

Auteur : Raymond Moché

**But de l'activité :** Comprendre la différence entre la notion de moyenne et de médiane ; approfondir la notion de quartile ; constater l'insensibilité de la médiane et des quartiles aux variations des valeurs extrêmes ; constater l'effet de la multiplication des données par une constante.

### Compétences engagées :

revoir la notion de moyenne d'une série statistique, illustrer de façon concrète les notions de médiane, premier et troisième quartiles, étendue.

### Pré-requis :

- ✓ initiation au tableur : insérer et dérouler une formule, lire et interpréter des résultats ; connaître les fonctions de tri (« Données > Trier ») et « MOYENNE »

### Matériels utilisés :

- ✓ Classe informatique.

**Durée indicative :** Une heure.

### Nom des logiciels utilisés :

- ✓ Suite bureautique OOo (Calc)

**Documents utiles à télécharger :** « Fiche Élève » et « Fiche Professeur », classeur « Calculs ».

✓

### À propos des pré-requis :

Cette activité est rédigée en supposant que

- les élèves ont été initiés aux notions d'étendue, de médiane, de premier et troisième quartiles d'une série statistique, dont nous rappelons les définitions ci-dessous ;
- ils connaissent la fonction de tri (« Données>Trier ») et la fonction MOYENNE des tableurs (sinon, ils devront consulter l'« assistant des fonctions »).
- L'usage du tableur est indispensable parce que la série statistique étudiée est longue (423 termes) conformément aux recommandations du Projet de document d'accompagnement du programme.
- Des difficultés de vocabulaire de français standard et de mathématiques sont à prévoir (tableau 2 notamment).

### Comment on détermine la médiane et les premier et troisième quartiles :

- On ordonne la série statistique observée dans l'ordre croissant. Si elle est de taille  $2n+1$ , la médiane est la valeur du terme de rang  $n+1$  de la série ordonnée (terme du milieu) ; si elle est de taille  $2n$ , la médiane est la demi-somme des termes de rang  $n$  et  $n+1$  (les deux termes du milieu) de la série ordonnée.
- Le premier quartile est le plus petit élément  $q$  de la série tel qu'au moins 25% des données soient inférieures ou égales à  $q$ . Cela signifie, en clair, que si la taille de la série statistique observée est  $4n$  (respectivement  $4n+1$ ,  $4n+2$ ,  $4n+3$ ), c'est la valeur du terme de rang  $n$  (respectivement  $n+1$ ,  $n+1$ ,  $n+1$ ) de la série ordonnée (en gros, un quart des termes de la série ordonnée sont à gauche du premier quartile, trois quarts à droite).
- Le troisième quartile est le plus petit élément  $q$  de la série tel qu'au moins 75% des données soient inférieures ou égales à  $q$ . Cela signifie, en clair, que si la taille de la série statistique observée est  $4n$  (respectivement  $4n+1$ ,  $4n+2$ ,  $4n+3$ ), c'est la valeur du terme de rang  $3n$  (respectivement  $3n+1$ ,  $3n+2$ ,

3n+3) de la série ordonnée.

### **Remarque importante : pourquoi il ne faut pas utiliser la fonction « QUARTILE » du tableur**

La feuille « Professeur » du classeur « Calculs Professeur » contient le calcul de la médiane (S12) et des premier et troisième quartiles (S13:S14) à l'aide du tableur (voir les fonctions MEDIANE et QUARTILE dans l'assistant des fonctions), ce qui n'est pas demandé aux élèves. Il est très tentant d'utiliser ces fonctions parce qu'elles rendent inutile d'ordonner la série statistique A1:A423 (tout en masquant aux élèves la signification de ces notions). *En fait, la fonction MEDIANE du tableur peut être utilisée car elle utilise la même définition de la médiane que nos programmes. Par contre, il ne faut pas utiliser la fonction QUARTILE du tableur parce qu'elle est basée sur une définition des quartiles qui n'est pas celle de nos programmes. On constate par exemple que les résultats des plages L13:M13 et L14:M14 sont différents. Cela n'est pas alarmant. En fait, les définitions du tableur d'une part, des programmes d'enseignement d'autre part, traduisent correctement la même idée. Mais pour éviter d'employer la bonne définition, qui provient de la « fonction quantile » et qui est très technique et incompréhensible pour les élèves, on fait des choix simplificateurs, autrement dit, on adopte des conventions qui varient d'un pays à l'autre. C'est l'explication du mystère.*

**Remerciements** au Collègue anonyme dont le message m'a convaincu d'améliorer la rédaction de cette activité.

### **Déroulement de la séance :**

- Les élèves remplissent les tableaux de la « Fiche Élève » au fur et à mesure qu'ils font les calculs de la feuille « Élève » du classeur « Calculs Élève ».
- Le tableau 1 se déduit de la statistique ordonnée de la colonne B, en appliquant les définitions des notions concernées. On doit déterminer à la main les cellules contenant la médiane et les premier et troisième quartiles. Bien sûr, *on manque d'une macro* qui donnerait ces quantités à partir de la série statistique brute ou, mieux, à partir de la série statistique ordonnée (car il paraît indispensable pour la bonne compréhension des choses, d'ordonner toujours la série statistique proposée).
- Les effectifs du tableau 2 se déduisent de l'observation de la colonne B. Les pourcentages sont à rapprocher des définitions des caractéristiques de position et de dispersion utilisées.
- La question qui mène au tableau 3 est intéressante : tous les nombres sont multipliés par 1,0143. Ce n'est pas facile.
- La question 5 est très intéressante et demande de la réflexion. Elle est plutôt difficile. L'activité est entièrement traitée à la feuille « Professeur » du classeur « Calculs ».
- L'activité peut être arrêtée à la fin de la première ou de la deuxième partie. Ce serait dommage. En cas de manque de temps, il vaut mieux sacrifier le tableau 2.