

Clustering Report

1. Introduction

This report outlines the results of a customer segmentation analysis performed using clustering techniques. The objective of this study was to group customers based on their profile and transaction data to derive actionable insights. Both customer profile data (e.g., region) and transactional data (e.g., total expenditure, frequency, average value) were utilized for clustering. Key metrics, including the Davies-Bouldin (DB) Index, were used to evaluate the quality of the clusters.

2. Clustering Methodology

Data Preparation

1. Data from two sources—Customers.csv and Transactions.csv—were merged using the CustomerID field.
2. Features engineered for clustering included:
 - **Total Expenditure:** Sum of transaction values for each customer.
 - **Transaction Count:** Number of transactions per customer.
 - **Average Transaction Value:** Average value of transactions for each customer.
 - **Region:** Encoded as a numerical feature.
3. Features were normalized using standard scaling to ensure uniformity in clustering.

Clustering Algorithm

- K-Means clustering was selected due to its efficiency and ability to handle large datasets.
 - The optimal number of clusters was determined by minimizing the Davies-Bouldin (DB) Index across cluster counts ranging from 2 to 10.
-

3. Clustering Results

Number of Clusters Formed

The optimal number of clusters was determined to be **10**, as this configuration yielded the lowest Davies-Bouldin Index.

Davies-Bouldin Index (DB Index)

- **Value: 0.979**
- The DB Index evaluates the compactness and separation of clusters. Lower values indicate better clustering quality.

Cluster Characteristics

1. **High Spenders:**

- Customers with the highest total expenditure and frequent transactions.
 - Likely to be loyal and high-value customers.
 - 2. **Moderate Spenders:**
 - Customers with average expenditure and transaction counts.
 - Target group for upselling strategies.
 - 3. **Low Spenders:**
 - Customers with low transaction frequency and expenditure.
 - Potential for targeted campaigns to increase engagement.
 - 4. **Regional Insights:**
 - Certain clusters were concentrated in specific regions, indicating regional patterns in spending behavior.
-

4. Visual Representations

Visualization Techniques

1. **Pair Plot:**
 - Relationships between total expenditure, transaction count, and average value were visualized, with cluster labels represented by distinct colors.
2. **Cluster Distribution:**
 - A bar chart showed the number of customers in each cluster, highlighting imbalances and opportunities for targeted strategies.

Key Insights

- Cluster 1 had the highest average expenditure but a lower number of customers, representing a small but highly valuable segment.
 - Clusters 4 and 6 exhibited moderate expenditure and high transaction counts, suggesting frequent but lower-value purchases.
 - Low-value clusters (e.g., Cluster 9) represent a significant portion of the customer base, indicating potential for growth.
-

5. Business Insights from EDA

1. **Customer Segmentation:** High-spending customers contribute disproportionately to total revenue. Prioritizing retention strategies for these customers can maximize profitability.
2. **Transactional Patterns:** Customers with frequent transactions but lower average values indicate potential for upselling through targeted offers.

3. **Regional Variations:** Spending behavior varies significantly by region, suggesting a need for region-specific marketing strategies.
 4. **Engagement Opportunities:** Low-frequency, low-expenditure customers represent a large segment. Incentives can be designed to boost their transaction frequency.
 5. **Product Affinity:** Insights from transaction-level data reveal clusters with distinct product preferences, enabling personalized product recommendations.
-

6. Conclusion

This clustering analysis successfully segmented customers into distinct groups based on transactional behavior and profile data. The following conclusions can be drawn:

1. **Target High-Value Customers:** Focus on retaining and rewarding customers in high-spending clusters.
2. **Upsell Opportunities:** Develop marketing strategies for moderate spenders to increase their transaction frequency and expenditure.
3. **Engage Low Spenders:** Design campaigns to convert low-value customers into more engaged and frequent buyers.