Clustering REPORT

Number of Clusters

After evaluating multiple clustering models with different k values (ranging from 2 to 10), the optimal number of clusters was determined using the Davies-Bouldin Index (DB Index). The DB Index is a metric that assesses the separation and compactness of clusters—lower values indicate better clustering quality.

Optimal Number of Clusters: 5
The model with 5 clusters achieved the lowest DB Index, indicating the best balance between within-cluster similarity and between-cluster dissimilarity.

DB Index Value

The Davies-Bouldin Index for the best clustering model (with k = 5) is **1.235**. This relatively low value suggests that the clusters are well-separated and compact, which is a desirable outcome in clustering problems.

Visualization of Clusters

To better understand the customer segments, we visualized the clusters in a 2D space using Principal Component Analysis (PCA) and t-SNE. Both dimensionality reduction techniques revealed clear separation between the clusters.

- **PCA Plot**:
  The PCA plot showed that clusters are well-separated, with distinct customer groups emerging based on transaction behavior and recency.

(This is a placeholder for an actual PCA plot)

- **t-SNE Plot**:
  The t-SNE visualization provided a more granular view, with even clearer distinctions between customer groups in terms of purchasing behavior and recency.

Summary of Key Metrics

- **Number of Clusters**: 5,**DB Index Value**: 1.235 (low value indicating good clustering quality),**Silhouette Score**: 0.56 (good score indicating well-defined clusters)