

Customer Segmentation Report

Objective

To segment customers using clustering techniques based on profile and transaction information. The evaluation focuses on the Davies-Bouldin (DB) Index and the visual representation of clusters.

Methodology

Data Preparation:

Imported data from Customers.csv and Transactions.csv.

Engineered features:

- Total spending (total_spending).
- Number of transactions (num_transactions).
- Average transaction value (avg_transaction_value).
- Merged transaction features with customer profile information.
- Handled missing values by filling with zeros.

Feature Scaling:

Applied StandardScaler to normalize feature values for clustering.

Clustering Approach:

Used K-Means clustering with cluster numbers ranging from 2 to 10.

Evaluated clustering configurations using:

- Inertia (Elbow Method).
- Davies-Bouldin Index (DB Index).

Optimal Cluster Selection:

Based on DB Index and Elbow Method plots, the optimal number of clusters was determined to be 3.

Evaluation Metrics:

Calculated DB Index for the final clustering configuration.

Visualization:

Used PCA (Principal Component Analysis) to reduce dimensionality and visualize clusters in 2D.

Output:

Saved cluster assignments to Customer_Segmentation.csv.

Results

Optimal Number of Clusters:

3 clusters were selected as the optimal configuration.

Davies-Bouldin Index:

DB Index: 0.82 (lower values indicate better clustering performance).

Other Metrics:

Inertia values were analyzed using the Elbow Method plot.

Cluster Visualizations:

Clusters were visualized in 2D using PCA, revealing clear separations between groups.

Visual Representations

1. Elbow Method Plot:

Illustrates the inertia values for different cluster numbers (2-10), showing diminishing returns after 3 clusters.

2. Davies-Bouldin Index Plot:

Plots DB Index for cluster numbers 2-10, with the lowest value at 3 clusters

3. PCA-Based Cluster Visualization:

Visualizes customer clusters in 2D space, colored by cluster assignment.

Conclusion

The clustering analysis successfully segmented customers into 3 distinct groups based on transaction behavior and profile information. The segmentation can provide insights for targeted marketing strategies and personalized customer engagement.