CLUSTERING REPORT

1. Introduction

This report presents the results of customer segmentation using Agglomerative Clustering. The dataset consists of customer demographics and transaction history, which were preprocessed and clustered to identify distinct customer groups.

2. Data Preprocessing

The dataset includes customer demographic and transactional information. Key features used for clustering:

- Region (Encoded)
- Total Spent
- Total Transaction
- Average Transaction Value
- Recency (Days since last purchase)

Feature scaling was performed using StandardScaler to normalize numerical attributes.

3. Optimal Number of Clusters

The optimal number of clusters was found to be **k = 7**. The cluster quality metrics for this optimal k are:

Davies-Bouldin Index (DBI): 1.1624

Silhouette Score: 0.2130

4. Clustering Results

Customers were grouped into **7 distinct segments** based on their purchasing behavior. The clusters were visualized using scatter plots with **Total Spent vs. Recency** as key differentiators.