Mini Project Report: Unsupervised and Supervised Learning on Mall Customers Dataset

# 1. Problem Statement

The objective of this mini project is to explore customer segmentation using both supervised and unsupervised machine learning techniques. We aim to:  
- Understand customer data through exploratory data analysis (EDA).  
- Use K-Means clustering to segment customers (unsupervised learning).  
- Use a classification algorithm to predict customer segments (supervised learning).  
- Evaluate the performance of the applied models.  
- Derive actionable business insights based on findings.

# 2. Data Understanding

The Mall Customers dataset includes demographic information and spending behavior of customers. It has the following features:  
- CustomerID: Unique ID for each customer  
- Gender: Male or Female  
- Age: Age of the customer  
- Annual Income (k$): Customer’s yearly income  
- Spending Score (1–100): Score assigned by the mall based on customer behavior  
The goal is to understand customer groups and predict behavior using these features.

# 3. Model Building

We applied the following models:  
  
Unsupervised Learning:  
- K-Means Clustering: Applied to identify different segments of customers based on income and spending score.  
- PCA: Used for dimensionality reduction and visualization of clusters.  
  
Supervised Learning:  
- Logistic Regression / Decision Tree: Used to predict the cluster label assigned by K-Means.  
- Dataset was split into training and testing sets.

# 4. Evaluation

Unsupervised Learning Evaluation:  
- K-Means identified 5 optimal clusters using the Elbow Method.  
- Clusters were visualized using PCA, showing clear separation.  
  
Supervised Learning Evaluation:  
- Accuracy of the classifier in predicting cluster labels was high.  
- Confusion matrix and classification report confirmed strong prediction capability.

# 5. Conclusion

The analysis enabled us to segment mall customers effectively:  
- Cluster analysis revealed distinct customer groups with varied spending behavior.  
- Supervised models successfully predicted customer groupings.  
These insights can help the mall develop targeted marketing strategies and improve customer engagement.