

Ex 1 : Filtrage & Tri des données d'un fichier Tableur CSV

1. a) Créer un dossier de travail nommé « TP8_Données_SNT »
b) Télécharger le fichier « [Les Prénoms attribués depuis 2000](#) » (on choisira le fichier CSV – 3 Mo - allégé) et enregistrer ce fichier dans votre dossier
2. Ouvrez le fichier au format .ZIP ainsi obtenu, et extraire le fichier **Dpt27depuis2000.csv** puis enregistrer sous « **TP8_Données.ods** »
3. a) Recopier toutes les données des prénoms des garçons dans une nouvelle feuille nommée « **garçons** » (supprimer les « prénoms rares »)
b) Recopier toutes les données des prénoms des filles dans une nouvelle feuille nommée « **filles** » (supprimer les « prénoms rares »)
4. a) Créez les tableaux suivants indiquant le nombre de prénoms (garçons) en fonction de la 1ère lettre et insérer la formule
=NB.SI(\$B\$2:\$B\$3690;"A")
en cellule H1
b) Quel est le rôle des symboles « \$ » ?
c) Effectuer une recopie de cette formule jusqu'en cellule H26 et compléter la somme de la colonne H
d) Effectuer la même démarche pour les prénoms des filles dans la dernière feuille du classeur
e) Existe-t-il des « premières lettres » non utilisées pour les prénoms de garçons et de filles ?
5. a) Effectuer une mise en forme des colonnes G et H
b) Construire un **diagramme** en barres (verticales) permettant de visualiser les fréquences des différentes lettres des prénoms des filles & garçons

| | F | G | H |
|---|---|---|---|
| 1 | | A | |
| 2 | | B | |
| 3 | | C | |
| 4 | | D | |
| 5 | | E | |
| 6 | | F | |

Ex 2 : Filtrage & Tri des données d'un fichier CSV avec PYTHON

- 1) a) Créez le script Python ci-dessous & enregistrer sous « **prenoms.py** »

```

1 import pandas
2 import matplotlib.pyplot as plt
3 #-----
4 fichier = pandas.read_csv("Dpt27depuis2000.csv", sep=";")
5 prenoms = fichier.loc[fichier["preusuel"] != "_PRENOMS_RARES"]
6 print(prenoms)
7 print("-----")

```

- b) Exécuter le script et expliquer le résultat obtenu
- c) Expliquer le rôle de chaque ligne de ce script

- 2) a) Modifier le script comme dans l'encadré ci-dessous
b) Exécuter le script et expliquer le résultat obtenu

```

8 recherche1 = prenoms.loc[(prenoms["annais"] == 2023), :]
9 print(recherche1)
10 print("-----")

```

(Rque : on pourra déposer le symbole « # » en début de la ligne 6)

- 3) a) Modifier le script comme dans l'encadré ci-dessous
b) Exécuter le script et expliquer le résultat obtenu

```

11 recherche2 = prenoms.loc[(prenoms["preusuel"] == "HUGO"), :]
12 print(recherche2)
13 print("-----")

```

- 4) a) Modifier le script comme dans l'encadré ci-dessous
b) Exécuter le script et expliquer le résultat obtenu

```

14 recherche3 = prenoms.loc[(prenoms["sexe"] == 1), :]
15 print(recherche3)
16 print("-----")

```

- 5) a) Modifier le script comme dans l'encadré ci-dessous
b) Exécuter le script et expliquer le résultat obtenu

```

17 recherche4 = prenoms.loc[(prenoms['sexe'] == 2)
                            & (prenoms['preusuel'] == "CAMILLE"), :]
18 print(recherche4)
19 tri = recherche4.sort_values(by="nombre")
20 print(tri)
21 print("-----")

```

- 6) a) Modifier le script comme dans l'encadré ci-dessous
b) Exécuter le script et expliquer le résultat obtenu

```

23 recherche5 = prenoms.loc[(prenoms['preusuel'] == "LOUISE")
                            & (prenoms['sexe'] == 2), :]
24 recherche5.plot(x="annais", y="nombre")
25 plt.show()

```

c) Modifier le script ci-dessus afin d'observer l'évolution d'apparition de **votre prénom** depuis l'an 2000

- 7) **BONUS :** Modifier le script PYTHON afin de créer le diagramme en barres des fréquences des premières lettres des prénoms usuels