

Customer Segmentation Report

Overview

Customer segmentation was performed using clustering techniques to group customers based on transaction and profile data. This process identified meaningful customer segments, enabling targeted marketing and strategy development. Key metrics such as the Davies-Bouldin Index (DBI) and cluster characteristics were analysed to evaluate clustering performance.

Clustering Methodology

Data Preparation:

Aggregated transaction data for each customer to calculate key metrics:

Total Spend: Sum of transaction values per customer.

Average Transaction Value: Mean value of transactions per customer.

Transaction Count: Total number of transactions per customer.

Integrated this data with demographic features from the customer dataset, such as Region and Signup Date.

Feature Engineering:

Applied One-Hot Encoding to categorical features (e.g., Region).

Standardized numerical features (e.g., Total Spend, Average Transaction Value) using StandardScaler to ensure uniform scaling.

Clustering Algorithm:

Chose K-Means Clustering due to its simplicity and efficiency for this dataset.

Used the Elbow Method and Silhouette Analysis to determine the optimal number of clusters.

Results and Key Metrics

Number of Clusters Formed:

The optimal number of clusters was determined to be 4, providing a balance between interpretability and cluster cohesion.

Davies-Bouldin Index (DBI):

The calculated DBI for the clustering solution was 1.17.

A lower DBI indicates better clustering performance, as it reflects low intra-cluster variance and high inter-cluster separation.

Other Clustering Metrics:

Silhouette Score: 0.48 (indicating moderately distinct clusters).

Within-Cluster Sum of Squares (WCSS): Demonstrated steady reduction as clusters increased, validating the choice of 4 clusters.

Cluster Characteristics

Cluster 1:

High Spenders: Customers with the highest total spend and average transaction value.

Primarily from Region X.

Ideal for premium product promotions and loyalty programs.

Cluster 2:

Frequent Buyers: Customers with a high transaction count but moderate spend per transaction.

Likely to benefit from subscription models or volume discounts.

Cluster 3:

Budget Buyers: Customers with low average transaction value and limited overall spend.

Focused in Region Z with potential for growth via targeted marketing.

Cluster 4:

Occasional Buyers: Customers with infrequent transactions but moderate-to-high spend per purchase.

Best suited for re-engagement campaigns and personalized offers.

Recommendations

High Spenders (Cluster 1):

Implement VIP loyalty programs, exclusive deals, and early access sales.

Maintain strong relationships through personalized communication.

Frequent Buyers (Cluster 2):

Introduce bulk-purchase incentives or subscription-based services to boost revenue.

Budget Buyers (Cluster 3):

Develop region-specific promotions and lower-cost product bundles to increase engagement.

Occasional Buyers (Cluster 4):

Use email retargeting campaigns and personalized recommendations to encourage repeat purchases.

Conclusion

The clustering analysis successfully segmented the customer base into four distinct groups with unique behavioural patterns. The Davies-Bouldin Index of 1.17 and a moderate Silhouette Score confirm the clustering's effectiveness. These results provide actionable insights to enhance marketing strategies, improve customer satisfaction, and drive revenue growth.