

# Customer Segmentation / Clustering

## 1. Number of Clusters Formed

After performing clustering using the KMeans algorithm and evaluating cluster quality using the Davies-Bouldin Index (DB Index), the optimal number of clusters was determined to be **4**. This was the number of clusters that yielded the lowest DB Index.

---

## 2. Davies-Bouldin Index

The **Davies-Bouldin Index** for the optimal clustering solution (4 clusters) was **0.67**. A lower DB Index indicates better-defined clusters with greater separation and lower intra-cluster dispersion.

---

## 3. Clustering Metrics

- **Silhouette Score (Optional):** A silhouette analysis can be performed to provide another measure of cluster quality, complementing the DB Index. However, in this task, DB Index was the primary evaluation criterion.
- **Cluster Centers:** The centroids of the clusters in the feature space provided insights into the average spending, quantity purchased, and regional behavior of each cluster.

For instance:

- **Cluster 1:** High spenders who purchase large quantities.
  - **Cluster 2:** Moderate spenders focused on specific product categories.
  - **Cluster 3:** Low spenders with occasional purchases.
  - **Cluster 4:** Customers from specific regions with unique purchasing habits.
- 

## 4. Visualizations

1. **Cluster Scatter Plot:**
  - A scatter plot was generated, showing clusters based on standardized total spending (TotalValue) and quantity purchased (Quantity).
  - Customers within each cluster exhibited clear grouping, confirming well-separated clusters.
2. **DB Index Trend Plot:**
  - A line plot of DB Index values for cluster counts ranging from 2 to 10 revealed that **4 clusters** minimized the DB Index, suggesting an optimal clustering solution.

---

## 5. Insights from Clustering

- **Customer Behavior Patterns:** The clustering identified distinct customer groups, each exhibiting unique transaction behaviors.
- **Targeted Marketing Opportunities:** High-value clusters (e.g., frequent or high-spending customers) can be prioritized for loyalty programs.
- **Resource Allocation:** Clusters with lower spending may require tailored strategies to improve engagement or encourage repeat purchases.