# Clustering Results Report

**Name :** Anisha Kumari

**Date:** 28/01/2025

**Task :** Report on clustering

**1. Number of Clusters Formed:**

The optimal number of clusters is determined by evaluating the Davies-Bouldin Index across cluster counts ranging from 2 to 10. The value of K corresponding to the lowest DB Index is selected as optimal.

**Optimal Number of Clusters**: {{optimal_k}}
(Note: Replace {{optimal_k}} with the actual value printed by the code.)

**2. Davies-Bouldin Index (DB Index):**

The DB Index for the optimal number of clusters (K={{optimal_k}}) is calculated as follows:
**DB Index**: {{min(db_scores):.4f}}

**3. Other Clustering Metrics:**

- **Cluster Size Distribution**:
  You can calculate and present the size of each cluster to understand the distribution of customers across segments. For example:

  ```
  print(customer_summary['Cluster'].value_counts())
  ```

- **Feature Importance for Clusters**:
  Analyzing the centroid values of each cluster can help interpret the characteristics of each segment.

**4. Visualizations:**

- **DB Index vs. Number of Clusters**:
  The plotted line graph shows the DB Index values for cluster counts from 2 to 10, allowing for the visualization of the optimal K.

- **Cluster Scatter Plot**:
  A scatter plot of the clusters (using the first two normalized features) provides a visual representation of the segmentation.

**Insights from Clustering:**

1. **High-value customers**: Identify clusters with high total spending.
2. **Frequent buyers**: Segments with a high transaction count can represent repeat customers.
3. **Regional differences**: If clusters are influenced by the Region feature, tailor marketing efforts regionally.
4. **Low-value or inactive customers**: Segments with minimal spending or fewer transactions.