REPORT ON CLUSTERING

Customer Segmentation Report

Objective

The primary goal of this analysis was to segment customers based on their transaction behavior, specifically looking at the total transaction value, quantity of purchases, and signup year. The results are intended to identify distinct customer groups to enable targeted marketing strategies and optimize customer relationship management.

Clustering Process

1. ***Data Preparation:***

Transactional data was aggregated by CustomerID to calculate total value and quantity of purchases.

Signup dates were converted to extract the year of signup.

Features were standardized using the StandardScaler to ensure consistent scaling across variables.

2. ***Clustering Methodology:***

Algorithm Used: KMeans clustering was employed, with the number of clusters set to 4.

Features Used:

TotalValue, Quantity, and SignupYear.

REPORT ON CLUSTERING

Evaluation Metric: Davies-Bouldin Index (DB Index) was used to assess cluster quality.

3. Results

Number of Clusters Formed:

Four distinct clusters were identified:

Cluster 0

Cluster 1

Cluster 2

Cluster 3

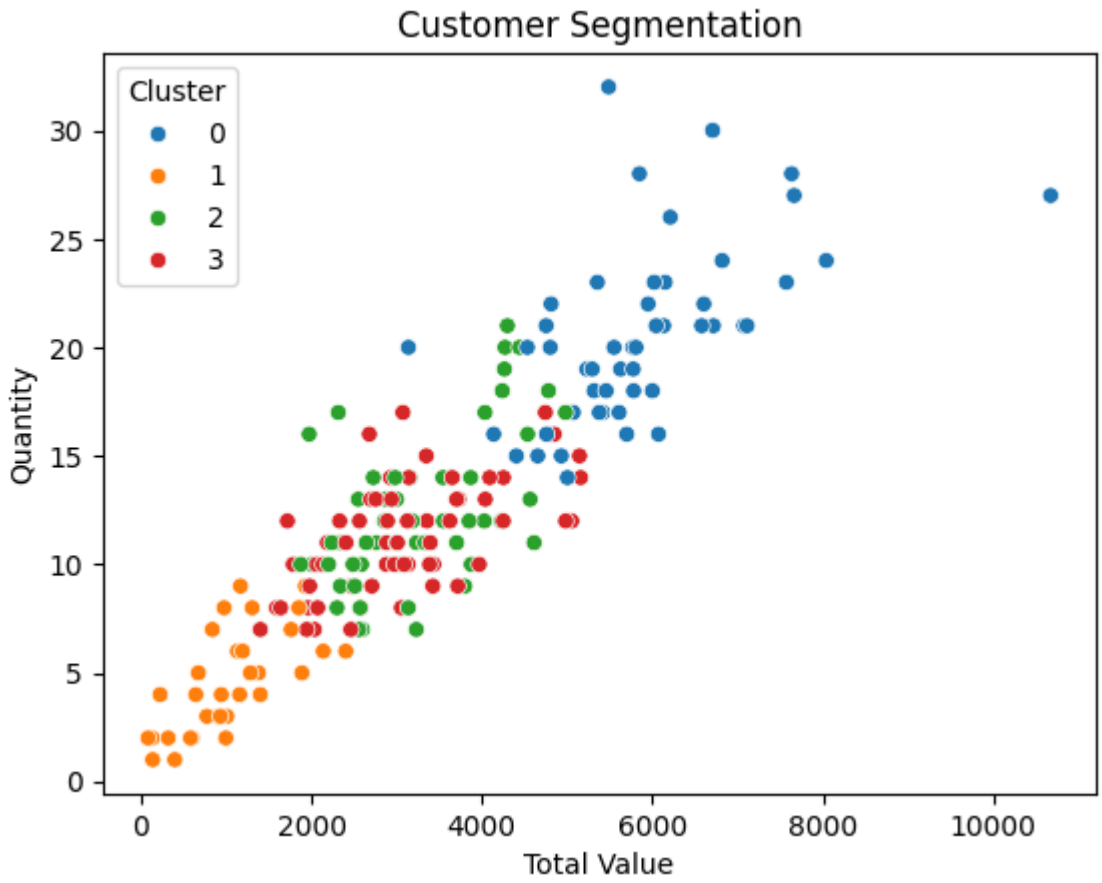
4. Evaluation Metrics:

Davies-Bouldin Index: The DB Index for the clustering was 0.8318, indicating well-defined clusters with minimal overlap. A lower DB Index reflects strong inter-cluster separation and intra-cluster cohesion.

5. ***Cluster Visualization:*** *The scatterplot of Total Value vs. Quantity (with clusters labeled) demonstrates a clear segmentation of customers into groups based on their behavior.*

REPORT ON CLUSTERING

Cluster Descriptions



Cluster 0 (Blue):

Customers in this group exhibit high total transaction value and purchase quantity, identifying them as high-value or loyal customers. This group likely consists of frequent buyers and is crucial for long-term profitability.

Cluster 1 (Orange):

Representing customers with low total value and low purchase quantity, this group likely includes casual or new customers who have minimal engagement with the business.

REPORT ON CLUSTERING

Cluster 2 (Green):

Mid-tier customers with moderate transaction values and quantities. This segment likely consists of customers with growth potential if engaged appropriately.

Cluster 3 (Red):

Similar to Cluster 2, with mid-tier transaction values, though slightly differing in terms of purchase quantity. Customers in this group might require further analysis to differentiate behaviors.

Recommendations

Conclusion

The clustering analysis effectively segmented customers into four distinct groups based on purchasing behavior. The low Davies-Bouldin Index demonstrates the quality of the clusters. These insights can be leveraged to develop tailored strategies to enhance customer engagement, improve retention, and boost overall revenue. Future work could include incorporating additional features such as product categories or geographic data to refine segmentation further.

DB Index: 0.8318

This analysis highlights the importance of data-driven decision-making in customer segmentation to drive meaningful business outcomes.

REPORT ON CLUSTERING