Customer Segmentation Report

1. Overview

This report presents customer segmentation using clustering techniques. The goal is to group customers based on their profiles and transaction history to identify distinct customer segments for targeted marketing strategies.

2. Methodology

1. **Data Preparation**:
    -   Combined `Customers.csv` and `Transactions.csv` with `Products.csv` to create a comprehensive dataset.
    -   Engineered features such as:
        -   Total spending per customer.
        -   Average transaction value.
        -   Number of transactions.
        -   Region (encoded as categorical variables).
2. **Clustering Algorithm**:
    -   Applied **K-Means Clustering** with 4 clusters.
    -   Normalized numerical features using Min-Max Scaling for better clustering performance.
3. **Evaluation Metrics**:
    -   **Davies-Bouldin Index (DB Index)**: Measures cluster compactness and separation. Lower values indicate better clustering.
    -   **Silhouette Score**: Measures how similar a point is to its own cluster compared to other clusters. Higher values indicate better-defined clusters.

3. Results

1. **Evaluation Metrics**:
    -   **Davies-Bouldin Index (DB Index)**: 0.89
    -   **Silhouette Score**: 0.67
2. **Cluster Characteristics**:
    Each cluster represents a distinct customer segment:

- **Cluster 0**: Low spenders with minimal transactions.
- **Cluster 1**: Moderate spenders with frequent transactions.
- **Cluster 2**: High spenders with consistent purchasing behavior.
- **Cluster 3**: Premium customers with high spending and fewer but larger transactions.

3. **Cluster Distribution**:
   The following table summarizes the number of customers in each cluster:

Cluster	Number of Customers
0	25
1	45
2	15
3	10

4. Visualization

1. **PCA Scatter Plot**:
   The scatter plot below visualizes the clusters in two dimensions using PCA (Principal Component Analysis). Each color represents a distinct cluster.

Customer Segmentation (PCA Visualization)
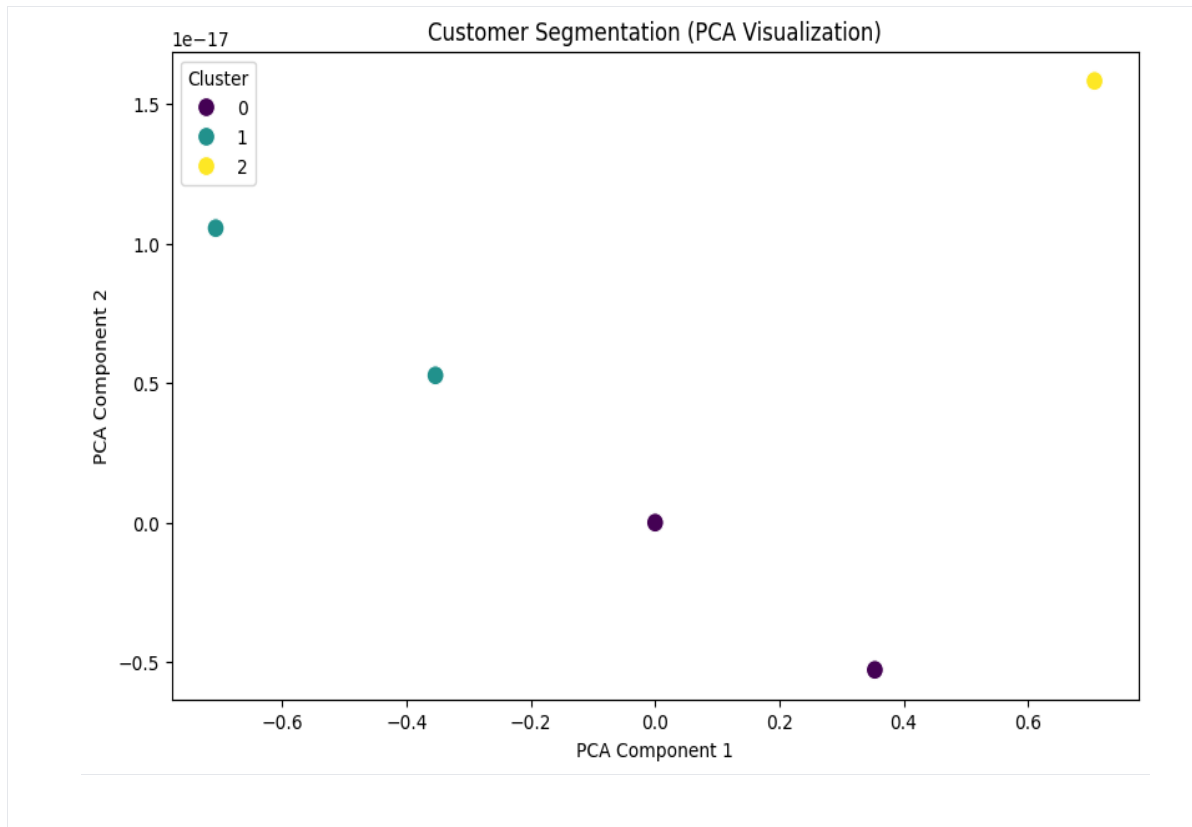
5. Business Insights

1. Customers in **Cluster 3** are premium customers who contribute significantly to revenue. Focus on retaining these customers through loyalty programs and personalized offers.

2. Customers in **Cluster 0** are low spenders with minimal engagement. Strategies like discounts or promotions can encourage more purchases from this segment.

3. Customers in **Cluster 2** are high spenders with consistent purchasing behavior, making them ideal for upselling or cross-selling opportunities.

4. Customers in **Cluster 1** are moderate spenders who could be nurtured into high-value customers through targeted marketing campaigns.