

TP*1 Traitement de données

L'objectif du TP est de traiter les données pour confirmer ou infirmer si oui ou non la population à Auxerre ou dans l'Yonne est en chute libre depuis ces dernières années.

ÉTAPE NUMÉRO 1 :

Importer le fichier .csv :

Dans le TD1 de la saé 1.05 on a déjà le code disponible nous permettant d'importer les données sur notre terminal, nous allons juste l'importer :

```
import csv
table=[]
with open('donnees_2008.csv',newline=" ") as csvfile:
    reader=csv.reader(csvfile,delimiter=',')
    for row in reader:
        table.append(row)
```

J'ai choisi de prendre comme exemple le fichier "donnees_2008.csv" mais cela aurait pu fonctionner avec n'importe quelle autre fichier csv.

Extraire et isoler les informations qui nous intéresse dans ce fichier .csv :

Dans le fichier 'Donnees_2008.csv' nous avons différentes données dans notre table dont beaucoup qui ne serviront jamais n'a rien et qui empêche la lisibilité de nos données, nous allons donc filtrer et choisir ce que nous allons garder comme le Code de la région, Le nom de la région et le nom de la Commune et la Population Totale.
(Dans le screen je n'ai pas rajouté row[6] le nom de la Commune et row[2] le Code de département mais je l'ai fait après)

```

données_importantes = []
for row in table[1:]:
    données_importantes.append([row[0], row[1], row[9]])
print(données_importantes)

```

Cela nous permettra donc d'obtenir le code ci-dessus :

J'ai importé une nouvelle table vide intitulé "données_importantes" et nous allons la remplir avec les données qui nous manque, pour faire cela dans un premier temps on retire la première ligne qui ne sert à rien et on sélectionne avec row nos tables les plus importantes données, il ne reste plus qu'à print pour regarder le résultat.

Enfin, nous arrivons à la dernière et unique étape, mettre une condition afin qu'elle n'affiche qu'uniquement une ville ou un code de département comme ci-dessous :

```

données_importantes = []
for row in table[1:]:
    if row[6] == 'Auxerre':
        données_importantes.append([row[1], row[2], row[6], row[9]])
print(données_importantes)

```

Additionner les différentes données pour pouvoir constater d'une évolution plus tard :

La dernière étape que j'ai à faire avec le fichier 'Donnees_2008.csv' est d'ajouter toute la population totale du département du 89, pour cela j'ai besoin de code python ci-dessous :

```

total_population = 0
for row in table[1:]:
    if row[2] == '89':
        données_importantes.append([row[1], row[2], row[6], row[9]])
        if row[9].isdigit():
            total_population += int(row[9])
print(total_population)

```

Voici comment j'ai fait, j'ai créé une nouvelle table variable total_population, j'ai utilisé un tout nouveau module qui permet de transformer une chaîne en int qui est .isdigit() et j'ai additionner les int(row[9]) avec lui-même à chaque nouvelle table.

```
me831169@aux-r2inf2-015:~/Documents/traitement de donnees$ python3 tp1.py  
353611
```

Voici le nombre d'habitants totale du 89 en 2008 ci-dessus :

Le nombre d'habitants totale du 89 en 2016 est de 350 970

Le nombre d'habitants totale du 89 en 2021 est de 344 022

Le nombre d'habitants totale du 89 en 2023 est de 341 738

```
me831169@aux-r2inf2-015:~/Documents/traitement de donnees$ python3 tp1.py  
[353611, 350970, 344022, 341738]
```

A partir de cela nous pouvons nous rendre compte que la population totale d'habitants à diminuer de 12.000 personnes dans le 89.

ÉTAPE NUMÉRO 2 :

Faire un graphique pour constater de la diminution :

Au préalable je stock tous mes résultats dans une liste que je vais nommer population évolution, et j'importe mon module matplotlib, je réserve l'abscisse x pour les années et l'abscisse Y pour l'évolution de la population et j'ai ceci :

```

total_population = int(row[10])

population_evolution.append(total_population)
population_evolution.append(total_population2)
population_evolution.append(total_population3)
population_evolution.append(total_population4)

print(population_evolution)

année = [2008, 2016, 2021, 2023]

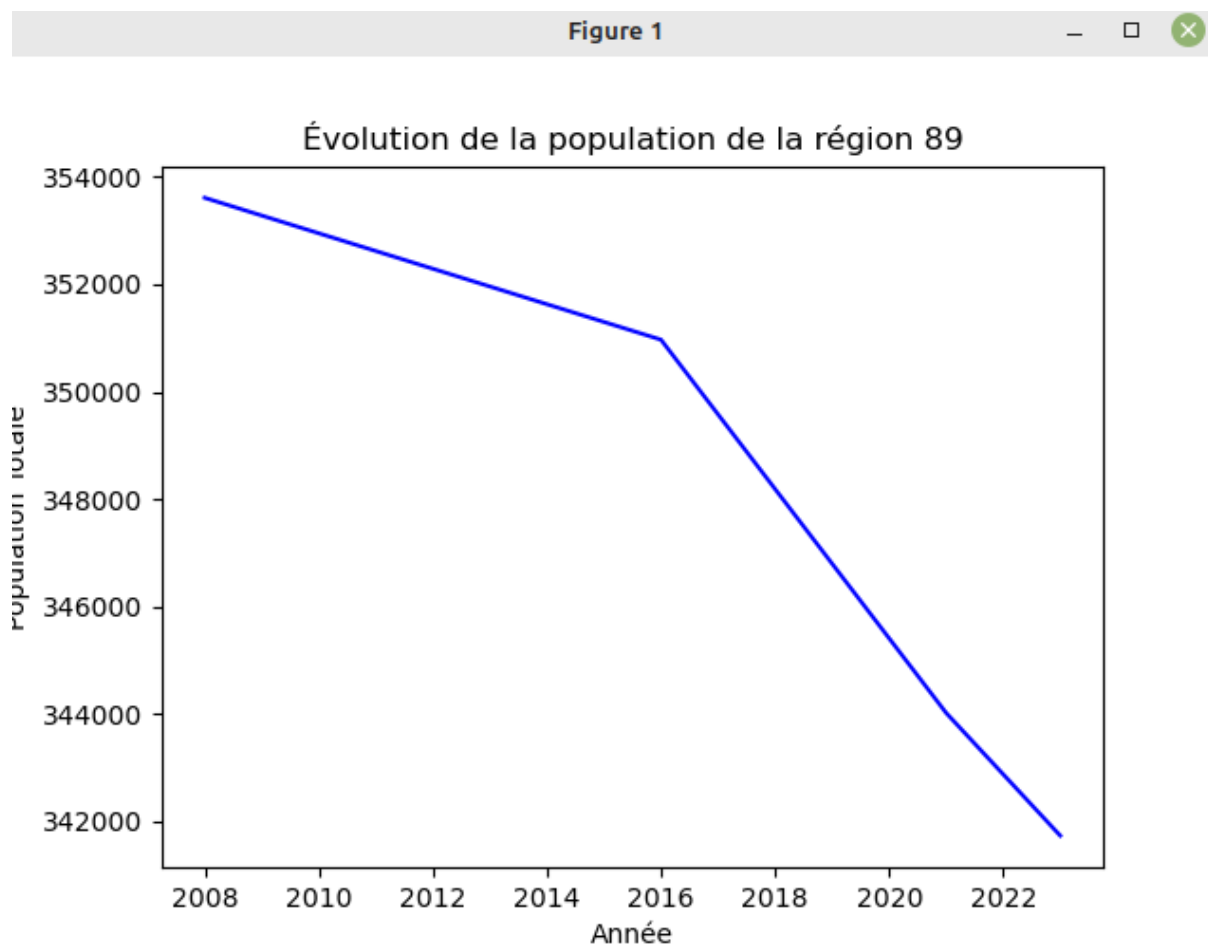
plt.plot(année, population_evolution, color='b')

plt.title('Évolution de la population de la région 89')
plt.xlabel('Année')
plt.ylabel('Population Totale')

plt.show()

```

Avec comme résultat de graphique comme cela :



Ainsi pour conclure nous pouvons voir que les rumeurs à propos de la diminution de la population de la région de l'Yonne n'est pas totalement fausse, plus de 15 000 habitants ont quitté la région en 15 ans.

