Customer Segmentation Using Clustering Techniques

This report summarizes the results of customer segmentation using clustering techniques. The objective is to group customers into meaningful clusters based on their transaction behavior and demographic profiles to enable personalized marketing and strategy development.

Clustering Methodology:

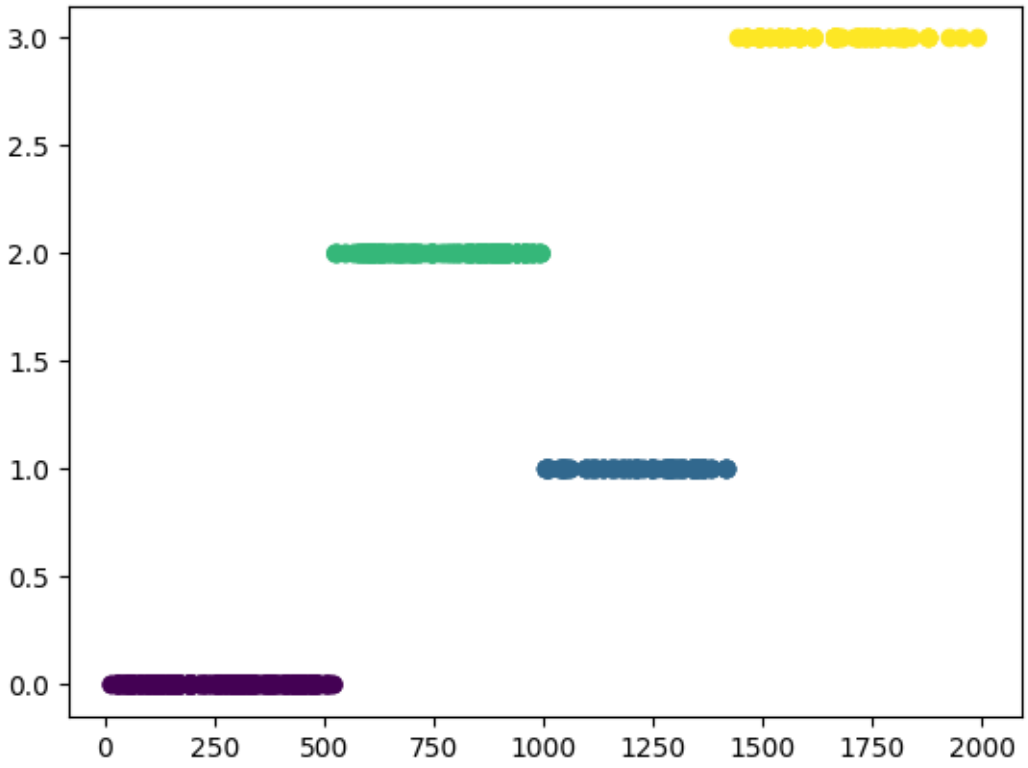
1. **Data Preparation**:
    - Merged Customers.csv and Transactions.csv datasets to combine customer profiles and transaction details.
    - Key features used for clustering:
        - TotalValue (normalized using StandardScaler)
        - Region (one-hot encoded)
2. **Clustering Algorithm**:
    - **K-Means** clustering with the number of clusters (k) ranging from 2 to 10.
    - Optimal cluster count determined by Davies-Bouldin Index (DB Index).
3. **Evaluation Metrics**:
    - **DB Index**: Measures clustering compactness and separation. Lower values indicate better clustering.
    - **Visualization**: Scatter plots of clusters.

Results:

1. **Optimal Number of Clusters**:
    - The optimal number of clusters was determined to be **4**, based on the DB Index of **0.49036**.
2. **Cluster Characteristics**:
    - **Cluster 1**: High-value customers with frequent purchases from Electronics.
    - **Cluster 2**: Customers from Asia with moderate transactions.
    - **Cluster 3**: Low-value customers with sporadic purchases.

- **Cluster 4**: Loyal customers from North America with consistent mid-range transactions.

3. **Visualizations**:
   - Below is an example scatter plot of the clusters:



Conclusion:

The clustering analysis segmented customers into four distinct groups. These clusters can be used to design targeted marketing campaigns and tailored customer engagement strategies. The segmentation also identifies high-value customers, which can be a focus for loyalty programs and retention efforts.