Executive Summary

This project aims to perform customer segmentation using clustering techniques on the Mall Customers dataset. The objective is to categorize customers based on their purchasing behavior and demographic traits to enable targeted marketing strategies.

Key Steps and Findings:

- **Data Exploration**:
    - The dataset contains demographic and behavioral attributes of mall customers, including age, gender, annual income, and spending score.
    - Initial exploratory data analysis (EDA) involved visualizing distributions and relationships using histograms, KDE plots, boxplots, and pair plots.
    - Gender-based distribution analysis revealed slight variations in spending habits and income.
- **Clustering Analysis**:
    - **Univariate Clustering**: KMeans clustering was first applied on a single feature — *Annual Income*. Three clusters were identified, representing low, medium, and high-income groups.
    - **Bivariate Clustering**: A more insightful segmentation was performed using both *Annual Income* and *Spending Score*. Using the elbow method, 5 clusters were found to be optimal.
    - These clusters revealed distinct customer personas, such as high-income low spenders and low-income high spenders, which are valuable for targeted business strategies.
- **Insights**:
    - Cluster analysis effectively grouped customers into segments with similar behavior.
    - Visualization of clusters provided a clear understanding of customer groupings and helped identify target segments for customized marketing.

Analysis

Target Cluster

• Target group would be cluster 1which has a high Spending Score and high income

• 54 percent of cluster 1 shoppers are women. We should look for ways to attract these customers using a marketing campaign targeting popular items in this cluster

• Cluster 2 presents an interesting opportunity to market to the customers for sales event on popular items.