

Business Report for Task 3: Customer Segmentation / Clustering

Objective: The goal of this task was to segment customers based on both demographic and transactional data to better understand customer behavior. The clustering technique applied was KMeans, and the dataset used was a combination of customer demographic data (from `Customers.csv`) and transaction data (from `Transactions.csv`).

Clustering Analysis:

1. Clustering Algorithm and Number of Clusters: For customer segmentation, KMeans clustering was applied, using 4 clusters to categorize customers based on their total spending and purchase frequency.

2. Clustering Metrics:

- **Davies-Bouldin Index (DB Index):** The DB Index value for the clustering is **0.8595**. This value indicates that the clustering performance is good, as a lower DB Index value suggests that the clusters are well-separated and there is less overlap or scatter within clusters.
-

Cluster Characteristics:

Cluster 0 (High-Value Customers):

- **Total Spend:** 4000 to 10000
- **Purchase Frequency:** 6 to 11
- **Behavior:** These customers represent the highest-value segment, characterized by frequent and high-value purchases.
- **Insights:** This group should be targeted with loyalty programs, personalized offers, and exclusive deals to maintain and increase engagement. They contribute significantly to the revenue.

Cluster 1 (Low-Engagement Customers):

- **Total Spend:** 0 to 4000
- **Purchase Frequency:** 1 to 5
- **Behavior:** Customers in this cluster are infrequent shoppers with lower total spend.
- **Insights:** This group requires focused marketing efforts to increase their engagement. Strategies like discounts, targeted promotions, and product recommendations could help boost spending and purchase frequency.

Cluster 2 (Moderate-Value Customers):

- **Total Spend:** 2000 to 8000
- **Purchase Frequency:** 4 to 9
- **Behavior:** These customers exhibit moderate to high spending behavior and purchase frequency.
- **Insights:** This group represents an opportunity for growth. Offering them incentives or promotions could encourage them to increase their spending and purchase frequency, potentially transitioning them into high-value customers.

Cluster 3 (Balanced Customers):

- **Total Spend:** 0 to 6000
 - **Purchase Frequency:** 3 to 6
 - **Behavior:** Customers in this segment have balanced spending and frequency, neither extremely high nor low.
 - **Insights:** This cluster could benefit from targeted promotions and engagement strategies to either push them towards higher engagement or maintain their current behavior with regular, personalized offers.
-

Cluster Centers:

The KMeans algorithm identified the following cluster centers, representing the central tendency of each cluster's characteristics:

- **Cluster 1 (High Spend & Frequency):** Coordinates: [1.4797, 1.4739]
- **Cluster 2 (Low Spend & Frequency):** Coordinates: [-1.2402, -1.2770]
- **Cluster 3 (Moderate Spend & Frequency):** Coordinates: [0.4405, 0.4283]
- **Cluster 4 (Balanced Spend & Frequency):** Coordinates: [-0.3948, -0.3602]

These cluster centers provide a geometric representation of the average characteristics of each group. They help identify key segments, which can be crucial for personalized marketing efforts.

Visual Representation:

A scatter plot was generated to visually represent the clusters based on total spend and purchase frequency. The clusters were clearly distinguishable, with each group showing distinct patterns. The clusters were color-coded as follows:

- **Cluster 0:** Red
- **Cluster 1:** Blue
- **Cluster 2:** Green

- **Cluster 3: Purple**

The visualization confirms that the algorithm has successfully divided the customer base into meaningful segments based on their spending behavior and engagement.

Insights and Recommendations:

1. **High-Value Customers (Cluster 0):**

- **Strategy:** This group is highly engaged and should be rewarded with exclusive offers, loyalty programs, and early access to new products. Engaging them through personalized experiences will foster brand loyalty and sustain high lifetime value.

2. **Low-Engagement Customers (Cluster 1):**

- **Strategy:** Target this group with re-engagement campaigns, such as discounts, limited-time offers, and personalized communication. Understanding the reasons for their low engagement (e.g., lack of product variety or inadequate marketing) will be key to improving their behavior.

3. **Moderate-Value Customers (Clusters 2 and 3):**

- **Strategy:** Focus on nurturing these customers through tailored offers to move them towards higher spending and frequency. Consider cross-selling, upselling, and loyalty rewards to transition them into high-value customers.
-

Conclusion:

The customer segmentation analysis provides a clear view of different customer segments based on spending behavior and engagement. With these insights, businesses can create targeted marketing campaigns that address the unique needs and behaviors of each group, ultimately driving customer retention and revenue growth.

This segmentation model helps businesses understand their customer base better and adopt data-driven strategies for improving engagement and maximizing lifetime value.