = \frac{\overbrace{\sum_{i=1}^5 n_i \times x_i}^{\text{somme des produits entre les valeurs et leurs effectifs}}}{\underbrace{\sum_{i=1}^5 n_i}_{\text{somme des effectifs}}} = \frac{n_1 \times x_1 + n_2 \times x_2 + \dots + n_5 \times x_5}{\underbrace{n_1 + n_2 + \dots + n_5}_{\text{effectif total}}} = \frac{42 \times 850 + 49 \times 950 + \dots + 16 \times 1225}{\underbrace{42 + 49 + \dots + 16}_{\text{effectif total}}} = \frac{198600}{200} = 993$$

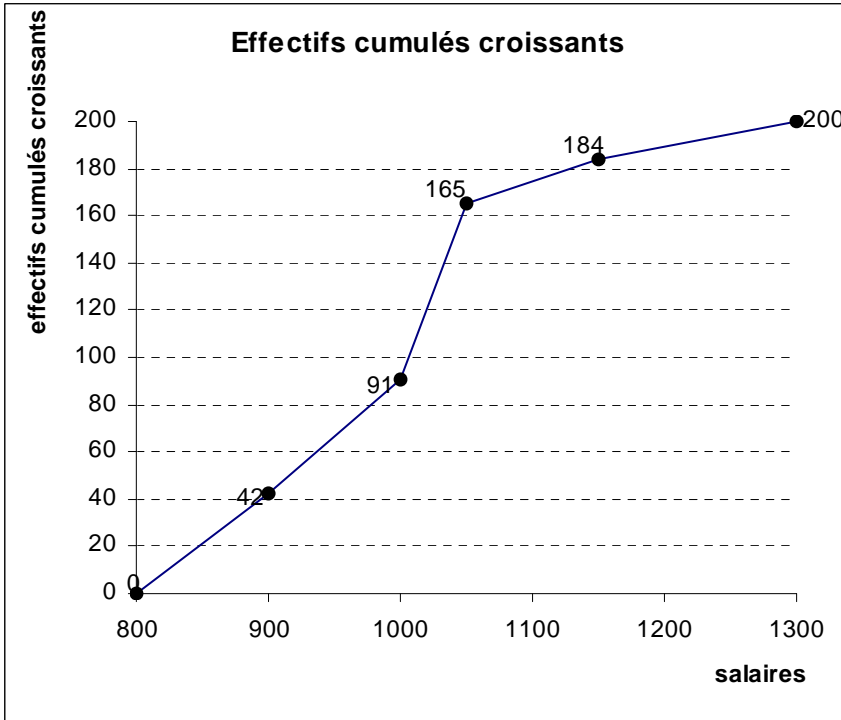
Le salaire moyen dans cette entreprise est donc de 993 €. Il n'est pas forcément très représentatif de cette entreprise, car plus de la moitié des employés y gagnent plus de 1000 euros !

- 3) 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 euros, au sein de cette entreprise

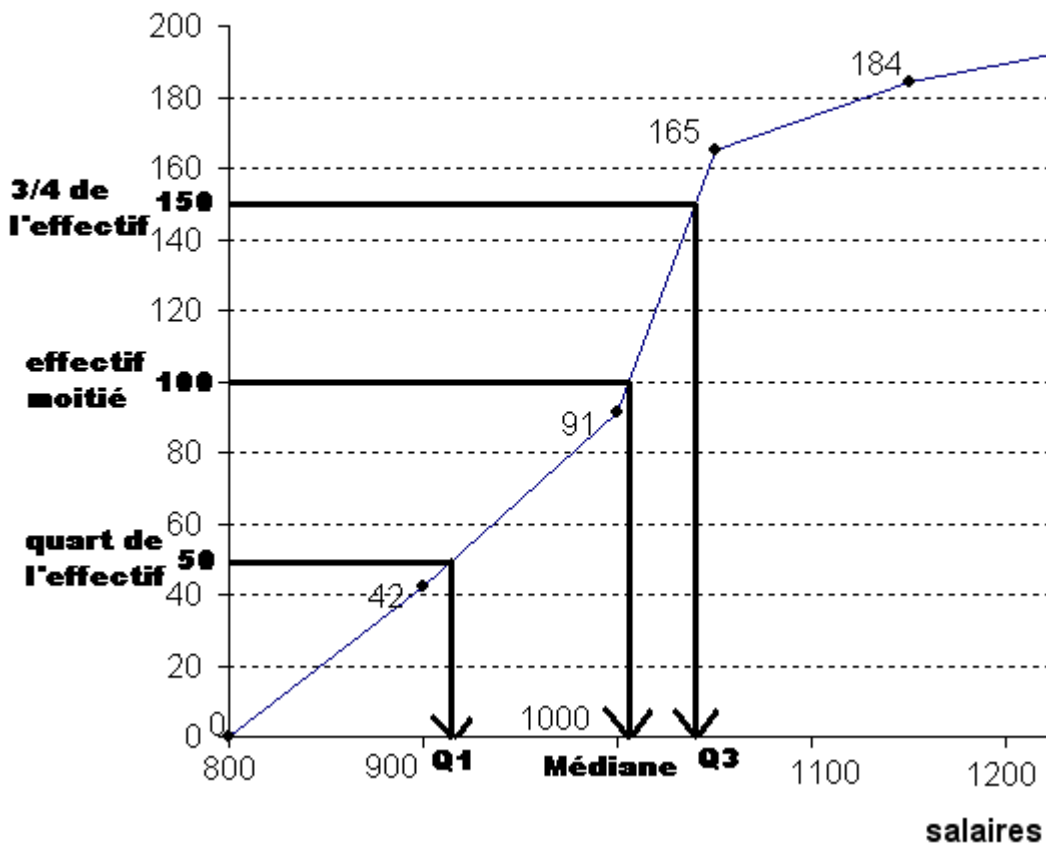
A partir de ce tableau, on dresse le polygone des effectifs cumulés croissants



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 Médiane ≈ 1010 , $Q_1 \approx 915$ et $Q_3 \approx 1050$

Effectifs cumulés croissants



4) 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 $y = 1,48x + p$.

Pour trouver la valeur de p , on utilise les coordonnées de A (ou B !): $y_A = 1,48x_A + p$ donc $p = y_A - 1,48x_A = 91 - 1,48 \times 1000 = -1389$. L'équation 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 $f : x \rightarrow 1,48x - 1389$, c'est-à-dire en résolvant l'équation $1,48x - 1389 = 100 \Leftrightarrow x = \frac{1489}{1,48} \approx 1006,08$. Ainsi $\boxed{Me \approx 1006}$

Puisque le quartile Q_3 semble lui aussi appartenir à l'intervalle $[1000;1050[$, on utilise la même droite, et on résout l'équation $1,48x - 1389 = 150 \Leftrightarrow x = \frac{1539}{1,48} \approx 1039,86$. Ainsi $\boxed{Q_3 \approx 1040}$

De la même manière, pour déterminer le quartiles Q_1 , 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 $0,49x - 399 = 50 \Leftrightarrow x = \frac{449}{0,49} \approx 916,33$ fournit $\boxed{Q_1 \approx 916}$

5) Commençons par calculer la Variance qui est la moyenne des carrés des écarts à la valeur, c'est-à-dire

$$V(x) = \frac{\overbrace{\sum_{i=1}^5 n_i \times (x_i - \bar{x})^2}^{\text{somme des produits entre les carrés des écarts des valeurs à la moyenne et leurs effectifs}}}{\underbrace{\sum_{i=1}^5 n_i}_{\text{somme des effectifs}}} = \frac{n_1 \times (x_1 - \bar{x})^2 + n_2 \times (x_2 - \bar{x})^2 + \dots + n_5 \times (x_5 - \bar{x})^2}{\underbrace{n_1 + n_2 + \dots + n_5}_{\text{effectif total}}}$$

$$= \frac{42 \times (850 - 993)^2 + 49 \times (950 - 993)^2 + \dots + 16 \times (1225 - 993)^2}{\underbrace{42 + 49 + \dots + 16}_{\text{effectif total}}}$$

$$= \frac{2103950}{200} = 10519,75$$

Enfin, $\sigma = \sqrt{V} \approx 102,6$

6) La moyenne de la série statistique constituée des deux sous séries (salaires inférieurs à 1300 euros, d'effectif 200 et

$$\text{tranche } [1300 ;1500[, \text{ d'effectif } n), \text{ vaut } \bar{y} = \frac{\overbrace{993}^{\text{moyenne de la première sous-série}} \times \overbrace{200}^{\text{effectif de la première sous-série}} + \overbrace{1400}^{\text{moyenne de la deuxième sous-série}} \times \overbrace{n}^{\text{effectif de la deuxième sous-série}}}{\underbrace{200 + n}_{\text{effectif total}}}$$

Puisque $\bar{y} = 1200$, on résout l'équation

$$\frac{993 \times 200 + 1400 \times n}{200 + n} = 1200 \Leftrightarrow 198600 + 1400n = 240000 + 1200n \Leftrightarrow 200n = 41400 \Leftrightarrow n = 207.$$

Il y aura donc 207 personnes dont le revenu appartient à la tranche $[1300 ;1500[$.