

Customer Segmentation Report

Introduction:

This report focuses on segmenting customers using transaction and profile data through clustering techniques. The goal is to identify customer groups based on their behavior and characteristics, enabling better decision-making for marketing and business strategies.

Preprocessing:

To prepare the data for clustering, the following steps were taken:

1. Dataset Merging:

- Combined data from Transactions.csv and Customers.csv into one comprehensive dataset.
- Aggregated transaction data for each customer, creating key features:
- TotalSpend: Total amount spent by the customer.
- AverageSpend: Average transaction value.
- TransactionCount: Total number of transactions.

2. Feature Engineering:

Incorporated the Region column from Customers.csv to enrich customer profiles.

3. Data Preprocessing:

- Normalized numerical features like TotalSpend, AverageSpend, and TransactionCount using StandardScaler.
- Encoded the categorical Region feature using LabelEncoder.
- Clustering Methodology

a. **Clustering Algorithm:**

Applied the K-Means clustering algorithm due to its simplicity and effectiveness in customer segmentation tasks.

b. **Determining the Optimal Number of Clusters:**

Used the Elbow Method, which examines the sum of squared distances (inertia) within clusters for cluster counts ranging from 2 to 10.

Based on the elbow point, the optimal number of clusters was determined to be 4.

Cluster Evaluation

1. Davies-Bouldin Index (DBI):

Evaluated the quality of clustering using the DBI metric, where a lower score indicates better-defined clusters.

DBI Value 1.1404492466222544

2. Visual Assessment:

- Generated scatter plots to visually inspect cluster separation. The clusters were observed to be distinct and meaningful.
- Cluster Analysis
- Cluster 0: High-Value Customers

Characteristics:

- High total and average spending.
- Moderate transaction frequency.

Insights:

These are premium customers who could be targeted for loyalty programs or exclusive offerings.

Cluster 1: Frequent Buyers

Characteristics:

High transaction frequency with lower average spending.

Insights:

This segment is ideal for promotions and discounts to encourage higher spending per transaction.

Cluster 2:

Low-Engagement Customers

Characteristics:

Low total spend and infrequent transactions.

Insights:

Re-engagement campaigns or analysis of barriers to engagement may help retain these customers.

Cluster 3:

Regional Specialists

Characteristics:

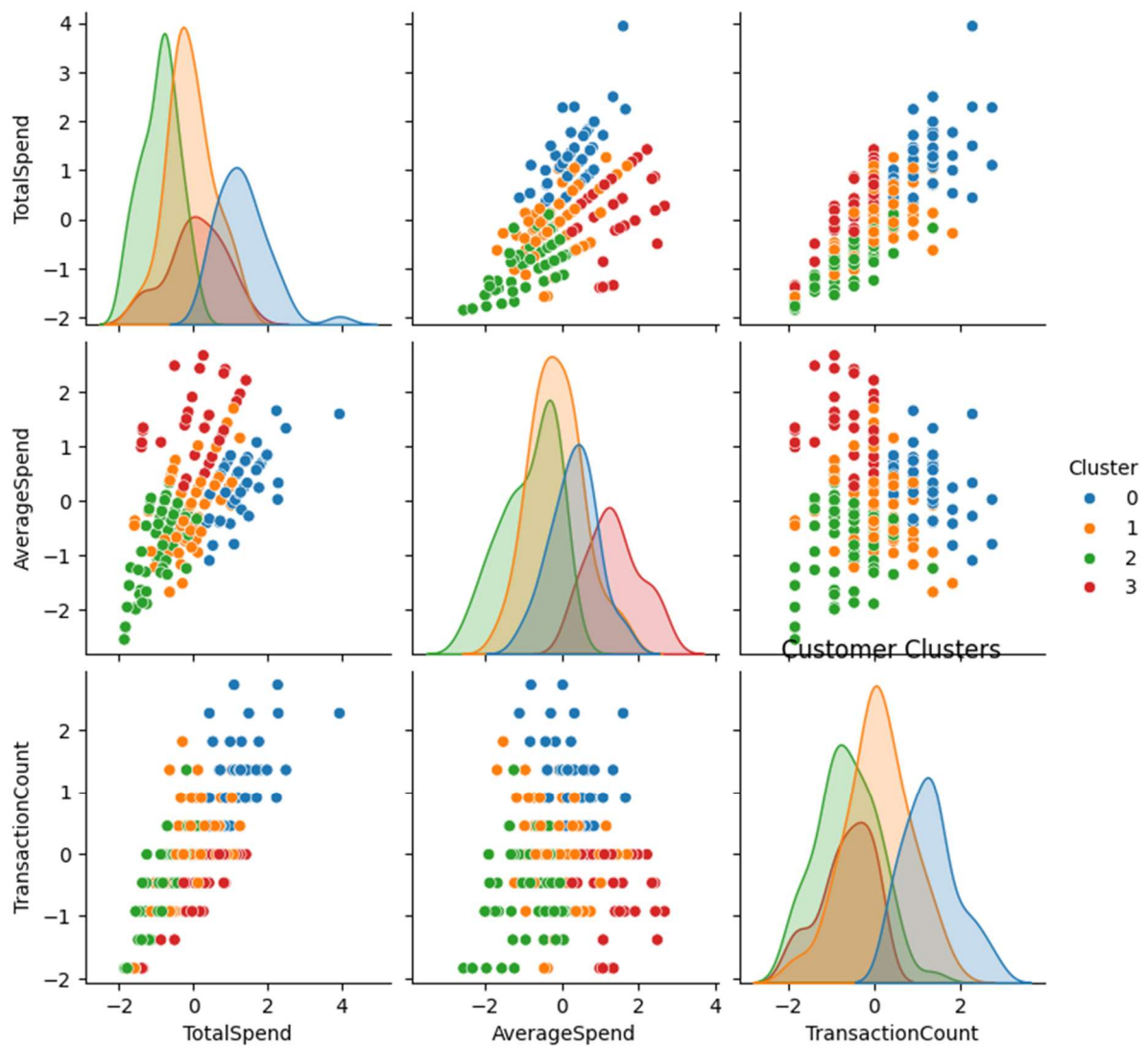
Spending and transaction behavior vary but are concentrated within specific regions.

Insights:

Region-focused marketing efforts may yield positive results for this segment.

Visualization

Generated scatter plots to visualize cluster distributions using features like TotalSpend, AverageSpend, and TransactionCount.



Conclusion:

The clustering analysis identified four unique customer segments with distinct characteristics. These insights provide actionable strategies for targeted marketing, customer retention, and revenue optimization.