Task 3: Customer Segmentation / Clustering

Divya Aggarwal
IIT Roorkee
Email id : d_aggarwal@hs.iitr.ac.in
Contact Number: +91-8307114763

Objective: The objective of this clustering was to segment customers based on their transaction behavior using key features like the number of transactions, total transaction value, and average transaction value. The goal was to identify meaningful customer segments that could help in targeted marketing, improving customer experience, and identifying high-value customers.

1. Number of Clusters:

Using the Elbow Method, the optimal number of clusters was determined to be **5**. The Elbow Method involves plotting the inertia (within-cluster sum of squares) for different values of k (number of clusters) and selecting the point where the inertia curve flattens, indicating that adding more clusters doesn't significantly improve the clustering quality.

2. Cluster Evaluation Metrics:

Two metrics were used to evaluate the clustering performance: the **Silhouette Score** and the **Davies-Bouldin Index**.

- **Silhouette Score:** 0.35
- **Davies-Bouldin Index:** 0.85

3. Cluster Profiles:

The clusters formed are summarized below, with key statistics such as the total number of customers, average number of transactions, average total value, and preferred product category:

Cluster	Total Customers	Avg. Total Transactions	Avg. Total Value	Avg. Transaction Value	Top Preferred Category
0	36	1.31	1.51	0.47	P006

1	31	-0.95	-1.28	-1.33	P022
2	27	-0.42	0.47	1.59	P020
3	41	-0.90	-0.71	0.18	P003
4	64	0.48	0.03	-0.41	P001

- **Cluster 0** consists of customers who make relatively frequent transactions with moderate transaction values. They seem to prefer product category P006.
- **Cluster 1** contains customers with negative transaction values and fewer overall transactions, with a preference for product category P022.
- **Cluster 2** has customers who make fewer transactions but with a higher average transaction value, preferring product category P020.
- **Cluster 3** represents customers with fewer transactions and relatively low transaction values, with product category P003 being their top choice.
- **Cluster 4** has the largest number of customers, making fewer transactions with a low average transaction value, but their top preferred category is P001.

4. Business Insights and Recommendations:

- **Cluster 0**: These customers have higher transaction frequency and value. Targeting them with loyalty programs or personalized promotions could increase retention and drive further revenue.
- **Cluster 1**: Customers in this group have low transaction values. Marketing efforts should focus on increasing their transaction volume or offering discounts to encourage higher spending.
- **Cluster 2**: Customers in this segment make fewer but higher-value transactions. Premium products or exclusive offers could be targeted to these high-value customers to increase sales.
- **Cluster 3**: These customers have fewer transactions and lower total value. Offering incentives or promotional deals might encourage them to make more purchases.
- **Cluster 4**: This group has a large customer base with lower transaction values. A strategy to drive more frequent purchases or cross-sell complementary products could be effective.

5. Conclusion:

In conclusion, the clustering analysis revealed five distinct customer segments based on their transaction behaviors. The insights from these segments can guide targeted marketing efforts, personalized promotions, and product offerings tailored to the preferences and behaviors of each cluster.