Customer Segmentation Report

Feature Selection:

I used Recency, Frequency and Monetary (RFM) as the features for classifying customers into clusters. The RFM model is widely used in customer analysis as it segments a company's customer base based on purchasing behavior. Specifically, it evaluates:

- Recency – How recently a customer made a purchase.
- Frequency – How often a customer makes purchases.
- Monetary Value – How much money a customer spends.

This approach helps identify high-value customers, improve engagement with lower-scoring ones and enhance retention strategies.

Number of Clusters:

The optimal number of clusters was determined using the within-cluster sum of squares (WCSS) and the elbow method. Since no clear elbow point was observed, I selected 10 clusters, as it resulted in the lowest Davies-Bouldin Index (DBI) and a high Silhouette Score, indicating well-defined clusters.

Visualization of Clusters:

Clusters were visualized in a two-dimensional space using Principal Component Analysis (PCA). The scatter plot shows cluster distribution and while some overlap exists, the clusters remain reasonably distinct. PCA helped reduce dimensionality while preserving most of the variance for meaningful visualization.

Clustering Metrics:

- Silhouette Score: 0.41
- Davies-Bouldin Index: 0.74