Mall Customer Clustering Report

# 1. Introduction

This report summarizes the methodology, findings, and business insights from a clustering analysis conducted on the Mall Customer Segmentation dataset. The objective was to explore customer groupings using unsupervised learning techniques, specifically K-Means and Hierarchical Clustering, and to provide interpretations that are useful for business.

# 2. Data Exploration and Preprocessing

The dataset includes information on 200 mall customers, covering both demographic and spending data. There are no missing values in the dataset. Gender was encoded using Label Encoding, and the features were standardized with StandardScaler to facilitate fair clustering. The relevant features selected for analysis include Gender, Age, Annual Income, and Spending Score.

# 3. Clustering Algorithms

## 3.1 K-Means Clustering

K-Means was implemented using scikit-learn. The Elbow Method and Silhouette Score were utilized to determine the optimal number of clusters, which was found to be k=5. A 2D plot effectively displayed clearly defined clusters based on Annual Income and Spending Score.

## 3.2 Hierarchical Clustering

Agglomerative clustering was conducted using the Ward linkage method. The dendrogram indicated distinct separation and supported a solution of 5 clusters.

# 4. Evaluation and Interpretation

Silhouette scores were calculated for K-Means clustering across multiple values of k, with the highest score observed at k=5. Clustering results were consistent between K-Means and Hierarchical Clustering.

## 4.1 Cluster Profiling & Business Insights

* Cluster 1 – High Income, High Spending

Business Action: Promote premium products and provide exclusive VIP experiences.

* Cluster 2 – Low Income, High Spending

Business Action: Target offers discounts and loyalty rewards to its customers..

* Cluster 3 – High Income, Low Spending

Business Action: Enhance engagement through tailored campaigns..

* Cluster 4 – Low Income, Low Spending

Business Action: Focus on value-driven promotions.

* Cluster 5 – Average Income & Spending

Business Action: Maintain retention with seasonal deals.

# 5. Conclusion

Customer segmentation through clustering techniques has uncovered significant groupings that businesses can utilize to enhance personalization, retention, and revenue. This analysis fosters data-driven decision-making in both marketing and operations.