

# **TD – Pratique Analyse exploratoire**

## **1. Exploration et Préparation des Données**

### **(a) Chargement et aperçu des données**

- Charger le dataset "Mall\_Customers\_modified.csv".
- Afficher les cinq dernières lignes du dataset.
- Décrire brièvement le dataset :
  - Combien de colonnes et de lignes contient-il ?
  - Quelles sont les colonnes présentes et leurs types de données ?

### **(b) Prétraitement des données**

- **Suppression des colonnes non pertinentes et gestion des variables catégorielles :**
  - Supprimer la colonne 'CustomerID' car elle ne fournit pas d'informations utiles pour les analyses.
  - **Encoder la variable catégorielle 'Genre' en utilisant un encodage numérique où 'Male' est représenté par 0 et 'Female' par 1.**
- **Traitements des valeurs manquantes :**
  - Vérifier s'il y a des valeurs manquantes et les traiter en les remplaçant par la médiane des valeurs existantes.
- **Normalisation des variables :**

- Normaliser les variables numériques afin d'assurer une bonne convergence de l'algorithme.

## 2. Visualisation des Données

Réaliser une analyse exploratoire à l'aide de graphiques :

- Tracer un histogramme de la distribution de l'âge des clients.
- Créer un nuage de points (scatter plot) représentant l'Annual Income (Revenu annuel) vs Spending Score (Score de dépenses).