

**Ojectif :** Utiliser un tableur pour le calcul des paramètres d'une série statistique et la représentation graphique de cette série.

Nous allons nous intéresser au nombre d'habitant dans les communes du département des Alpes de Haute Provence à partir de données fournies par l'INSEE <sup>1</sup> à la suite du recensement de la population française effectuée en 2007. Les données brutes seront être téléchargées sur le site de l'INSEE à l'adresse suivante :  
<http://www.insee.fr/fr/ppp/bases-de-donnees/recensement/populations-legales/france-departements.asp>

**1. Préparation du document de travail** (logiciel *OpenOffice.org*)

- a) Observer le fichier fourni par l'INSEE. Comment les communes sont-elles triées ?
- b) Créer un nouveau document de type classeur, dans lequel seront copiés le nom des communes du département et la population correspondante. Enregistrer votre document sous le nom : communes.ods.
- c) Sélectionner les cellules A2 à A201 puis dans la barre des menus Insertion-Noms-Définir : Communes.
- d) De même, sélectionner les cellules B2 à B201 puis dans la barre des menus Insertion-Noms-Définir : Population.

**2. Caractéristiques de la série**

On cherche la population minimale, la population maximale, l'étendue des populations, la population moyenne, la population médiane, le premier quartile, le troisième quartile et l'écart interquartile.

Toute formule de calcul commence par le signe =.

Par exemple, pour obtenir le premier quartile de la plage de données nommée Population, saisir la formule =QUARTILE(Population;1).

Donner la formule, utilisant les fonctions MIN, MAX, MOYENNE, MEDIANE et QUARTILE, dans les cellules F2, F3, F4, F5, F6, F7, F8, F9 pour le calcul correspondant.

E	F
<b>Caractéristiques des populations</b>	
population minimale	
population maximale	
étendue des populations	
Population moyenne	
premier quartile	
population médiane	
troisième quartile	
intervalle inter-quartile	

**3. Regroupement de la série par classes**

On voudrait représenter cette série par un diagramme pour mieux interpréter les données.

- a) Sélectionner les cellules A1 à B201, puis cliquer sur Diagramme et choisir l'un des diagrammes proposés par le tableur pour représenter les données. Cette représentation vous semble-t-elle satisfaisante ?
- b) On décide donc de regrouper les données en classes d'amplitude 2500, depuis [0;2500[ jusqu'à [20000;22500[.

- En I3, saisir 0 puis en I4 saisir =I3+2500 puis tirer vers le bas, avec la poignée de recopie, jusqu'en I11.
- Faire de même pour la plage J3 :J11.
- En K3, saisir ="["&I3&" ;"&J3&"[" et recopier jusqu'en K11.

I	J	K	L	M
<b>Regroupement par classes d'amplitude 2500</b>				
	Classes		Effectif	Centre de classe
	0	2500	[0;2500[	
	2500	5000	[2500;5000[	
	5000	7500	[5000;7500[	
	7500	10000	[7500;10000[	
	10000	12500	[10000;12500[	
	12500	15000	[12500;15000[	
	15000	17500	[15000;17500[	
	17500	20000	[17500;20000[	
	20000	22500	[20000;22500[	

On obtient le résultat ci-contre.

Pour compléter les effectifs, on utilise la fonction FREQUENCE :

- En L3, saisir =FREQUENCE(Population;J3-1), formule qui donne le nombre de communes de population inférieure ou égale à 2499; donc strictement inférieure à 2500.
- En L4, on compte les communes de population strictement inférieure à 5000 et on ôte celles de population strictement inférieure à 2500.

Saisir la formule =FREQUENCE(Population;J4-1)-FREQUENCE(Population;I4-1) et recopier jusqu'en L11.

- c) Représenter maintenant ces données à l'aide d'un diagramme? Cette représentation vous semble-t-elle plus satisfaisante ?
- d) Faire un choix plus judicieux sur le choix des classes (modifier uniquement les cellules I3 à J11).
- e) Si les classes ne sont pas de même amplitude, les diagrammes proposées sont-ils corrects ?

**4. Calcul d'une moyenne pondérée**

On veut calculer la moyenne de cette nouvelle série :

- En M3, on donne la formule pour le centre de la première classe et on la recopie jusqu'en M11 :  $=(I3+J3)/2$ .
- En L12, on donne la formule pour la moyenne de cette nouvelle série :  

$$\frac{\text{somme}(\text{effectifs} \times \text{centre})}{\text{somme}(\text{effectifs})} : =\text{SOMMEPROD}(L3 :L11 ;M3 :M11)/\text{SOMME}(L3 :L11)$$

Comparer les cellules L12 et F5.

**Remarque** : quand on regroupe une série en classes, on n'obtient qu'une valeur approchée de la moyenne. Pour que l'approximation soit bonne, il faut que les classes soient d'amplitude à peu près égales et que la répartition à l'intérieur de chaque classe soit à peu près uniforme.

## 5. Construction d'un histogramme

Un histogramme doit être constitué de rectangles dont la base est proportionnelle à l'amplitude de la classe et l'aire proportionnelle à l'effectif. Le logiciel *OpenOffice.org* ne permet pas la création de manière simple de ce type de diagramme. Nous allons utiliser un autre logiciel très utile en mathématiques : *Geogebra*.

Après avoir lancé ce logiciel :

- Choisir Affichage-Tableur.
- Recopier la plage I3 :L11 de *OpenOffice.org* dans la cellule A1 de *Geogebra*.
- Étant donné qu'il n'y a que 5 communes dont la population dépasse 5000, nous n'allons construire l'histogramme que pour les communes dont la population est inférieure à 5000. Pour avoir la hauteur de l'histogramme de façon à ce que l'aire de chaque surface soit proportionnelle à l'effectif, entrer  $=D1/(B1-A1)$  dans la cellule E1, puis recopier vers le bas cette formule.
- Pour fabriquer l'histogramme, entrer `Histogramme[A1 :A7,E1 :E6]` dans la zone Saisie (la cellule A7 doit contenir la valeur 5000).
- Afin de voir le résultat, dans Options-Graphique, puis dans l'onglet axeX choisir min : -100 et max : 5200 et dans l'onglet axeY choisir min : -0.01 et max : 0.5 et décocher Afficher.

Un résultat possible :

