Customer Segmentation Report

Task 3: Customer Segmentation / Clustering

Objective:

The goal of this task is to perform customer segmentation using clustering techniques on profile and transaction data.

Results:

1. Number of Clusters Formed: 3 (as determined by the optimal Davies-Bouldin Index).

2. Davies-Bouldin (DB) Index Value: 0.79.

3. Clustering Metrics:

   - Silhouette Score: 0.45

   - Within-Cluster Sum of Squares (WCSS): 129.6

Approach:

1. Data Preprocessing:

   - Converted date columns to datetime format.

   - Engineered features such as Total Spend, Average Spend, Days Since Signup, and Transaction Count.

   - One-hot encoded the 'Region' column for clustering.

2. Feature Scaling:

   - Applied StandardScaler for normalization of numerical features.

3. Clustering Algorithm:

- K-Means clustering was applied with k values ranging from 2 to 10.

- Optimal number of clusters (k=3) was chosen based on the lowest Davies-Bouldin Index.

4. Visualization:

  - Line plot for DB Index vs Number of Clusters (k).

  - Scatter plot of clusters visualized in 2D space using PCA.

Deliverables:

- Segmented customer data saved as 'Customer_Segments.csv'.

- Visual representations of clusters provided.

Conclusion:

The clustering results provide meaningful customer segments, enabling targeted marketing strategies and resource optimization.