

TP 5. Statistique exploratoire

Objectif : Trier des données et construire des graphiques avec un tableur, pour interpréter.

Ouvrir le fichier disponible sur le site contenant les statistiques d'immatriculations de véhicules neufs dans les départements de France métropolitaine de 1993 à 2007 (Source : fichier central des automobiles).

A. Extraction des immatriculations de voitures particulières

On souhaite extraire de ce fichier la série TR0101 relative aux voitures particulières et commerciales. Sélectionner la colonne C et créer un filtre, puis choisir la série visée. Sélectionner les données et les copier/coller sur une nouvelle feuille.

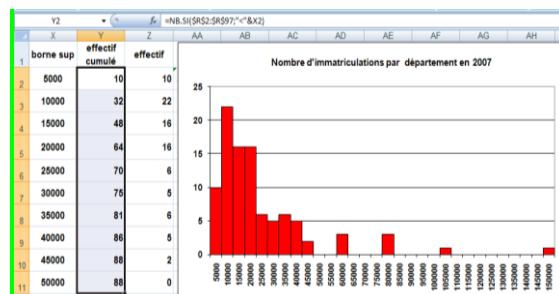
Aide Tableur

Pour créer un filtre sur une colonne comportant plusieurs noms de séries,
 – choisir dans l'onglet « Données » : Filtre / Filtre Automatique ou Filtrer ou AutoFiltre.
 – cocher le nom de la série désirée.

1. a. À la ligne 100, calculer, pour chaque année, le nombre total d'immatriculations de véhicules particuliers et commerciaux en France métropolitaine.
- b. Quel a été le nombre moyen annuel d'immatriculations de 1993 à 2007 ?
- c. Représenter l'évolution du nombre total d'immatriculations (on choisira avec Excel : Nuage de points reliés par une courbe, et avec OpenOffice Calc : XY(dispersion) Points et lignes). Décrire cette évolution.
2. Utiliser le même type de diagramme pour représenter l'évolution des immatriculations dans les différents départements. Quels sont les départements qui, pour une période donnée que l'on indiquera, ont délivré le plus grand nombre d'immatriculations ?

B. Répartition des immatriculations en 2007

1. Indiquer dans combien de départements le nombre des immatriculations en 2007 a été inférieur à 5000.



2. On veut construire un histogramme montrant comment se répartissent les départements selon le nombre de leurs immatriculations en 2007, par classes de longueur 5000.

- a. Entrer en colonne X, comme sur l'image d'écran, les bornes supérieures des classes, de 5000 à 15000.
- b. Dans la cellule Y2, entrer une formule, utilisant la fonction NB.SI, donnant le nombre de valeurs de la colonne R qui sont inférieures à la valeur figurant dans la cellule X2. Que retrouve-ton ici ?
- c. Recopier cette formule jusqu'en Y31. Que signifie le nombre « 81 » apparu dans la cellule Y8 ?
- d. En déduire, en colonne Z, les effectifs de toutes les classes de [0,5000] à [145000 ;150000].
- e. Construire maintenant l'histogramme sur la même feuille de calcul du tableau.

3. Quelle particularité présente cet histogramme ? D'après son allure, quel indicateur représente au mieux la tendance centrale de cette série : la moyenne ou la médiane ? Expliquer pourquoi. Sans faire de calcul, sauriez vous dire lequel de ces deux indicateurs sera le plus grand ? Vérifier !