

Customer Segmentation/Clustering Report

1.Number of Clusters Formed

- Based on the **Davies-Bouldin (DB) Index**, the optimal number of clusters is **4**.

2.Clustering Metrics

- The DB Index measures the quality of clustering by evaluating intra-cluster similarity and inter-cluster separation.
- For **k = 4**, the DB Index is **1.15**, indicating good cluster quality.

Cluster Characteristics (Key Metrics)

Cluster	Tenure (days)	Total Transactions	Total Spent	Avg Quantity	Unique Products
0	521.98	4.95	3456.25	2.55	4.73
1	548.08	4.68	3325.09	2.52	4.62
2	422.22	5.30	3311.16	2.47	5.26
3	598.41	5.15	3717.84	2.58	4.97

3. Insights from Clustering

- Cluster 0 (Moderate Spenders)

- Average tenure: ~522 days.
- Moderate spending (**\$3456**) and transaction frequency (~5 transactions).
- Represents steady customers with balanced engagement.

- Cluster 1 (Low Engagement Customers)

- Long tenure (~548 days) but lowest spending (**\$3325**).
- Fewer transactions and lower product diversity.
- Likely customers with reduced engagement over time.

- Cluster 2 (Frequent Shoppers)

- Lowest tenure (~422 days), suggesting newer customers.
- Highest number of transactions (~5.3) and most unique products purchased (~5).
- Likely frequent shoppers exploring a variety of products but spending slightly less overall.

- Cluster 3 (High Value Customers)

- Highest tenure (~598 days).
- High spending (**\$3717**) and relatively high transaction frequency (~5 transactions).
- Represents loyal, high-value customers who purchase a diverse range of products.

4. PCA Visualization

- The PCA scatter plot shows the distribution of customers in a reduced two-dimensional space.
- Clusters are well-separated, indicating that the features used for clustering effectively differentiate customer groups.