

## Semaine 2, première partie : Dessine-moi un résumé

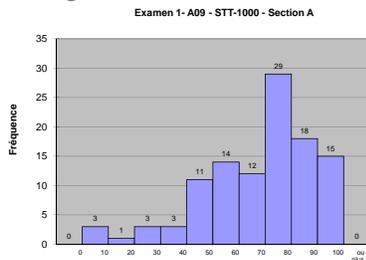
# 1 Statistique descriptive :

## 2 1-Description graphique

### 1-Description graphique : les diagrammes

- histogramme
- diagramme en boîte (de quartiles)
- diagramme tiges-feuilles
- diagramme à bandes
- pictogramme
- diagramme circulaire (secteurs, pointes de tarte, camembert)
- diagramme de Carroll
- diagramme de Venn
- diagramme cartésien (nuage de points)
- diagramme sagittal
- diagramme à ligne brisée
- diagramme en mosaïque

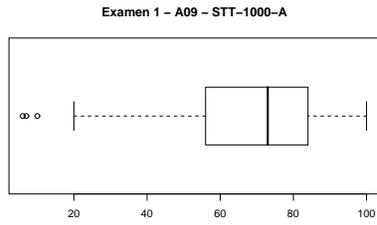
### Histogramme



#### À retenir :

- Représente la distribution d'une variable continue.
- La surface des rectangles est proportionnelle à la fréquence des classes.
- Les rectangles sont collés.
- Le nombre de classes est "suffisant" sans être exagéré.
- L'axe vertical commence à 0.

### Diagramme en boîte



À retenir :

- Représente la position des quartiles et des extrêmes d’une variable continue.
- Un quart des observations se situe dans chaque partie.
- Plusieurs logiciels tracent les moustaches jusqu’au minimum et maximum sans identifier de valeurs extrêmes.
- L’axe numérique est essentiel (vertical ou horizontal).
- Surtout utile pour placer plusieurs échantillons côte à côte.

### Diagramme tige-feuilles

bas : 6,0,7,0,10,0

```

2 | 0 1 7
3 | 0 2 6 8
4 | 2 2 3 4 5 7 7 7 8 9
5 | 0 3 4 4 6 7 7 8
6 | 0 0 0 2 2 2 2 3 4 4 7 9
7 | 0 0 1 1 1 1 2 3 4 4 5 5 5 6 6 8 8 8 8 9 9
8 | 0 1 2 2 4 4 4 5 7 8 9
9 | 0 0 0 1 2 2 4 4 5 6 6 7 7
10| 0
    
```

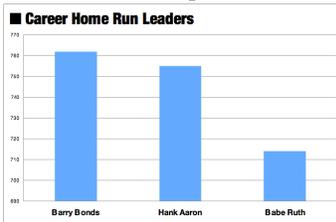
Lecture : 3|1 signifie 31

À retenir :

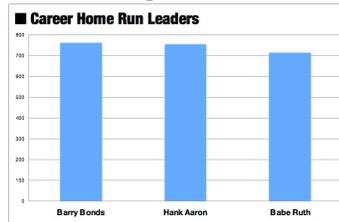
- 
- 
- 
- 

### Diagramme à bandes

Le mauvais exemple :



Le bon exemple :



<http://henkayu.blogspot.ca/2011/01/rules-of-thumb-for-visualization.html>

À retenir :

- Représente la distribution d’une variable...
- Les rectangles sont séparés.
-

**Pictogramme**

Un mauvais exemple du Erickson times :



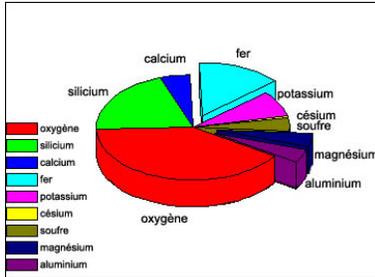
[www.forbes.com/sites/naomirobins/2012/02/16/misleading-graphs-figures-not-drawn-to-scale/](http://www.forbes.com/sites/naomirobins/2012/02/16/misleading-graphs-figures-not-drawn-to-scale/)

À retenir :

—

**Diagramme circulaire**

Un "mauvais" exemple de l'Agence spatiale canadienne :



À retenir :

—

—

—

—

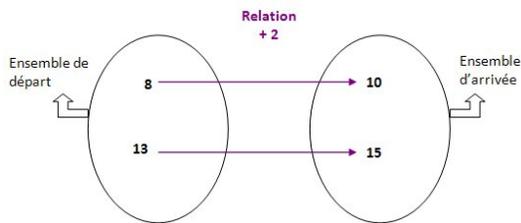
**Diagramme de Carroll**

	Prime	Not prime
Even	2	4, 6, 8, 10, 12, 14, 16, 18, 20
Not even	3, 5, 7, 11, 13, 17, 19, 23, 29	1, 9, 15, 21, 25, 27, 33, 35, 39

[http://en.wikipedia.org/wiki/File:Carroll\\_diagram.2.png](http://en.wikipedia.org/wiki/File:Carroll_diagram.2.png)

À retenir :

—  
—  
—

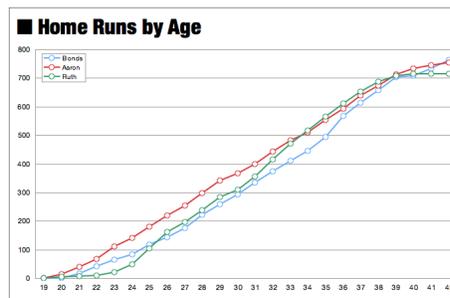


[biblio.alloprof.qc.ca/PagesAnonymes/DisplayFiches.aspx?ID=2096](http://biblio.alloprof.qc.ca/PagesAnonymes/DisplayFiches.aspx?ID=2096)

À retenir :

—  
—  
—

**Diagramme à ligne brisée**



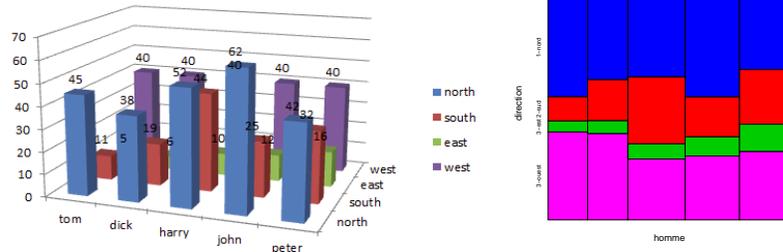
Un autre exemple par Joshua Maciel :

<http://henkayu.blogspot.ca/2011/01/rules-of-thumb-for-visualization.html>

À retenir :

—  
—

**Diagramme en mosaïque**



[www.forbes.com/sites/naomirobins/2012/05/30/winner-of-the-bad-graph-contest-announced-2/](http://www.forbes.com/sites/naomirobins/2012/05/30/winner-of-the-bad-graph-contest-announced-2/)

À retenir :

- 
- 
- 
- 

**3 2-Description numérique**

**3.1 a) Mesures de position**

**2-Description numérique a) Mesures de position**

- moyenne ( $\bar{X}$ )
- médiane ( $\tilde{X}$ )
- mode ( $M$ )
- moyenne tronquée d'ordre  $k$  ( $\bar{X}_k$ )
- quartiles ( $Q_1, Q_2, Q_3$ )
- quantiles d'ordre  $\gamma$  ( $X_{(\gamma n + 0,5)}$ )
- rang cinquième ( $R_5(x_i)$ )
- rang centile ( $R_{100}(x_i)$ )

**Rang cinquième**

Procédure à suivre pour assigner un rang cinquième à une donnée

1. Placer les données en ordre décroissant.
2. Compter le nombre de données  $n$ .
3. Séparer les données en cinq groupes de fréquence approximativement égale, en veillant à ce que deux données identiques (s'il y a lieu) soient dans le même groupe.
4. Assigner un rang à chaque groupe, en commençant par les valeurs les plus élevées.

[biblio.alloprof.qc.ca/Pages/Anonymes/DisplayFiches.aspx?ID=2751](http://biblio.alloprof.qc.ca/Pages/Anonymes/DisplayFiches.aspx?ID=2751)

Remarque : Plus le rang cinquième est élevé, plus la donnée a une valeur faible.

**Rang centile**

**Définition 1** : Le rang centile correspond au pourcentage de données inférieures ou égales à celle qu'on tente de positionner dans la distribution.

**Définition 2** :

$$R_{100}(x) = \frac{(\text{nb données} < x) + \left(\frac{\text{nb données} = x}{2}\right)}{n} \times 100$$

**Remarque** : Plus le rang centile est élevé, plus la donnée a une valeur élevée.

biblio.alloprof.qc.ca/Pages/Anonymes/DisplayFiches.aspx ?ID=2752#

**Exemple**

Déterminer le rang cinquième et le rang centile de 4 dans l'échantillon ci-dessous.

1	8
2	8
3	9
3	10
4	11
4	12
5	12
6	14
7	17
8	

**Exercice**

Pour alimenter la réflexion sur les caractéristiques de la moyenne, de la médiane et du diagramme en boîte, voici une application intéressante, sur le site du *National Council of Teachers of Mathematics (NCTM)* :

[illuminations.nctm.org/ActivityDetail.aspx ?ID=160](http://illuminations.nctm.org/ActivityDetail.aspx ?ID=160)

**3.2 b) Mesures de dispersion****2-b) Mesures de dispersion**

- étendue :  $X_{(n)} - X_{(1)}$
- variance :  $S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$
- écart-type :  $S = \sqrt{S^2}$
- écart moyen :  $\frac{1}{n} \sum |X_i - \bar{X}|$
- écart interquartile :  $EIQ = Q_3 - Q_1$
- étendue des quarts :  $EQ_1 = Q_1 - X_{(1)}, EQ_2, EQ_3, EQ_4$

**Quelques questions**

- Pourquoi préfère-t-on l'écart-type à la variance ?
- Pourquoi préfère-t-on l'écart-type à l'écart moyen ?
- Devrait-on préférer la moyenne ou la médiane ?
- Combien de données faut-il pour calculer un écart-type ?
- Combien de données faut-il pour tracer un diagramme en boîte ?

**Quelques questions**

Pouvez-vous construire deux jeux de données distincts avec  $n = 3$  ayant les caractéristiques suivantes ?

- $\bar{x}_1 = \bar{x}_2$  et  $\tilde{x}_1 \neq \tilde{x}_2$
- $\bar{x}_1 \neq \bar{x}_2$  et  $\tilde{x}_1 = \tilde{x}_2$
- $\bar{x}_1 = \bar{x}_2$  et  $\tilde{x}_1 = \tilde{x}_2$
- $\bar{x}_1 = \bar{x}_2$  et  $s_1 \neq s_2$
- $\bar{x}_1 \neq \bar{x}_2$  et  $s_1 = s_2$
- $\bar{x}_1 = \bar{x}_2$  et  $s_1 = s_2$