**Task 2: Customer Segmentation Report (Mall Customers Dataset)**

# 1. Problem Statement and Objective

The objective of this project is to perform customer segmentation using unsupervised learning. By clustering customers based on their demographic and spending behavior, we aim to identify distinct groups that can help businesses design tailored marketing strategies to maximize customer satisfaction and revenue.

# 2. Dataset Description and Loading

The Mall Customers dataset consists of customer demographic information including CustomerID, Gender, Age, Annual Income, and Spending Score. These features provide insights into customer profiles. The dataset was loaded into Python for further analysis using pandas.

# 3. Data Cleaning and Preprocessing

The dataset was checked for missing values and inconsistencies. No missing values were present. Features such as Gender were encoded into numeric form, and continuous features like Age, Annual Income, and Spending Score were scaled for clustering algorithms.

# 4. Exploratory Data Analysis (EDA)

EDA revealed important customer insights:  
• Age distribution shows that most customers are between 25-40 years old.  
• Annual income varies widely, with some customers earning very high incomes.  
• Spending score indicates that customers fall into low, medium, and high spenders.  
This analysis helped understand the overall profile of customers before applying clustering.

# 5. Model Building – K-Means Clustering

K-Means clustering was applied to segment customers based on Age, Annual Income, and Spending Score. The Elbow Method was used to determine the optimal number of clusters, which was found to be 5. This resulted in 5 distinct customer segments, each representing a unique profile of spending habits.

# 6. Visualization of Clusters

To visualize high-dimensional data, PCA (Principal Component Analysis) and t-SNE (t-distributed Stochastic Neighbor Embedding) were applied. The clusters were plotted in 2D space, clearly showing separation between the customer segments. These visualizations confirmed the effectiveness of clustering.

# 7. Marketing Strategies for Identified Segments

Based on the clustering results, the following marketing strategies are suggested:  
• Cluster 1 (Young High Spenders): Focus on luxury and premium products with personalized offers.  
• Cluster 2 (Older Low Spenders): Provide discounts, loyalty programs, and essential value products.  
• Cluster 3 (Average Income & Moderate Spending): Introduce mid-range product bundles to increase spending.  
• Cluster 4 (High Income but Low Spending): Target with exclusive promotions to encourage spending.  
• Cluster 5 (Students/Young Low Income): Affordable product lines and campus-based campaigns.

# 8. Final Conclusion with Insights

The customer segmentation analysis successfully grouped mall customers into 5 distinct clusters. Each cluster reflects different age groups, income levels, and spending habits. By tailoring marketing strategies to each group, businesses can improve customer engagement, increase sales, and maximize retention.