

## Suite - Apprentissage non supervisé

### Exercice :

1. Chargez un jeu de données de votre choix à partir de Kaggle ou de UCI
2. Explorez le jeu de données pour comprendre sa structure, y compris le nombre de caractéristiques, les statistiques descriptives, et la distribution des classes de fleurs.
3. Effectuez une réduction de dimension à l'aide de l'analyse en composantes principales (PCA) pour visualiser les données en deux dimensions.
4. Appliquez un algorithme de clustering tel que K-Means pour regrouper les individus en fonction de leurs caractéristiques. Utilisez la métrique nécessaire pour trouver un nombre optimal de clusters.
5. Visualisez les clusters obtenus à l'aide d'un diagramme de dispersion en 2D.
6. Évaluez la qualité des clusters obtenus en utilisant des métriques telles que l'inertie et le coefficient de silhouette.
7. Quel est l'objectif de l'apprentissage non supervisé, et en quoi diffère-t-il de l'apprentissage supervisé ?
8. Comment choisissez-vous le nombre optimal de clusters (k) pour l'algorithme K-Means ?
9. Pouvez-vous expliquer l'idée derrière la standardisation des caractéristiques dans le contexte de l'apprentissage non supervisé ?
10. Quels sont les avantages et les limites de l'algorithme K-Means ?
11. Comment évaluez-vous la qualité des clusters obtenus ? Quelles mesures ou critères pourriez-vous utiliser ?