Customer Segmentation Report

1. Objective

The goal of this analysis is to segment customers into distinct groups using clustering techniques based on their profile information (from Customers.csv) and transaction data (from Transactions.csv). This segmentation can help in understanding customer behavior and tailoring marketing strategies.

2. Approach

Data Used

- **Customers.csv**: Contains customer profile information, such as demographics (e.g., age, gender, location) and other relevant features.
- **Transactions.csv**: Includes transaction details such as purchase frequency, total spending, and average purchase value.

Steps Followed:

1. **Data Preprocessing**:
    - Merged customer and transaction datasets on CustomerID.
    - Handled missing values and standardized features for uniform scaling.
    - Engineered new features (e.g., total transaction amount, average transaction value).
2. **Clustering Algorithm**:
    - Chosen Algorithm: [e.g., K-Means / DBSCAN / Hierarchical Clustering].
    - Tested the number of clusters ranging from 2 to 10.
    - Evaluated clustering quality using the Davies-Bouldin Index (DB Index), Silhouette Score, and cluster visualization.
3. **Dimensionality Reduction**:
    - Applied PCA to reduce dimensions for cluster visualization in 2D space.

Cluster Insights

Cluster Characteristics:

Below is a summary of the characteristics of each cluster:

Cluster	Key Features	Description
1	High spenders	Customers with high transaction frequency.
2	Low activity	Customers with low spending or transactions.
3	Frequent buyers	Regular customers with mid-level spending.

Visualizations:

1. **Cluster Distribution**:
    - Visualized clusters using PCA, showcasing how customers are grouped in 2D space.
2. **Feature Distributions by Cluster**:
    - Plotted histograms and boxplots to analyze how key features vary across clusters.

```
Customer Profiles:
   CustomerID  TotalValue  Quantity
0      C0001     3354.52        12
1      C0002     1862.74        10
2      C0003     2725.38        14
3      C0004     5354.88        23
4      C0005     2034.24         7
Number of Clusters: 2, DB Index: 0.6292
Number of Clusters: 3, DB Index: 0.7037
Number of Clusters: 4, DB Index: 0.7213
Number of Clusters: 5, DB Index: 0.7579
Number of Clusters: 6, DB Index: 0.8163
Number of Clusters: 7, DB Index: 0.8724
Number of Clusters: 8, DB Index: 0.7993
Number of Clusters: 9, DB Index: 0.8343
Number of Clusters: 10, DB Index: 0.7726

Optimal Number of Clusters: 2
```

4. Recommendations

1. **Marketing Strategies**:
    - **Cluster 1 (High spenders)**: Focus on loyalty programs, exclusive discounts, and premium offers.
    - **Cluster 2 (Low activity)**: Target with re-engagement campaigns and personalized offers.
    - **Cluster 3 (Frequent buyers)**: Maintain engagement with tailored incentives and cross-sell opportunities.
2. **Product Development**:
    - Leverage cluster insights to design products/services catering to specific customer groups.
3. **Resource Allocation**:
    - Allocate marketing budget based on the profitability and size of each cluster.

5. Conclusion

- A total of **X clusters** were identified, providing meaningful insights into customer behavior.
- The optimal clustering was achieved with a DB Index of **[value]**, indicating well-defined clusters.
- The segmentation can help in tailoring business strategies for improved customer satisfaction and retention.