Pranav Dhobi - Customer Segmentation Report

1. Introduction

Customer segmentation is essential for businesses to identify distinct customer groups and tailor marketing strategies accordingly. In this task, we used clustering techniques to group customers based on their profiles and transaction histories. The goal is to find meaningful clusters and provide insights into customer behavior to improve engagement and retention strategies.

2. Methodology

1. Data Aggregation:

- Combined transaction data with customer profiles to create a summary of each customer's total spending (TotalValue) and purchase quantity (Quantity).
- Incorporated regional information using one-hot encoding.

2. Feature Scaling:

- Standardized the features using StandardScaler to ensure equal contribution of all attributes during clustering.

3. Clustering Algorithm:

- Applied the KMeans clustering algorithm with K values ranging from 2 to 10.
- Calculated the Davies-Bouldin (DB) Index for each K to evaluate cluster compactness and separation.

4. Optimal Number of Clusters:

- The K value with the lowest DB Index was chosen as the optimal number of clusters.

5. Cluster Visualization:

- Created scatter plots of the clusters using TotalValue and Quantity to observe the distribution of customer groups.

3. Results

1. Optimal Number of Clusters:

- The optimal number of clusters was determined to be 8, with the lowest DB Index of 0.632.