

# Chapitre 6 : Statistiques

## I) Vocabulaire :

Population: C'est l'ensemble des "individus" sur les quels porte l'étude statistique.

Echantillon: C'est une partie de la population de taille fixe  $n$

Caractère: C'est la propriété étudiée

→ Qualitatif = lorsque qu'il ne prend pas de valeurs numériques

→ Quantitatif = valeurs numériques -

→ discret = s'il prend des valeurs isolées

→ continu dans le cas contraire -

à chaque valeur du caractère on lui associe un effectif. (un nombre d'occurrences)

### Exercice :

Dans une classe on étudie la couleur des yeux

Population: les élèves d'une classe

Caractère: la couleur des yeux

→ qualitatif.  
discret

Dans une usine on étudie la longueur des pièces produites

Population: les pièces produites

Caractère: longueur

→ quantitatif.  
continu

Dans une entreprise on étudie le nombre de pièces défectueuses

Population: les pièces

Caractère: le nombre de pièces défectueuses

→ quantitatif.  
discret.

## II) Représentations

### 1) Tableaux :

Serie A: notes obtenues à un contrôle :

2-3-3-4-5-6-6-7-7-7-8-8-8-8-8-9-9-9-9-9-9-10-10-11-11-11-13-13-15-16

Serie B: Salaires dans une entreprise -

entre 900€ et 1200€ : 30 salariés | entre 1400€ et 1600€ = 60 salariés | entre 1800€ et 2000€ : 40 salariés

entre 1200€ et 1400€ : 30 salariés | entre 1600€ et 1800€ = 80 salariés | entre 2000€ et 2400€ : 40 salariés

Série C: Dans un club multisports, 17% jouent au handball, 25% au rugby et 58% au tennis

Pour chacune des séries: \* indiquer la population, le caractère, le type de caractère  
\* résumer ces séries dans 3 tableaux.

Série A: Population: les élèves caractère: notes type: quantitatif, discret.

|              |         |         |   |   |   |    |    |    |    |    |    |    |    |                   |
|--------------|---------|---------|---|---|---|----|----|----|----|----|----|----|----|-------------------|
| Notes        | $x_1=2$ | $x_2=3$ | 4 | 5 | 6 | 7  | 8  | 9  | 10 | 11 | 13 | 15 | 16 | $\rightarrow x_i$ |
| effectifs    | $n_1=1$ | $n_2=2$ | 1 | 1 | 2 | 3  | 5  | 6  | 2  | 3  | 2  | 1  | 1  | $\rightarrow m_i$ |
| fréquences % | 3       | 7       | 3 | 3 | 7 | 10 | 17 | 20 | 7  | 10 | 7  | 3  | 3  |                   |

$N=30$

Définition: On considère une série statistique  $X$  à caractère quantitatif. dont le  $p$  valeurs sont  $x_1, x_2, x_3, \dots, x_p$  d'effectifs associés  $n_1, n_2, \dots, n_p$  on note  $N = n_1 + n_2 + \dots + n_p$   
à chaque valeur  $x_i$  on associe sa fréquence  $f_i = \frac{n_i}{N}$  ou  $f_i = \frac{n_i}{N} \times 100 \%$ .

Remarque: lorsque l'on a beaucoup de valeurs on peut faire un regroupement en classes.

|            |          |           |            |            |
|------------|----------|-----------|------------|------------|
| Notes      | $[0; 5[$ | $[5; 10[$ | $[10; 15[$ | $[15; 20]$ |
| effectifs  | 4        | 17        | 7          | 2          |
| fréquences | 13       | 57        | 23         | 7          |

Série B: Population: les salariés caractère: salaire type: quantitatif, continu.

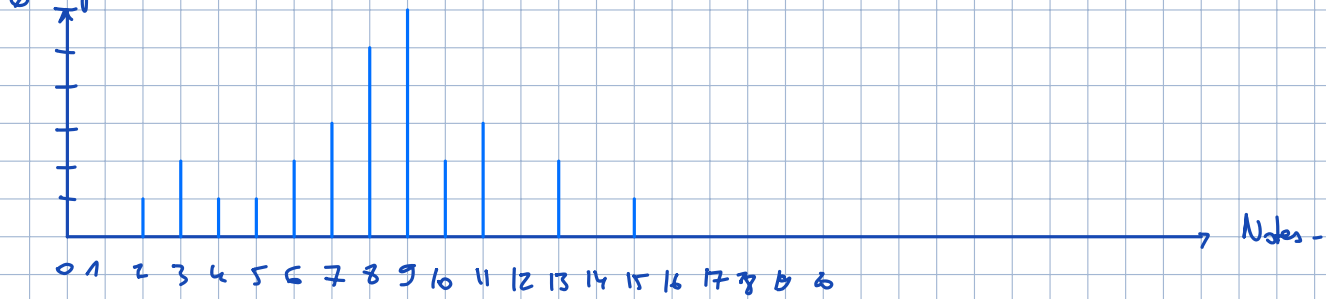
|       |          |               |                |                |                |                |                |       |
|-------|----------|---------------|----------------|----------------|----------------|----------------|----------------|-------|
|       |          | 1050          | 1300           | 1500           | 1700           | 1900           | 2200           |       |
| $x_i$ | Salaire  | $[900; 1200[$ | $[1200; 1400[$ | $[1400; 1600[$ | $[1600; 1800[$ | $[1800; 2000[$ | $[2000; 2400]$ | Total |
| $n_i$ | effectif | 30            | 30             | 60             | 80             | 40             | 40             | 280   |
|       | freq %   | 11            | 11             | 21             | 29             | 14             | 14             | 100   |

Série C: Population: les adhérents caractère: sport type: qualitatif, discret.

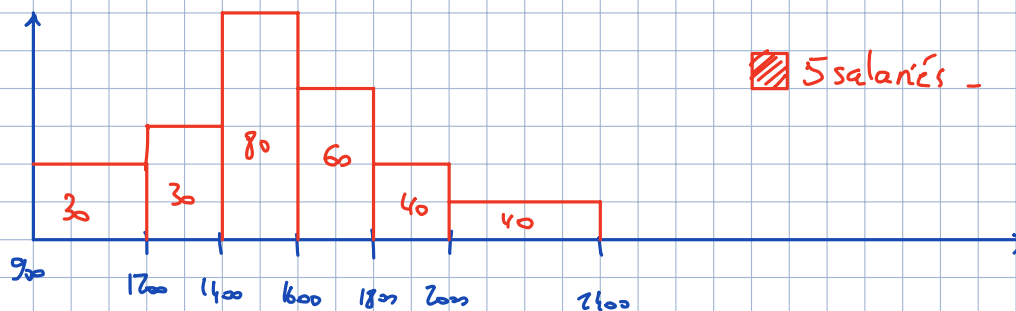
|        |          |       |        |        |
|--------|----------|-------|--------|--------|
| Sport  | Handball | rugby | tennis | total. |
| freq % | 17       | 25    | 58     | 100.   |

## 2) Graphiques:

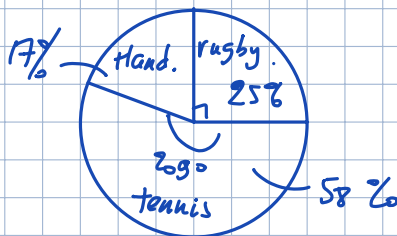
a) Diagramme en bâtons.



b) Histogramme



c) Diagramme circulaire



|     |    |
|-----|----|
| 100 | 58 |
| 360 | ?  |

$$\frac{58 \times 360}{100} \approx 209^\circ$$

## III) Caractéristiques:

1) Caractéristiques de position.

a) Moyenne  $\bar{x}$

Définition: Soit  $X$  une série statistique de  $p$  valeurs  $x_1, \dots, x_p$  d'effectifs associés  $m_1, \dots, m_p$  on note  $\bar{x}$  la moyenne de la série telle que:

$$\bar{x} = \frac{x_1 m_1 + x_2 m_2 + x_3 m_3 + \dots + x_p m_p}{m_1 + m_2 + m_3 + \dots + m_p} = \frac{1}{N} (x_1 m_1 + \dots + x_p m_p)$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^p x_i m_i = \sum_{i=1}^p x_i \times \frac{m_i}{N} = \sum_{i=1}^p x_i f_i$$

Ex:

Calculer les moyennes des séries A et B:

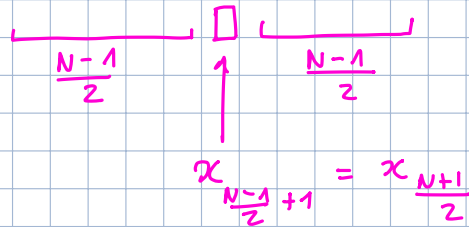
Série A:  $\bar{x} \approx 8,47$

Série B:  $\bar{x} \approx 1645 \text{ €}$

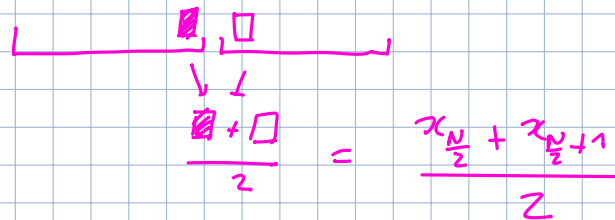
## b) Médiane:

Définition: Soit  $X$  une série statistique, on appelle médiane de  $X$  une valeur qui partage la série ordonnée en 2 séries de même effectif - telles que: Soit  $N$  l'effectif total

\* Si  $N$  est impair



\* Si  $N$  est pair



Exemple Série 1:

$N=9$   
impair

8 - 9 - 7 - 11 - 11 - 10 - 12 - 13 - 15  
7 - 8 - 9 - 10 - 11 - 11 - 12 - 13 - 15 (série ordonnée)  
4 4  
médiane = 11

Série 2:

$N=8$   
pair

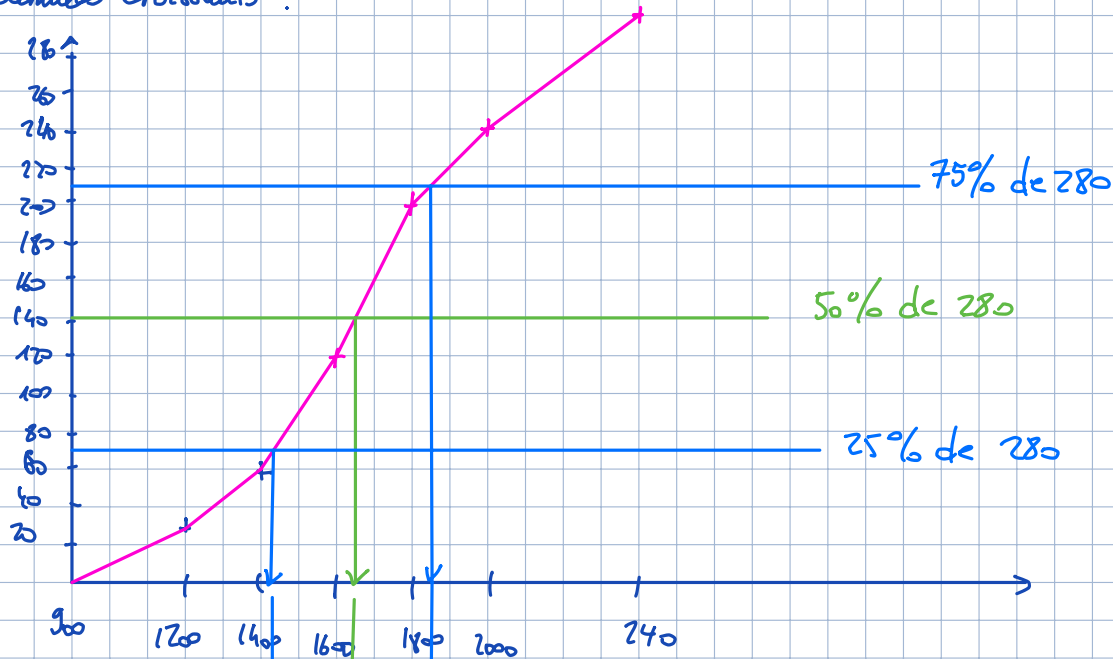
5 - 10 - 11 - 7 - 8 - 13 - 15 - 9  
5 - 7 - 8 - 9 - 10 - 11 - 13 - 15 (série ordonnée)  
9 + 10 = 9,5  
2  
médiane = 9,5

Exercice: Médiane de la série A:  $N=30$   $x_{\frac{30}{2}} = x_{15} = 8$   $x_{16} = 9$   
médiane:  $\frac{8+9}{2} = 8,5$ .

Cas d'un regroupement en classe:

| Salaires | $[900; 1200[$ | $[1200; 1400[$ | $[1400; 1600[$ | $[1600; 1800[$ | $[1800; 2000[$ | $[2000; 2400]$ | Total |
|----------|---------------|----------------|----------------|----------------|----------------|----------------|-------|
| Effectif | 30            | 30             | 60             | 80             | 40             | 40             | 280   |
| Ecc      | 30            | 60             | 120            | 200            | 240            | 280            |       |

Effectifs cumulés croissants.



le 1<sup>er</sup> quartile:  $Q_1 \approx 1440$

médiane  $\approx 1745$

le 3<sup>ème</sup> quartile:  $Q_3 \approx 1850$ .

c) les Quartiles ( $Q_1$  et  $Q_3$ )

Définition: Le premier quartile est la plus petite valeur de la série qui réalise au minimum 25% de l'effectif total. on note  $Q_1$

(idem pour la 3<sup>ème</sup> Quartile avec 75% on note  $Q_3$ )

Exemple: Série 1: 7 - 8 - 9 - 10 - 11 - 11 - 12 - 13 - 15 (série ordonnée)

$N = 9$   
impair

$25\% \text{ de } N = 25\% \text{ de } 9 = 2,25 \approx 3 \text{ par excès}$

donc  $Q_1$  est la 3<sup>ème</sup> valeur  $Q_1 = 9$

$75\% \text{ de } N = 75\% \text{ de } 9 = 6,75 \approx 7 \text{ par excès}$

donc  $Q_3$  est la 7<sup>ème</sup> valeur  $Q_3 = 12$

Sérez: 5 - 7 - 8 - 9 - 10 - 11 - 13 - 15

$N = 8$   
pair

$Q_1 = 7$   
2<sup>ème</sup> valeur

$Q_3 = 11$   
6<sup>ème</sup> valeur

### Exercice 1 : Notes

Le tableau ci-dessous indique la répartition des étudiants en BTS CRCI du Grand Est suivant leur note au CCF de mathématiques.

| Notes                        | 5  | 7  | 8   | 10  | 12  | 13  | 14          |
|------------------------------|----|----|-----|-----|-----|-----|-------------|
| Effectif                     | 54 | 45 | 56  | 21  | 17  | 11  | 6           |
| Effectifs cumulés croissants | 54 | 99 | 155 | 176 | 193 | 204 | 210 ← total |

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.

2.  $N = 210$  pair

105<sup>ème</sup> valeur = 8  
106<sup>ème</sup> valeur = 8

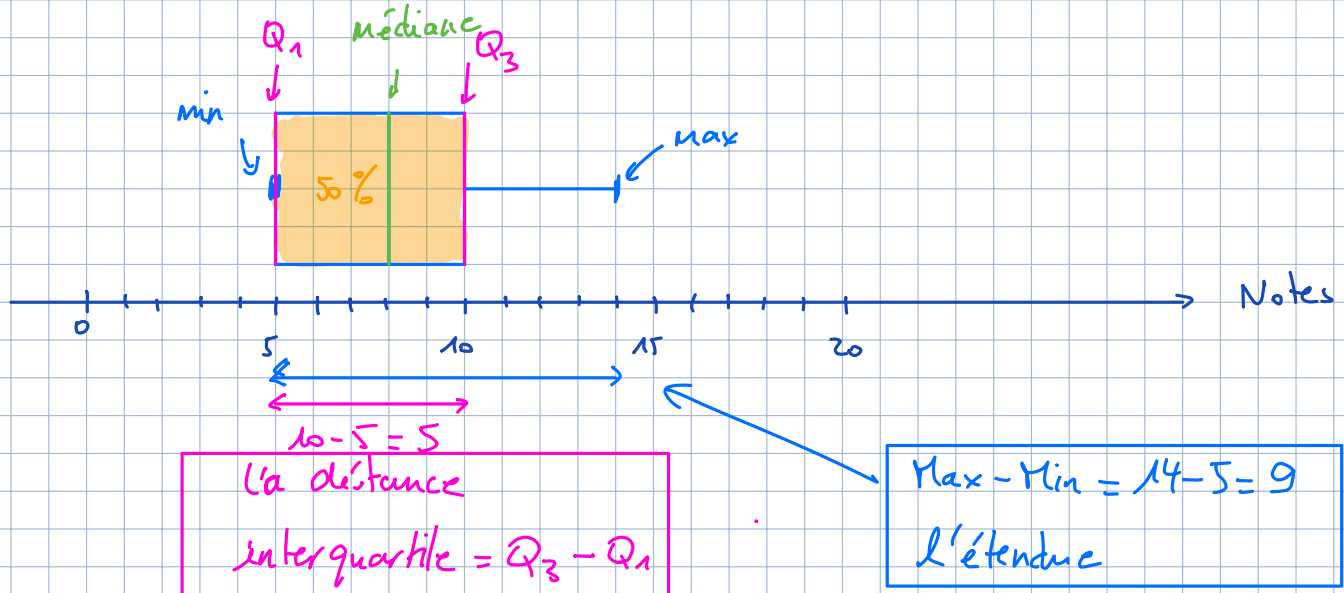
médiane = 8

$N = 210$

25% de 210 =  $\frac{210}{4} = 52,5 \approx 53^{\text{ème}}$  valeur  $Q_1 = 5$   
↑ arrondi par excès

75% de 210 =  $\frac{3 \times 210}{4} = 157,5 \approx 158^{\text{ème}}$  valeur  $Q_3 = 10$

non demandé Boîte à moustaches ou diagramme en boîte



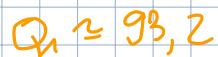
la distance interquartile et l'étendue  
sont des caractéristiques de dispersion.

Au cours d'un TP d'atelier, on a relevé les longueurs (mm) des barres usinées par les élèves d'une circonscription :

1. Calculer la longueur moyenne des barres pour ce TP.  $\bar{x} \approx 94,15 \text{ à } 10^{-2}$

- (TI)** ① STATS → modifier →

| L <sub>1</sub> | L <sub>2</sub> | quitter | L <sub>2</sub> /94 $\xrightarrow{\text{sto}}$ L <sub>3</sub> | STATS modifier |
|----------------|----------------|---------|--------------------------------------------------------------|----------------|
| 87,5           | 5              |         | x100                                                         | L <sub>3</sub> |
| 87,9           | 9              |         | done                                                         | 5, ...         |
| ⋮              |                |         |                                                              |                |
| 112,5          | 7              |         |                                                              |                |

$$\bar{x} \pm 94,15 \cdot 10^{-2}$$


$$Q_3 \approx 10^{4,8}$$

4) Par lecture graphique:  $87 - 22 = 65\%$  des longueurs sont entre 92 et 167mm

Ex 3:

|                             |   |    |    |    |    |    |    |    |    |    |    |    |
|-----------------------------|---|----|----|----|----|----|----|----|----|----|----|----|
| Vitesse du vent (en noeuds) | 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  |
| Effectif cumulés croissants | 1 | 3  | 4  | 5  | 9  | 14 | 17 | 21 | 25 | 27 | 28 | 30 |

2)  $\frac{27}{30} \times 100 = 90\%$  donc 10% sans production.

3)  $\bar{x} \approx 24,9$

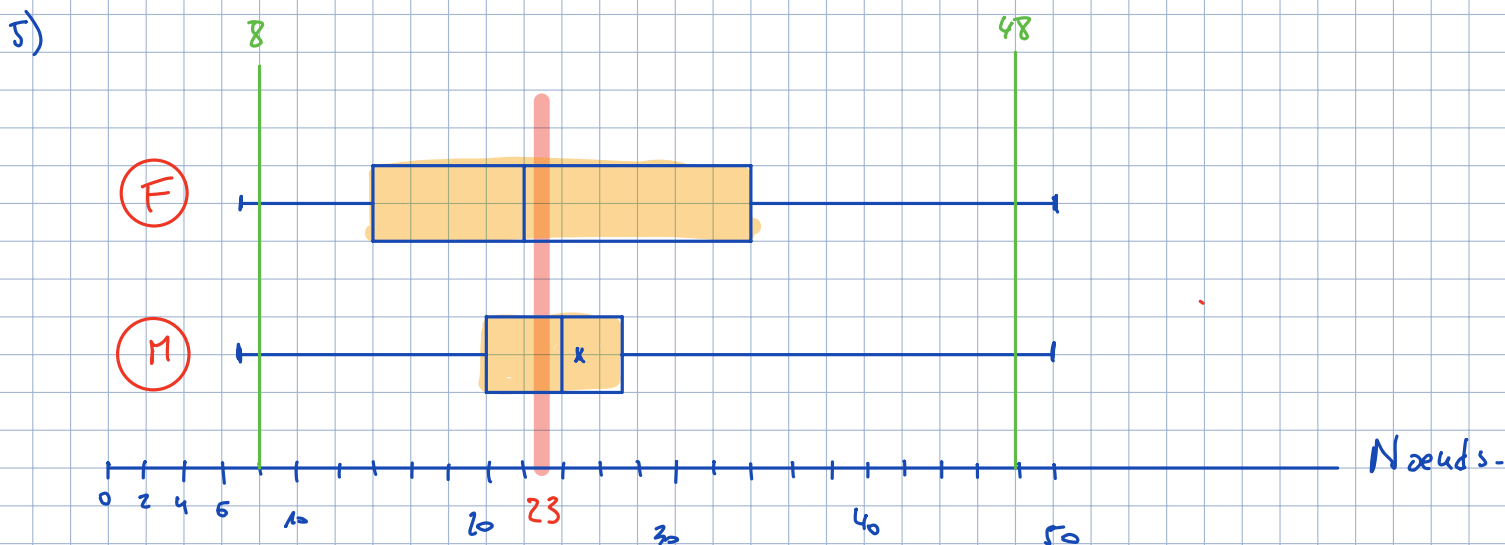
étendue =  $50 - 7 = 43$   
= Max - Min

4) médiane = 24

$Q_1 = 8^{\text{ème}} \text{ valeur} = 20$

$Q_3 = 23^{\text{ème}} \text{ valeur} = 27$

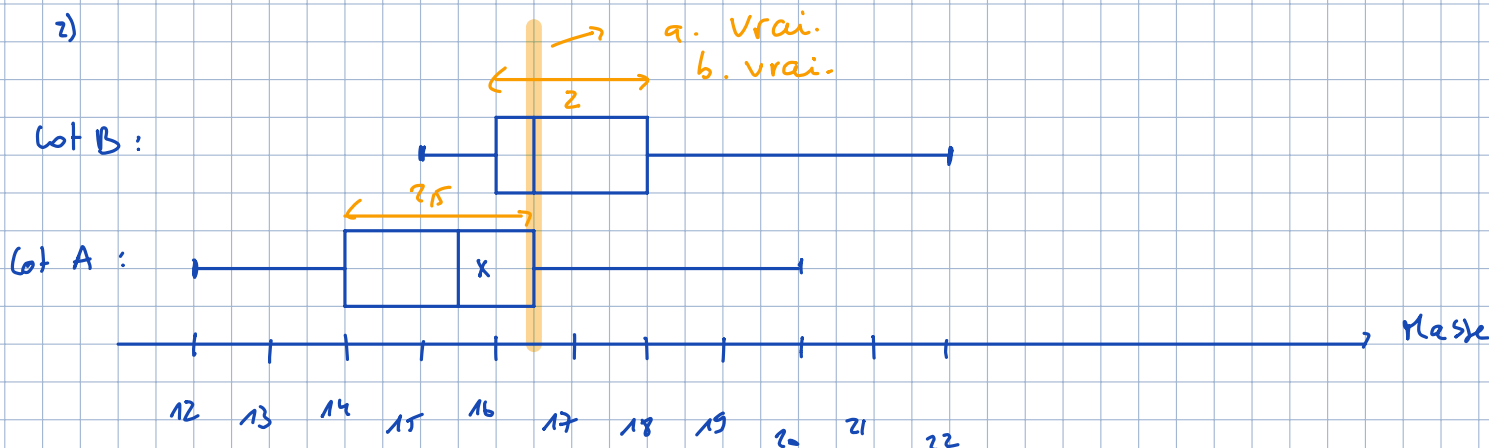
écart interquartile =  $Q_3 - Q_1 = 27 - 20 = 7$



la montage est plus adaptée car 50% des jours sont proches du rendement max.

Ex4: 1)  $\left\{ \begin{array}{l} \min = 15 \quad \max = 22 \quad Q_1 = 16 \quad Q_3 = 18 \quad \bar{x} \approx 16,81 \\ \text{étendue} = 22 - 15 = 7 \quad \text{médiane } 16,5 \end{array} \right.$

2)





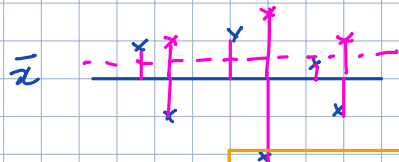
## 2) Caractéristiques de dispersion.

a) étendue Valeur maximale - Valeur minimale

b) écart interquartile:  $Q_3 - Q_1$  environ 50% de l'effectif et contenu dans l'intervalle  $[Q_1; Q_3]$ .

c) Ecart type:

on appelle variance:  $V(X) = \frac{m_1(x_1 - \bar{x})^2 + m_2(x_2 - \bar{x})^2 + \dots + m_p(x_p - \bar{x})^2}{N}$



$$V(X) = \frac{1}{N} \sum_{i=1}^p m_i (x_i - \bar{x})^2$$

l'écart-type:

$$\sigma(X) = \sqrt{\frac{1}{N} \sum_{i=1}^p m_i (x_i - \bar{x})^2}$$

(TI) Stats → Calc  
StatVar ed  
L1 L2  
→  $\sigma x \approx$

## II - Statistiques à 2 variables $(X; Y)$ une série statistique à 2 variables -

1) Nuage de points: l'ensemble  $(x_i; y_i)$  des valeurs de X et Y et appelé nuage de points

Exemple:

| Année                                       | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 |
|---------------------------------------------|------|------|------|------|------|------|
| nombre d'adhésés d'un club de rugby vosgien | 70   | 90   | 115  | 140  | 170  | 220  |
| rang.                                       | 1    | 2    | 3    | 4    | 5    | 6    |

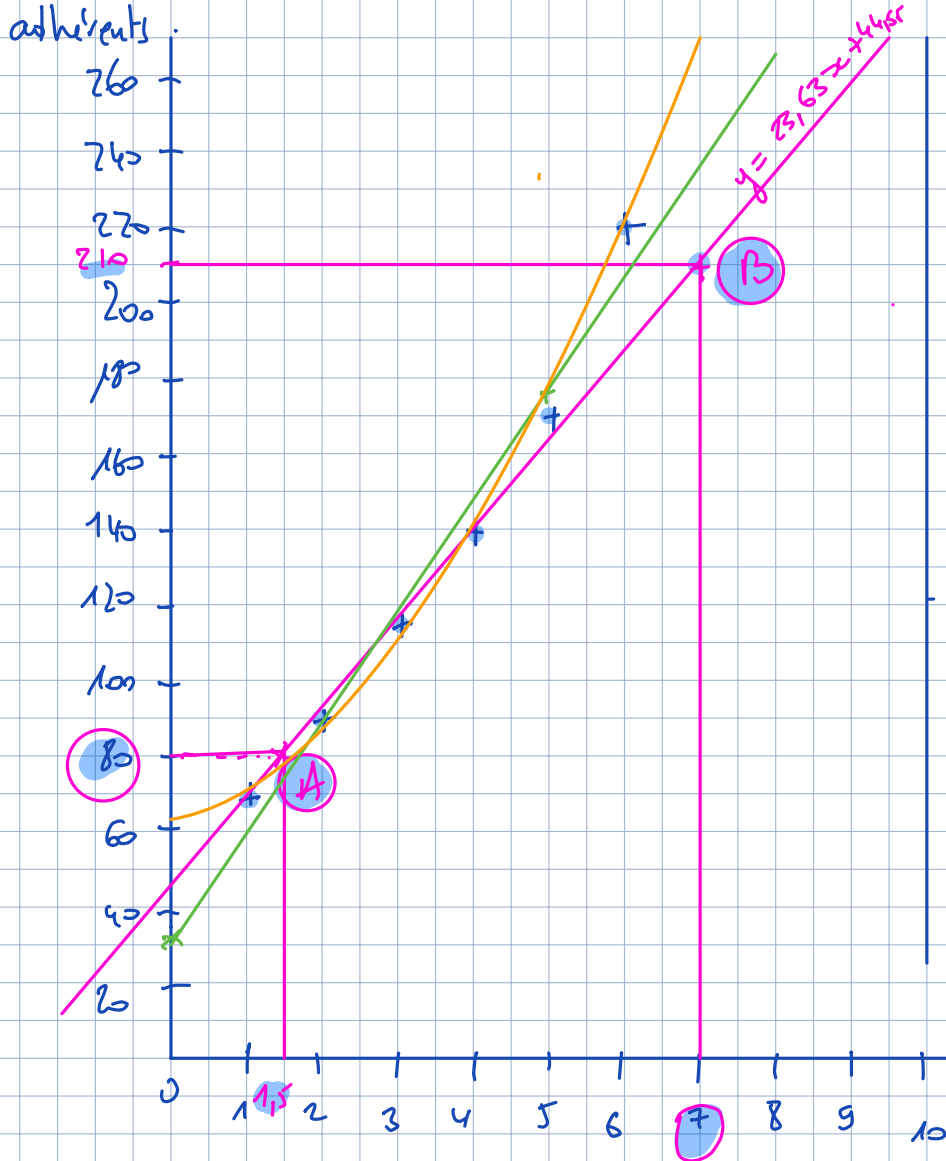
(TI) . nombre → L2 . (WW) régression.  
rang → L1

$X_1 \rightarrow$  rang  
 $Y_1 \rightarrow$  nombre.

- Graphstat - Aff  
L1 L2

zoom → zoom stat

a) Régression à la règle:



$$A(1,5; 80) \quad B(7; 210)$$

$$a = \frac{y_A - y_B}{x_A - x_B} = \frac{80 - 210}{1,5 - 7} \approx 23,63$$

$$y = ax + b$$

$$y = 23,63x + b$$

$$80 = 23,63 \times 1,5 + b$$

$$b = 80 - 23,63 \times 1,5 = 44,55$$

$$y = 23,63x + 44,55$$

en 2030 : rang 12

$$23,63 \times 12 + 44,55 \approx 328 \text{ adhérents.}$$

b) régression linéaire par la méthode des moindres carrés :

TI Stats, calc → Reglin(ax+b) L<sub>1</sub>, L<sub>2</sub> ⇒ a = 29 ; b ≈ 32,7

NN → regression → affine → a = 29 b ≈ 32,7.

$$y = 29x + 32,7$$

en 2030 :  $29 \times 12 + 32,7 \approx 380$  adhérents.

Coefficient de Corrélation : r

r<sup>2</sup> plus il est proche de 1 meilleure est la régression.

ici r<sup>2</sup> ≈ 0,973

c) Ajustement exponentiel.

| x rang :  | 1    | 2    | 3    | 4    | 5    | 6    |
|-----------|------|------|------|------|------|------|
| y adh.    | 70   | 90   | 115  | 140  | 170  | 220  |
| z = ln(y) | 4,14 | 4,49 | 4,74 | 4,94 | 5,13 | 5,39 |

TI  $\ln(L_2) \xrightarrow{\text{sto}} L_3$

# régression linéaire de z en x

(TI) Calc → reglin(ax+b) L1; L3

(NW) régression affine  $X_2 = X_1$  ;  $Y_2 = \ln(Y_1)$   
 $a \approx 0,224$   $b \approx 4,044$

$a \approx 0,224$   $b \approx 4,044$   
 ici  $r^2 \approx 0,998$

donc 
$$\begin{cases} z = 0,224x + 4,044 \\ z = \ln(y) \end{cases}$$

exp  $\ln(y) = 0,224x + 4,044$   
 $y = e^{0,224x + 4,044}$

$$e^{m+n} = e^m \times e^n$$

$$y = e^{4,044} \times e^{0,224x}$$

$$y = 57,054 e^{0,224x}$$

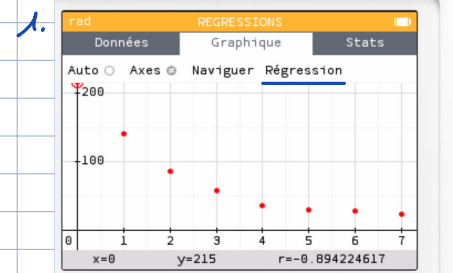
ajustement exponentiel.

en 2030 :  $x=12 \Rightarrow y \approx 838$  Adhérents

Ex1:

Ex1:

|                       |     |      |      |      |      |      |      |      |
|-----------------------|-----|------|------|------|------|------|------|------|
| $L_1 \leftarrow t_i$  | 0   | 1    | 2    | 3    | 4    | 5    | 6    | 7    |
| $L_2 \leftarrow v_i$  | 215 | 140  | 85   | 57   | 36   | 29   | 27   | 22   |
| $m_i = \ln(v_i - 15)$ | 5,3 | 4,83 | 4,25 | 3,74 | 3,04 | 2,64 | 2,48 | 1,95 |



le nuage ne ressemble pas à une droite.

(TI)  $\ln(L_2 - 15) \xrightarrow{\text{sto}} L_3$  (NW)  $t_i \rightarrow X_1$   $Y_1 \leftarrow v_i$   
 $X_2 = X_1$  ;  $Y_2 = \ln(Y_1 - 15)$

2) 3)  $m = -0,485t + 5,224$  (TI) Stats calc → reglin(ax+b) L1 L3

4) on a  $\begin{cases} m = \ln(v - 15) \\ m = -0,485t + 5,224 \end{cases}$   
 exp  $\ln(v - 15) = -0,485t + 5,224$   
 $v - 15 = e^{-0,485t + 5,224}$   
 $v - 15 = e^{5,224} \times e^{-0,485t}$

$$+15 \downarrow \quad V = e^{5,224} \times e^{-0,485t} + 15 \quad \downarrow +15$$

$$V = 185,7 \times e^{-0,485t} + 15$$

$$\text{done } \alpha = 185,7 \quad \beta = -0,485 \quad \gamma = 15$$