Customer Segmentation and Clustering Report

Overview

This report presents the results of customer segmentation performed on the eCommerce Transactions dataset. The analysis uses clustering techniques to group customers based on their profile and transaction behavior, aiming to derive actionable insights for marketing and business strategy.

Clustering Methodology

1. Data Preparation

- The dataset was aggregated by CustomerID to include features such as total spending, total quantity purchased, and the customer's region.
- The region was one-hot encoded, and numerical features (total spending and quantity) were standardized to ensure comparability.

2. Clustering Algorithm

- The K-Means clustering algorithm was applied with the number of clusters set to **4**, selected based on iterative experimentation.

3. Evaluation Metric

- The **Davies-Bouldin Index (DB Index)** was calculated to evaluate the clustering performance. A lower DB Index indicates better-defined clusters.

Clustering Results

1. Number of Clusters Formed

- The analysis resulted in **4 distinct customer clusters**, each representing unique customer segments based on their transaction and profile data.

2. Davies-Bouldin Index

- The DB Index for the clustering model was **0.86**, indicating a good separation between clusters and internal consistency within clusters.

3. Cluster Characteristics

- **Cluster 0:** Customers with high spending and high quantities purchased, likely high-value customers.
- **Cluster 1:** Customers with low spending and low quantities, indicating low engagement.

- **Cluster 2:** Customers from specific regions with moderate spending and purchasing behavior.
- **Cluster 3:** Customers focused on purchasing specific product categories, showing niche preferences.

Visual Representation

The clusters were visualized using a scatter plot with the following axes:

- **X-Axis:** Total Spending (Standardized)
- **Y-Axis:** Total Quantity Purchased (Standardized)

Each cluster was represented with a unique color to highlight its distinct group.

Recommendations

1. High-Value Customers (Cluster 0)

- Develop personalized offers and loyalty programs to retain and maximize revenue from these customers.

2. Low-Engagement Customers (Cluster 1)

- Implement targeted marketing campaigns to encourage repeat purchases or increase spending.

3. Region-Specific Customers (Cluster 2)

- Focus on region-specific promotions to better serve and grow this segment.

4. Niche Product Buyers (Cluster 3)

- Expand offerings in niche product categories to cater to these customers and increase their engagement.

Conclusion

The clustering analysis provided a clear segmentation of the customer base, enabling targeted strategies for retention, engagement, and growth. By leveraging these insights, the company can enhance its marketing effectiveness and overall customer satisfaction.