Clustering Report

1. Number of Clusters Formed

- The clustering was performed using K-Means with values of k ranging from 2 to 10.
- Based on the results, **10 clusters were formed**, as it had the lowest DB Index value (indicating better clustering quality).

2. DB Index Value

The DB Index (Davis-Bouldin Index) measures the compactness and separation of the clusters. Lower values indicate better clustering quality.

Number of Clusters (k)	DB Index
2	1.5307
3	1.3728
4	1.2144
5	1.1658
6	1.1937
7	1.1214
8	1.1252
9	1.0811
10	**1.0320** (Best)

- **Final DB Index Value (k=10): 1.032**

3. Visualizations

1. **Scatter Plot of Clusters**
    - A scatter plot of PCA-reduced features for 10 clusters was visualized (attached image). Each cluster was marked with a unique color to represent its data points.
2. **Key Variable Distribution**
    - Spending patterns, product categories, and customer regions were analyzed to identify unique characteristics of each cluster.

- Example insights:
  - Cluster 0: High spenders from Region A focusing on luxury items.
  - Cluster 3: Moderate spenders purchasing household essentials.

Key Insights

- **Cluster Characteristics**:
  Each cluster shows distinct behaviors based on transaction history and customer profile. This information can be used for targeted marketing strategies.

- **Optimal Number of Clusters**:
  Based on the DB Index and other metrics, **10 clusters** are identified as optimal.

- **Improvements**:
  While the DB Index is low, further feature engineering or alternative algorithms (e.g., Hierarchical Clustering) could be explored to improve separation between clusters.