

TP 4 : LES VILLES DE FRANCE (SUITE)

1 Introduction

L'objectif de ce TP est d'approfondir et d'améliorer le TP3. Si vous ne l'avez pas fait, finissez d'abord le TP 3 avant de faire ce TP, vous pourrez alors reprendre une grande partie du code pour débiter ce TP.

Sauf si cela est précisé, vous ne devez pas utiliser de boucle (for, while) ou de branchement conditionnel (if) durant ce TP.

2 Matrice des distances euclidiennes entre villes

Afin de rendre les calculs du TP 3 plus rapides, nous allons désormais écrire une fonction qui calcule la distance euclidienne entre deux villes. Vous reprendrez le code du précédent TP pour débiter celui-ci.

1. Nous allons dans un premier temps représenter les villes par leurs coordonnées 3D, et c'est à partir de ces coordonnées 3D que l'on pourra calculer les distances euclidiennes.

Pour calculer les coordonnées 3D, vous pourrez utiliser les équations suivantes :

$$\begin{aligned}x &= R * \cos(p_{lat}) * \sin(p_{lon}) \\y &= R * \cos(p_{lat}) * \cos(p_{lon}) \\z &= R * \sin(p_{lat});\end{aligned}$$

2. Calculez la distance entre Caen et Rouen avec cette nouvelle méthode. Comparez avec la valeur précédente et commentez.
3. En utilisant la fonction `plt.scatter` de `matplotlib`, affichez chaque ville sous la forme de point de coordonnée (x,y).

3 Villes les plus proches

Écrire un programme qui recherche les 10 villes les plus proches d'une ville donnée. Vérifiez en demandant les 10 villes les plus proches de Caen, au sens de la distance euclidienne.

Vous devriez trouver 'Caen', 'Hérouville-Saint-Clair', 'Louvigny', 'Mondeville', 'Saint-Germain-la-Blanche-Herbe', 'Epron', 'Colombelles', 'Fleury-sur-Orne', 'Giber-ville', 'Bretteville-sur-Odon'.

Vous ferez de même avec le calcul des 10 villes les plus éloignées d'une ville donnée.

4 Histogramme des distances des villes

Au moyen de la fonction `np.histogram` calculez l'histogramme des distances entre villes.

Affichez-le avec la librairie `matplotlib`.

5 Conclusion

1. Testez maintenant le TP avec toutes les villes de France.
2. Nous avons actuellement pas géré les doublons dans la base (par exemple une ville ayant plusieurs circonscription électorale ou des villes différentes ayant le même nom), en utilisant le code INSEE des villes faites en sorte de mieux gérer ces doublons. Si nécessaire, vous pouvez utiliser des boucles for.