

# Customer Segmentation Insights

## 1. Data Overview

Two datasets, Customers and Transactions, were merged using CustomerID.

The merged dataset has 1000 rows and 10 columns. Feature engineering was performed to derive:

- Total Spending per customer.
- Number of Transactions per customer.
- Average Transaction Value per customer.

The final dataset used for clustering had 199 customers and 7 features.

### Example Data Points:

CustomerID | TotalSpending | NumTransactions | AvgTransactionValue

C0001	INR 3354.52	5	INR 670.90
C0002	INR 1862.74	4	INR 465.68
C0003	INR 2725.38	4	INR 681.34
C0004	INR 5354.88	8	INR 669.36
C0005	INR 2034.24	3	INR 678.08

## 2. Feature Scaling and Clustering

Features used: Total Spending, Number of Transactions, and Avg. Transaction Value.

Data was standardized using StandardScaler. K-Means clustering was applied, testing k values from 2 to 10.

The optimal number of clusters was determined using the Davies-Bouldin Index.

### Davies-Bouldin Index for different cluster values:

Clusters (k) | DB Index

2	1.0072
3	0.9578 (Best Choice)
4	1.0604
5	0.8524
6	0.8693
7	0.8553
8	0.9125
9	0.8847

3. Cluster Characteristics

Cluster	Avg. Total Spending	Avg. Transactions	Avg. Transaction Value
0	INR 6,236.28	7.89	INR 800.22
1	INR 1,125.42	2.94	INR 372.11
2	INR 4,327.40	4.11	INR 1,066.05
3	INR 2,163.80	3.05	INR 730.43
4	INR 3,516.34	6.08	INR 589.91

4. Business Insights

- Cluster 0: High-value frequent customers (Avg Spending: INR 6,236.28, Avg Transactions: 7.89).
- Cluster 1: Low-spending, infrequent buyers (Avg Spending: INR 1,125.42, Avg Transactions: 2.94).
- Cluster 2: Customers with high average transaction values (Avg Spending: INR 4,327.40, Avg Transaction Value: INR 1,066.05).
- Cluster 3: Moderate spenders with fewer transactions (Avg Spending: INR 2,163.80, Avg Transactions: 3.05).
- Cluster 4: Customers with moderate spending and frequency (Avg Spending: INR 3,516.34, Avg Transactions: 6.08).