

Customer Segmentation Report

Objective

To perform customer segmentation using clustering techniques, utilizing customer profile and transaction data to identify distinct customer groups for tailored business strategies.

Data Overview

- **Datasets Used:**
 - Customers.csv: Customer demographic data.
 - Products.csv: Product information.
 - Transactions.csv: Purchase details.
- **Key Features for Clustering:**
 - TotalSpend: Total transaction value per customer.
 - TotalItems: Total quantity of products purchased.
 - UniqueProducts: Number of unique products purchased.

Clustering Methodology

1. **Data Preprocessing:**
 - Merged the three datasets to create a unified view of customer transactions.
 - Aggregated transactional data to compute customer-level metrics: TotalSpend, TotalItems, and UniqueProducts.
 - Normalized the features using StandardScaler to ensure uniform scaling for clustering.
2. **Clustering Algorithm:**
 - Used the K-Means algorithm for segmentation.
 - Determined the optimal number of clusters using the Elbow Method and Davies-Bouldin (DB) Index.

3. Evaluation Metrics:

- Calculated the Davies-Bouldin Index, with a lower value indicating better cluster separation and compactness.
- Conducted Principal Component Analysis (PCA) to reduce dimensionality and visualize clusters.

Clustering Results

1. Number of Clusters Formed:

- The optimal number of clusters was determined to be “4”, based on the Elbow Method and DB Index analysis.

2. Evaluation Metrics:

- **Davies-Bouldin Index:** 0.891 (indicating well-defined clusters).

3. Cluster Characteristics:

- **Cluster 1:** High spenders with frequent purchases of diverse products.
- **Cluster 2:** Moderate spenders with medium transaction frequency.
- **Cluster 3:** Low spenders with minimal product diversity.
- **Cluster 4:** Occasional high-value transactions with low frequency.

4. Visualization:

- PCA reduced the features to 2 dimensions for visualization.
- Clusters were well-separated and distinct, as shown in the scatter plot.