# 📌 Project Overview: Mall Customer Segmentation

## 1. Project Title

Mall Customer Segmentation using Unsupervised Learning

## 2. Objective

The main goal of this project is to analyze customer data from a mall and group customers into different segments based on their purchasing behavior and demographics. This will help businesses understand customer patterns, improve targeted marketing strategies, and enhance customer satisfaction.

## 3. Problem Statement

Retail businesses often face difficulty in identifying the right target audience for promotions. Treating all customers in the same way leads to ineffective marketing campaigns and revenue loss. By segmenting customers into meaningful groups, businesses can tailor their services and offers accordingly.

## 4. Dataset Description

The dataset typically includes the following features:  
- CustomerID: Unique ID for each customer  
- Gender: Male/Female  
- Age: Age of the customer  
- Annual Income (k$): Income level of the customer  
- Spending Score (1–100): Score assigned based on spending habits and mall behavior

## 5. Methodology

1. Data Collection & Cleaning – Import dataset, check for null values, handle inconsistencies.  
2. Exploratory Data Analysis (EDA) – Visualize distributions of income, age, spending score, gender ratio.  
3. Feature Selection – Choose important features for segmentation (Age, Income, Spending Score).  
4. Scaling – Normalize/Standardize data to bring features on the same scale.  
5. Clustering Algorithm – Apply K-Means Clustering (main model) and optionally compare with Hierarchical Clustering.  
6. Optimal Clusters – Use Elbow Method & Silhouette Score to decide the best number of clusters.  
7. Visualization – Plot customer groups in 2D/3D to interpret clusters.  
8. Insights – Describe each customer segment (e.g., “Young High Income High Spenders”, “Middle-aged Low Spenders”, etc.).

## 6. Expected Outcomes

- Segmentation of mall customers into distinct groups based on income and spending behavior.  
- Business insights such as:  
 - High income, high spenders → potential premium customers.  
 - Low income, high spenders → budget-conscious but valuable customers.  
 - Low spenders → customers requiring discount-based targeting.  
- Visualization of clusters for better decision-making.

## 7. Technologies & Tools

Python, Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn

## 8. Applications

Targeted marketing campaigns, Customer loyalty programs, Personalized offers & discounts, Better allocation of mall resources