

La liste du travail présentée ci-après est exhaustive. Elle concentre l'essentiel du travail et peut être améliorée et enrichie. Les données utilisées : [Test_Clusters_Atypiques.txt](#), [Test_Clusters_Distincts.txt](#), [Test_Clusters_Melanges.txt](#) et [Test_Clusters_Random.txt](#)



À faire

Faire les exercices 1 et 2 de la série de TD. Votre travail doit reprendre principalement les tâches suivantes :

1. Importer des données

- (a) Importer les librairies
- (b) Importer un fichier et le visualiser

2. Nuage de points des variables 2 à 2

3. K-means

- (a) Coordonnées des centres de chaque cluster
- (b) Visualisation des clusters avec leur centre

4. Classification Ascendante Hiérarchique

- (a) Dendrogramme
- (b) Résultats avec différentes métriques
- (c) Découpage en fonction du nombre de classes
- (d) Visualisation des clusters avec leur centre

5. DBSCAN

- (a) Distance maximale et voisinage d'une observation
- (b) Création des clusters
- (c) Le nombre d'observation par cluster
- (d) Visualisation des clusters avec leur centre

On doit retrouver dans votre travail sur python les configurations suivantes:

- # Définition du nombre de clusters qu'on souhaite avoir – affichage du numéro de cluster qu'on a attribué à chaque observation
- # Génération de la matrice des liens - affichage du dendrogramme
- # Utiliser la fonction NearestNeighbors – ordonner les distances des voisins en croissant – affichage graphique des distances