

Atelier 6 : Introduction au Clustering

Exercice 1 : Visualisation de données simples avec K-Means

1. **But :** Appliquer K-Means sur un ensemble de points en 2D pour observer la formation des clusters.
 2. **Instructions :**
 - Créer un ensemble de points en 2D (par exemple, avec NumPy ou un dataset généré).
 - Appliquer K-Means avec 3 clusters.
 - Visualiser les clusters sur un graphique scatter avec leurs centres.
 3. **Questions :**
 - Combien de points appartiennent à chaque cluster ?
 - Que remarquez-vous concernant la position des centres des clusters ?
-

Exercice 2 : Variation du nombre de clusters (kk)

1. **But :** Observer l'impact du choix de k sur les résultats du clustering.
 2. **Instructions :**
 - Reprendre le même ensemble de données que l'exercice 1.
 - Appliquer K-Means avec $k=2$, $k=3$, et $k=4$.
 - Comparer les graphiques obtenus pour chaque valeur de k .
 3. **Questions :**
 - Quel k semble le plus adapté ? Pourquoi ?
 - Quels problèmes peuvent survenir si k est mal choisi ?
-

Exercice 3 : Segmentation des clients avec un petit dataset

1. **But :** Utiliser K-Means pour regrouper des clients en fonction de deux caractéristiques simples (par ex., âge et dépenses).
2. **Instructions :**
 - Fournir un petit dataset avec des colonnes "Âge" et "Dépenses".
 - Appliquer K-Means avec $k=3$.
 - Visualiser les clusters sur un graphique avec les clients regroupés.
3. **Questions :**
 - Quels groupes peuvent être identifiés ?
 - Quelles hypothèses pouvez-vous tirer par rapport au comportement des clients de chaque groupe ?