

## TP N°2 : Statistique descriptive sur SPSS

### Objectif :

1. **Comment entrer une variable quantitative dans SPSS.**
2. **Comment calculer les paramètres de positions.**
3. **Comment calculer les paramètres de dispersion.**
4. **Comment calculer les paramètres de forme.**
5. **Comment tracer le diagramme en bâton et établit la forme de graph (étalé à droite, à gauche).**
6. **Comment faire la boîte à moustache (la boîte de Tukey ou le box plot).**

**Rappels :** Avant de commencer le travail sur SPSS, on fait petit rappel sur l'asymétrie

- Le premier paramètre de forme c'est le paramètre d'asymétrie qui nous permet de mesurer et conclut si la distribution est symétrique ou non, il existe plusieurs coefficients pour étudier l'asymétrie comme le coefficient de Pearson qui compare la moyenne avec le mode ou la médiane ainsi que la skewness (calculée par SPSS). On suit la règle suivante :
- Si la skewness est négative alors la queue de distribution étalée vers la gauche.
- Si la skewness est positive alors la queue de distribution étalée vers la droite.
- Si la distribution est symétrique alors la skewness est nul.

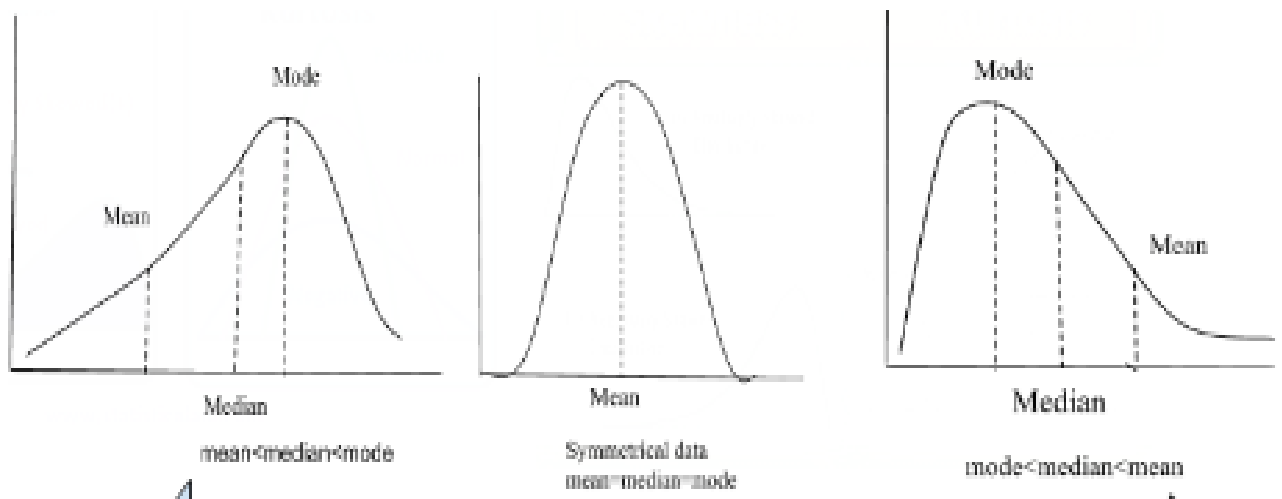
La même chose pour le coefficient de Pearson :

- Si La moyenne = mode = médiane, cela indique une distribution parfaitement symétrique où les données sont uniformément équilibrées des deux côtés de la moyenne.
- Si la moyenne > mode > médiane, cela suggère une distribution positivement asymétrique où la queue du côté droit est plus longue, et la majorité des points de données sont concentrés sur le côté gauche de la

moyenne. Dans ce cas on dit que la queue de la distribution est étalée à droite.

- Si la moyenne < mode < médiane, cela indique une distribution asymétrique négative où la queue du côté gauche est plus longue ou plus grosse, et la majorité des points de données sont concentrés sur le côté droit de la moyenne. Dans ce cas on dit que la queue de la distribution est étalée à gauche.

La figure suivante montre comment les données sont distribuées



**Pour lancer SPSS :** Cliquez sur **Démarrer**, puis **Programmes** puis sur **SPSS Statistics**, ou il est possible de cliquer sur l'icône SPSS sur le bureau de travail. Cliquez sur **Annuler** à la question « Que voulez-vous faire ? ».

Lorsque vous démarrez SPSS, il ouvre la fenêtre principale.

1. **Entrer les données statistiques :** En prend l'exemple suivant :

**Exemple 01 :** On a relevé le nombre de bactéries contenues respectivement dans 20 milieux biologique, où les résultats sont les suivants :

|    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|
| 40 | 42 | 32 | 38 | 40 | 42 | 32 | 38 | 36 | 38 |
| 32 | 36 | 31 | 36 | 38 | 31 | 36 | 38 | 30 | 40 |

### Questions :

- 1- Calculer les paramètres de tendance centrale
- 2- Calculer les paramètres de dispersion.
- 3- Calculer les paramètres de formes
- 4- Présenter le diagramme en bâton et la boîte à moustache de cette série statistique.
- 5- Que peut-on conclure ?

Pour saisir ces données, vous pouvez utiliser la méthode 1 ou 2 que nous avons déjà vue dans le premier TP.

**Etape 1 :** On choisit l'affichage des variables, et on entre le nom de la variable et ses caractéristiques. Car la variable est quantitative on choisit type : numérique et Mesure : Echelle.

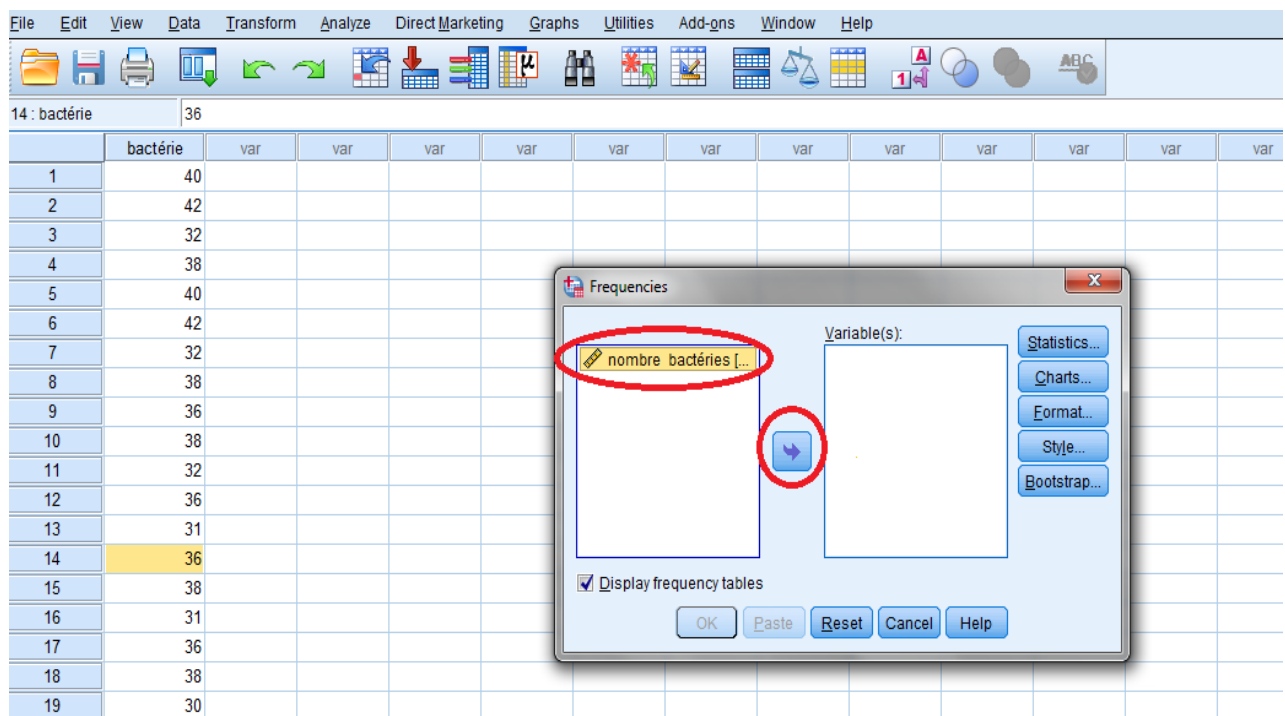
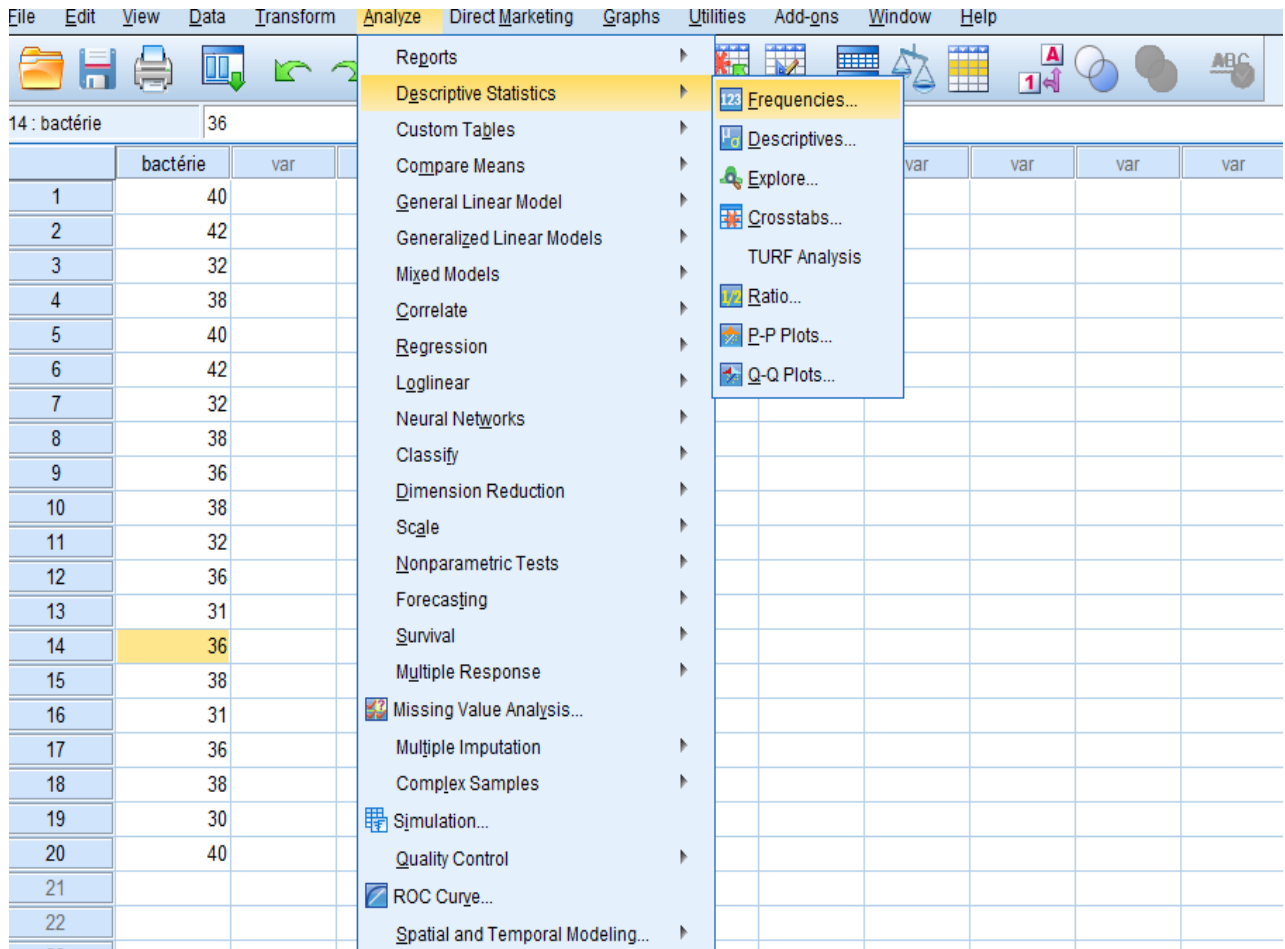


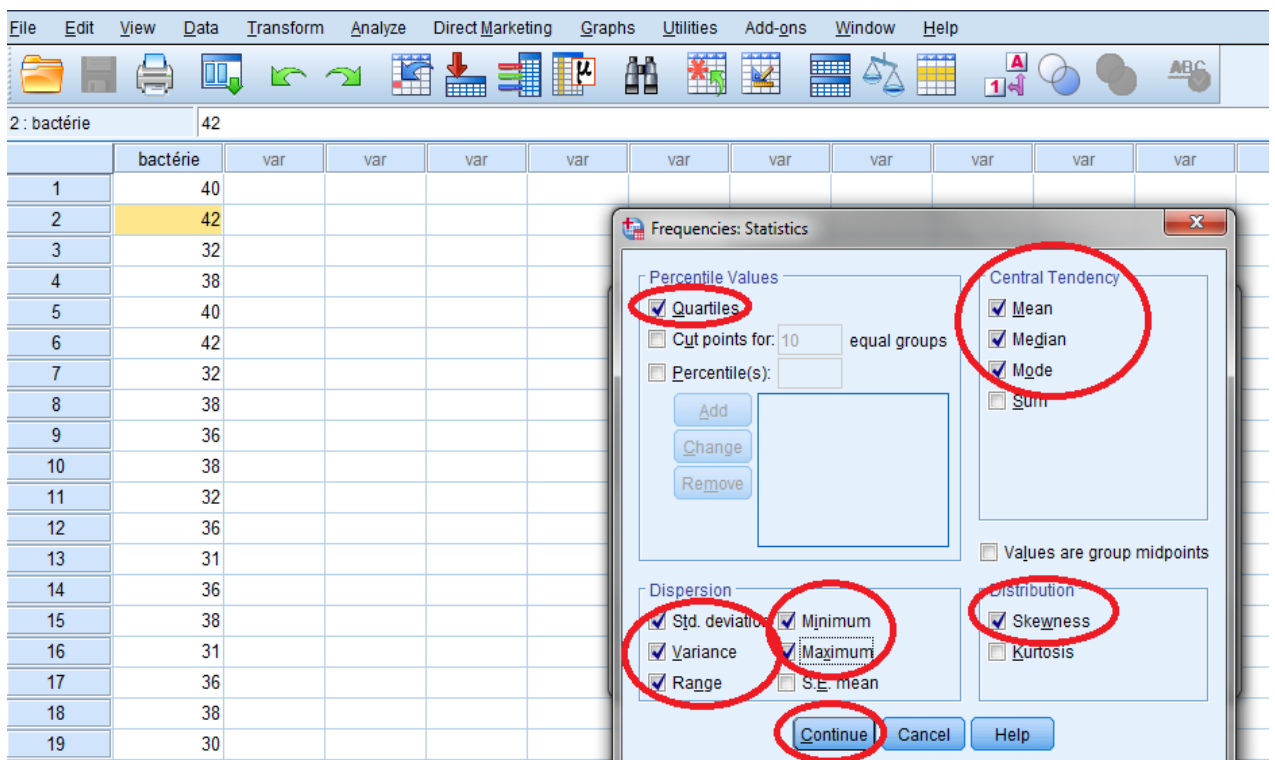
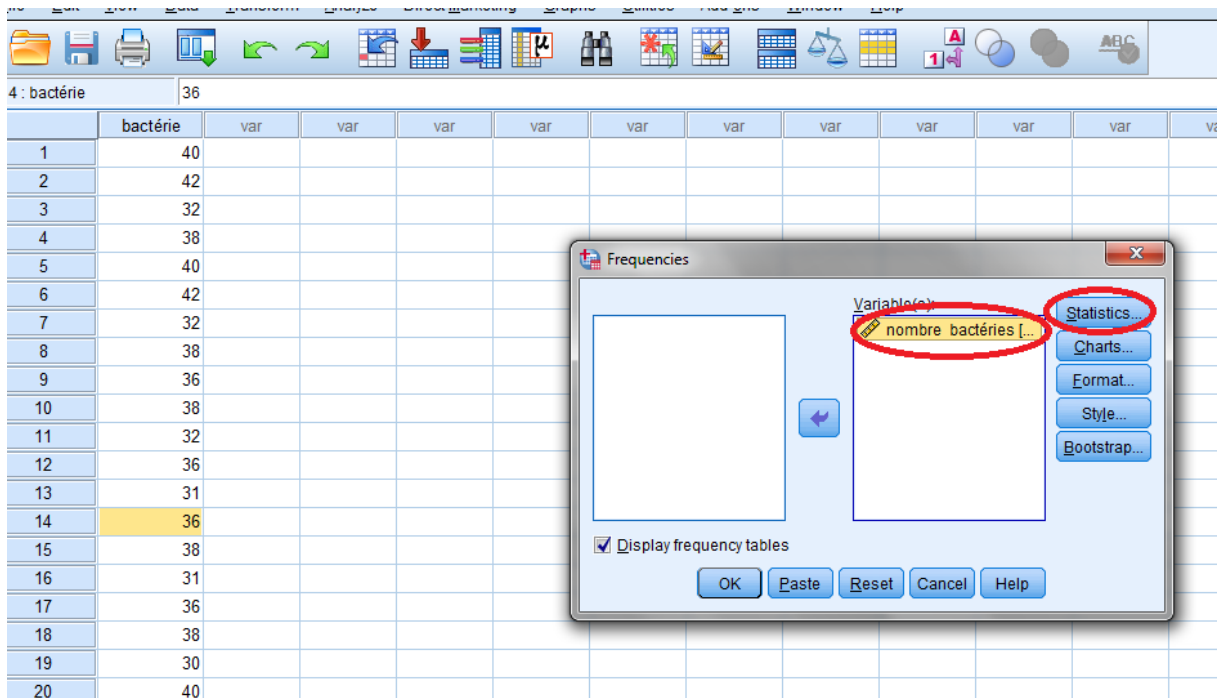
**Etape 2 :** Dans l'affichage des données, on remplit la colonne Bactérie par les modalités de cette variable, comme l'indique la figure suivante :

|    | bactérie | var | var | var | var | var | var | var | var | var | var |
|----|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1  | 40       |     |     |     |     |     |     |     |     |     |     |
| 2  | 42       |     |     |     |     |     |     |     |     |     |     |
| 3  | 32       |     |     |     |     |     |     |     |     |     |     |
| 4  | 38       |     |     |     |     |     |     |     |     |     |     |
| 5  | 40       |     |     |     |     |     |     |     |     |     |     |
| 6  | 42       |     |     |     |     |     |     |     |     |     |     |
| 7  | 32       |     |     |     |     |     |     |     |     |     |     |
| 8  | 38       |     |     |     |     |     |     |     |     |     |     |
| 9  | 36       |     |     |     |     |     |     |     |     |     |     |
| 10 | 38       |     |     |     |     |     |     |     |     |     |     |
| 11 | 32       |     |     |     |     |     |     |     |     |     |     |
| 12 | 36       |     |     |     |     |     |     |     |     |     |     |
| 13 | 31       |     |     |     |     |     |     |     |     |     |     |
| 14 | 36       |     |     |     |     |     |     |     |     |     |     |
| 15 | 38       |     |     |     |     |     |     |     |     |     |     |
| 16 | 31       |     |     |     |     |     |     |     |     |     |     |
| 17 | 36       |     |     |     |     |     |     |     |     |     |     |
| 18 | 38       |     |     |     |     |     |     |     |     |     |     |
| 19 | 30       |     |     |     |     |     |     |     |     |     |     |
| 20 | 40       |     |     |     |     |     |     |     |     |     |     |

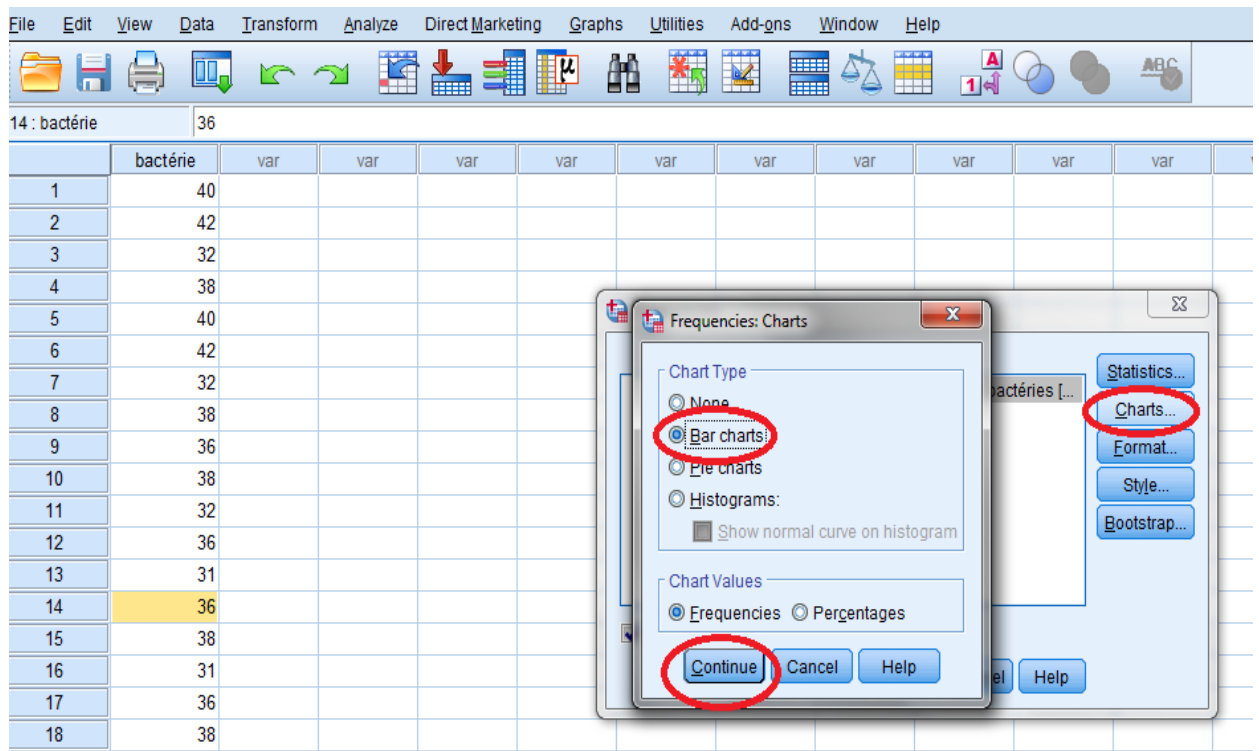
**Etape 3 :** Dans cette étape on calcule les mesures de tendance centrale de dispersion et de forme

1. **Mesures de tendance centrale, de dispersion et de formes :** Après la saisie des données dans SPSS ; on calcule les mesures de position telle que : la moyenne, le mode, la médiane et les quartiles ainsi que les mesures de dispersion telle que : la variance, l'écart type et l'étendu mesures de formes tel que la Skewness (coefficients d'asymétrie) et le Kurtosis (coefficient d'aplatissement). Pour les calculer avec SPSS nous suivons les étapes suivantes :
2. On choisit Analyse dans la barre menu,
3. On choisit Statistique descriptive,
4. On choisit effectif ou fréquence,
5. On choisit Statistique et on coche sur les mesures correspond au mesure de position et dispersion,
6. On clique sur poursuivre puis sur ok.





2. **Diagramme en bâton** : Pour tracer le diagramme à bâton on choisit Analyse puis statistique descriptive puis effectif ou fréquence et on choisit Diagramme ou graphique et on clique sur graphique à barre et ok



**Remarque :** Charts désigne diagramme et Bar charts désigne diagramme en bâton (version anglais)

- Interprétation des résultats :** On obtient une nouvelle fenêtre, c'est la fenêtre des résultats, elle nous donne : les tableaux et le diagramme

### Statistics

nombre bactéries

|                        |         |        |
|------------------------|---------|--------|
| N                      | Valid   | 20     |
|                        | Missing | 0      |
| Mean                   |         | 36.30  |
| Median                 |         | 37.00  |
| Mode                   |         | 38     |
| Std. Deviation         |         | 3.785  |
| Variance               |         | 14.326 |
| Skewness               |         | -.258  |
| Std. Error of Skewness |         | .512   |
| Range                  |         | 12     |
| Minimum                |         | 30     |
| Maximum                |         | 42     |
|                        | 25      | 32.00  |

|        |    |       |
|--------|----|-------|
| Perce- | 50 | 37.00 |
| ntiles | 75 | 39.50 |

La table statistics nous donne :

- **Les paramètres de tendance centrale** où la moyenne (mean) est égale à 36.30, le mode est égal à 38, la médiane est égale à 37 et les quartiles (Percentiles) sont égaux à 32 et 39.50 (Q1 et Q3 respectivement)
- **Les paramètres de dispersion** où la variance est égale à 14.32, l'écart type (Std.Deviation) est égale à 3.785 et l'Etendue (Range) est égale à 12.
- **Les paramètres de formes** où la Skewness est égale à -0, 258.

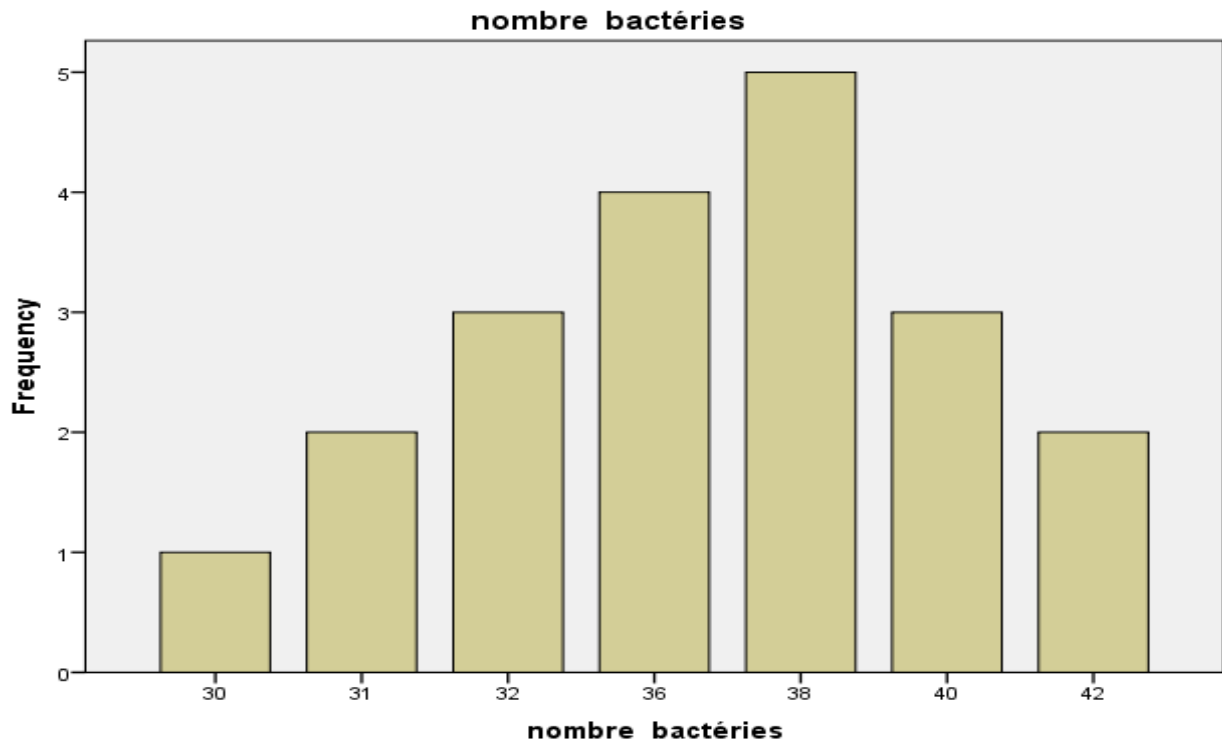
**Conclusion :** A travers ce tableau, nous concluons que la série statistique est asymétrique, car la moyenne est inférieure à la médiane et au mode, ainsi que la Skewness n'est pas 0. De plus, la Skewness est égale à -0, 258 et la moyenne < mode < médiane cela signifie que la queue de la distribution est légèrement plus longue du côté gauche et que la plupart des bactéries sont concentrés du côté droit de la moyenne. .

|       | Fre-   | Per-  | Valid Per- | Cumulative |
|-------|--------|-------|------------|------------|
|       | quency | cent  | cent       | Percent    |
| Valid | 30     | 1     | 5.0        | 5.0        |
| d     | 31     | 2     | 10.0       | 15.0       |
|       | 32     | 3     | 15.0       | 30.0       |
|       | 36     | 4     | 20.0       | 50.0       |
|       | 38     | 5     | 25.0       | 75.0       |
|       | 40     | 3     | 15.0       | 90.0       |
|       | 42     | 2     | 10.0       | 100.0      |
| Total | 20     | 100.0 | 100.0      |            |

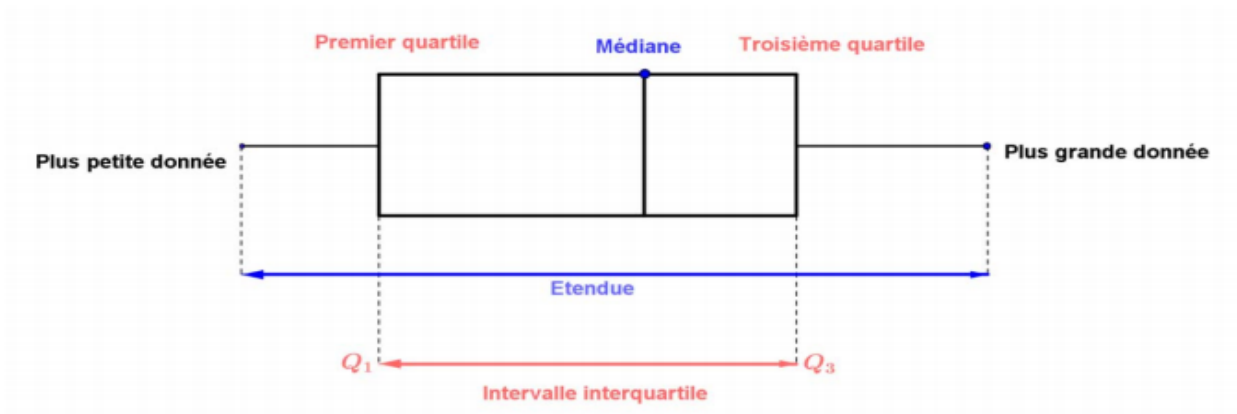
- Ce tableau représente un tableau des effectifs (frequency) en plus les fréquences fois 100 et les fréquences cumulés croissant.
- **Diagramme en bâton :** La figure suivante présente le diagramme en Bâton.



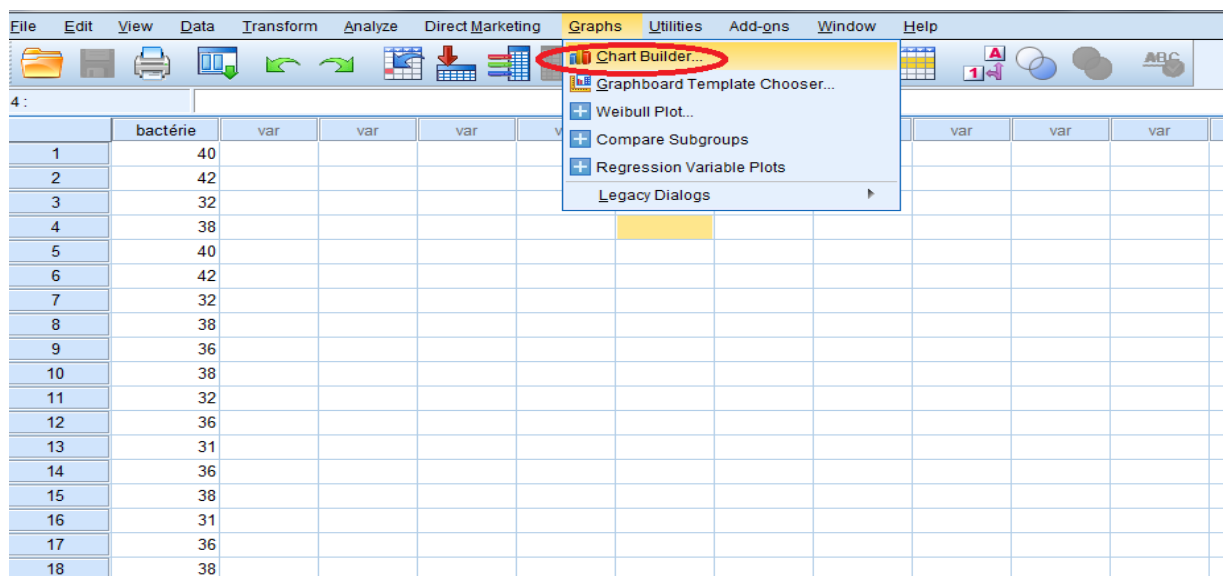
On peut voir que la distribution n'est pas symétrique.

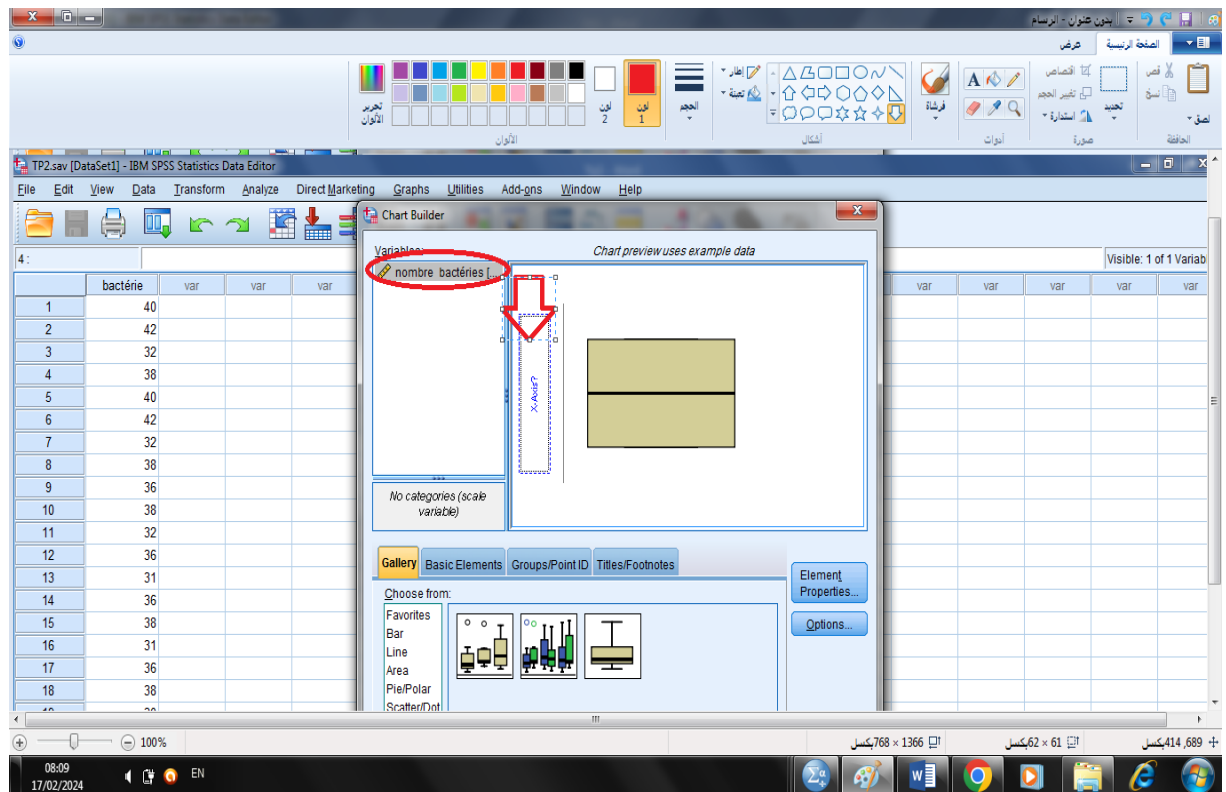
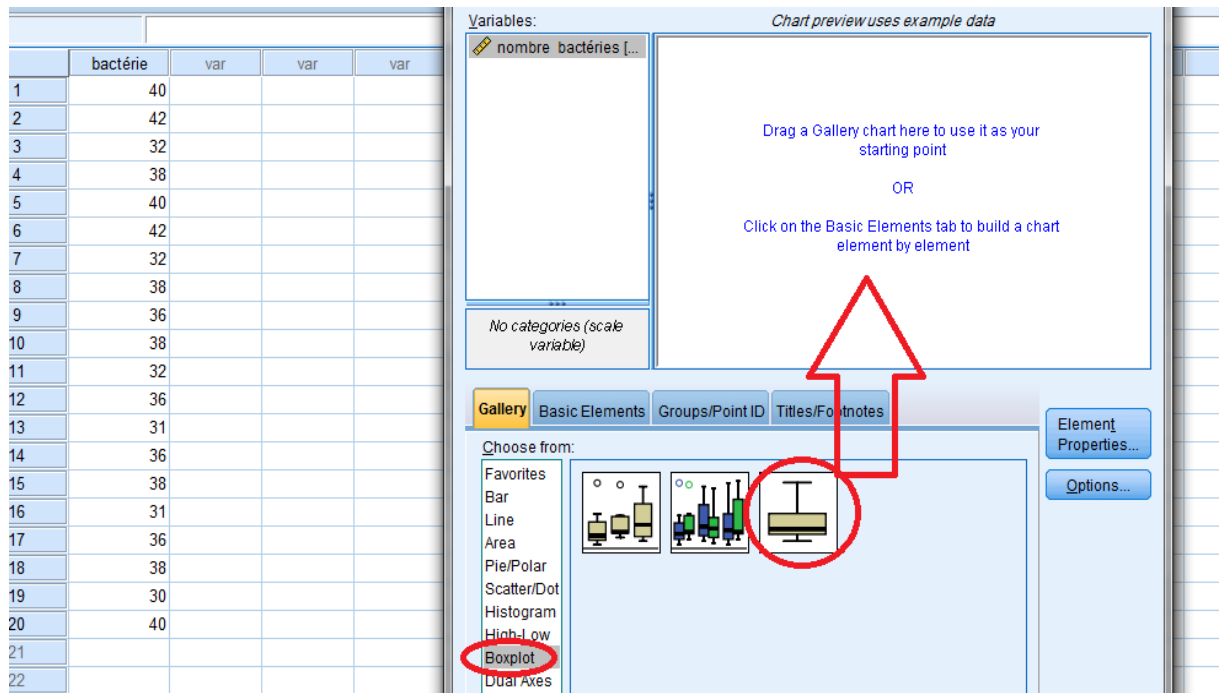


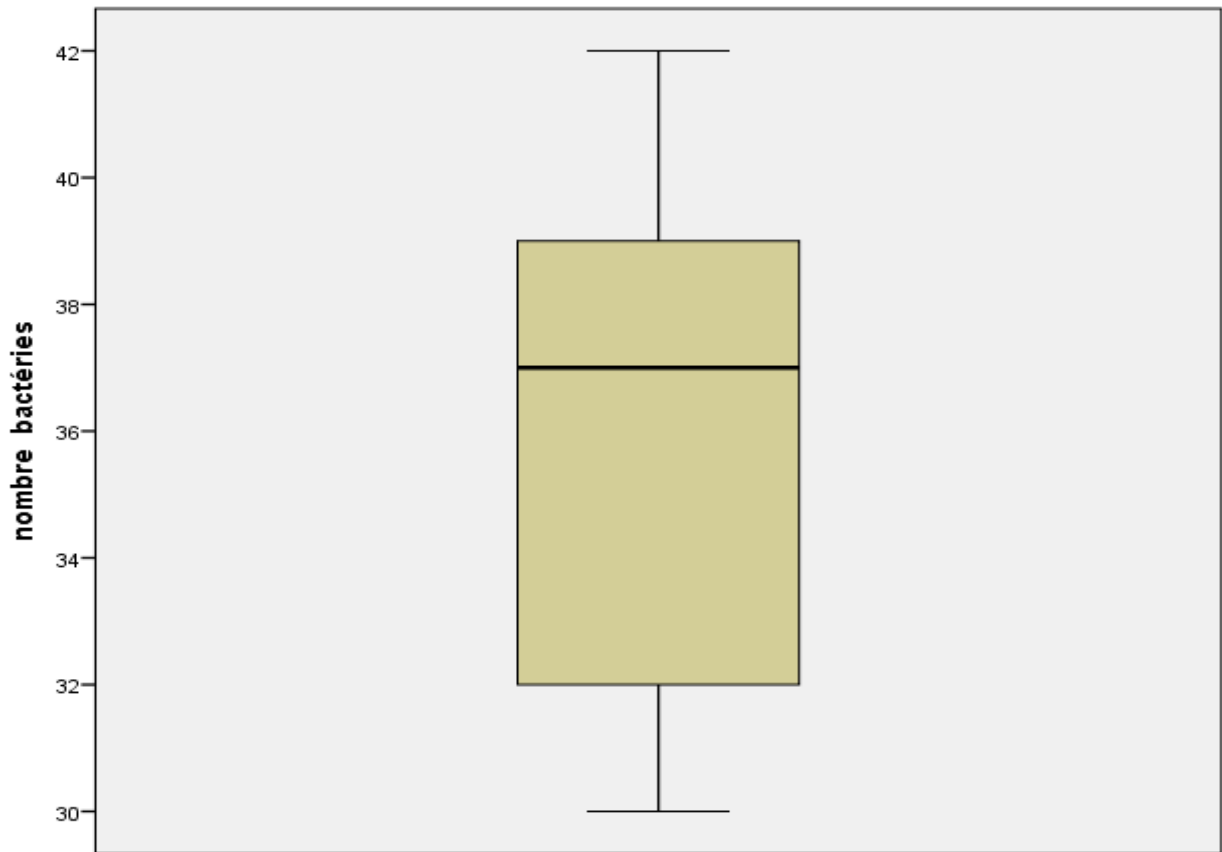
- **boîte à moustache ou boîte de Tukey** : La figure ci-dessus présente le graphique de boîte à moustache. Cette boîte nous permet d'établir l'asymétrie des données statistiques graphiquement.
- La ligne centrale dans la boîte indique la médiane des données. La moitié des données est supérieure à cette valeur et l'autre moitié est inférieure. Si les données sont symétriques, la médiane sera au centre de la boîte. Si les données sont asymétriques, la médiane sera plus proche du haut ou du bas de la boîte.



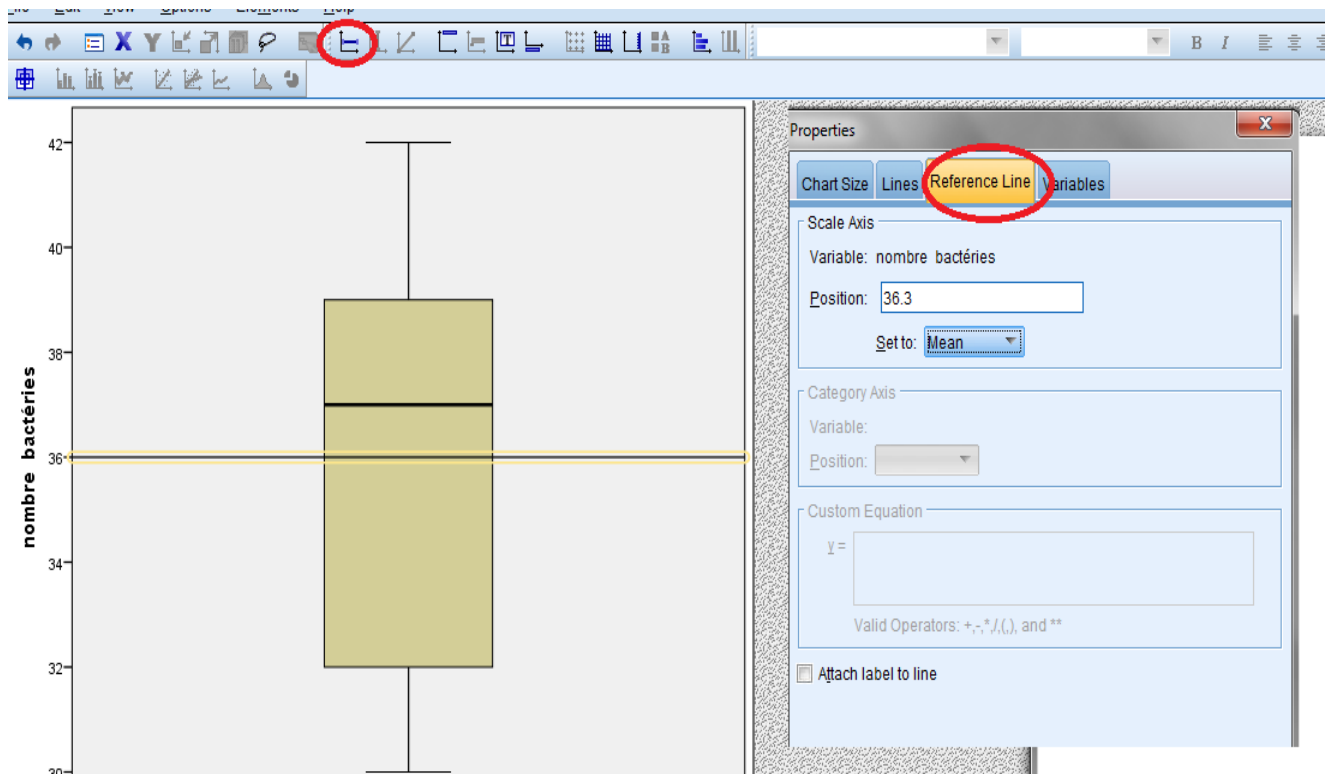
Pour présenter cette boîte dans SPSS, on choisit Graphiques puis Générateur de graphiques puis on sélectionne dans la boîte des graphes boîte à moustache et on clique sur ok.

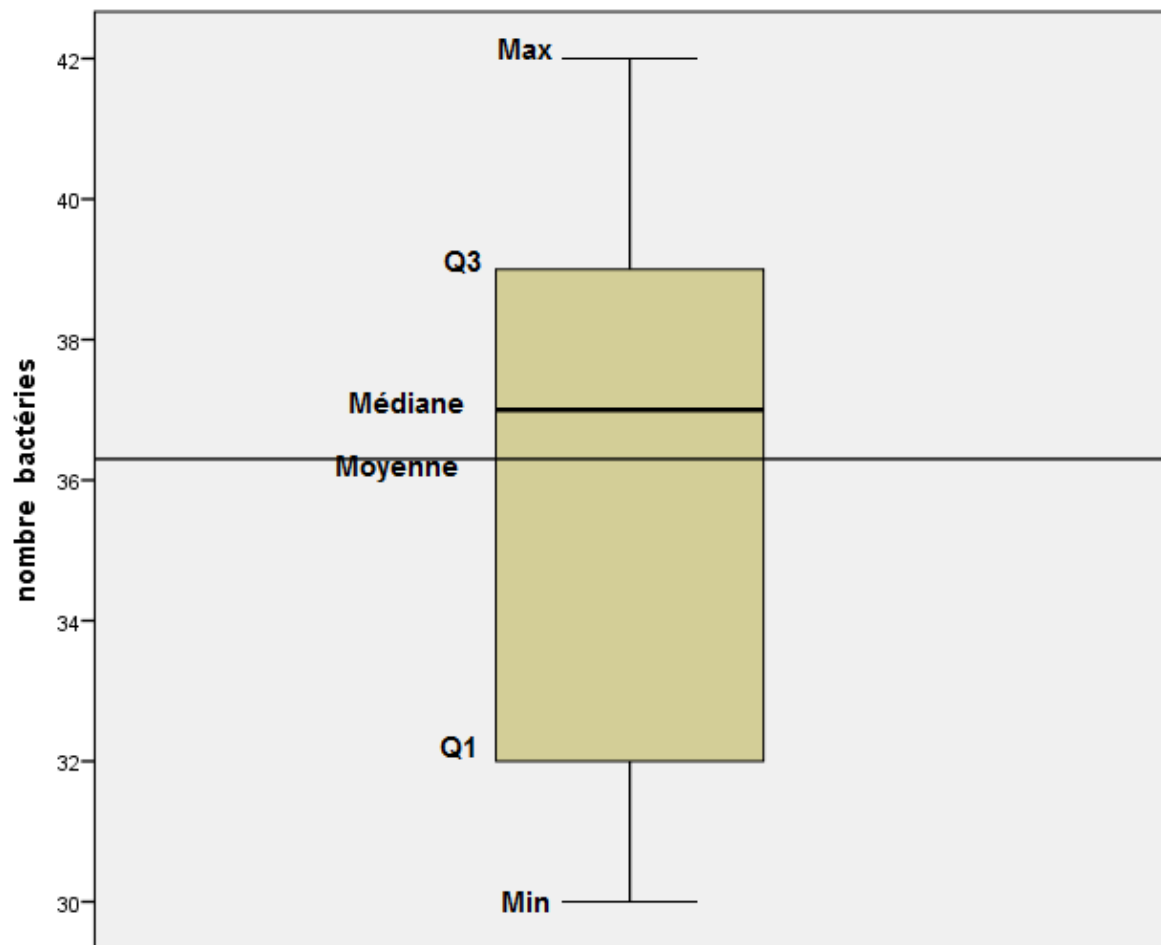






Si on clique double fois sur le graph on obtient une fenêtre graphique pour ajouter des options au graphe, comme exemple pour ajouter la moyenne.





La distribution est clairement asymétrique car la médiane n'est pas au milieu de la boîte, mais plutôt plus près du haut de la boîte.

