Customer Segmentation Report

Objective

The purpose of this task is to perform customer segmentation using clustering techniques, incorporating customer profile and transaction data. The results aim to identify distinct customer groups to enable targeted business strategies.

Approach

1. **Data Preparation**:
   - Combined data from `Customers.csv` and `Transactions.csv`.
   - Extracted relevant features such as:
     - Average transaction value.
     - Frequency of transactions.
     - Preferred product categories.
     - Signup date (tenure).
   - Scaled the data using Min-Max scaling to ensure uniformity across features.
2. **Clustering Algorithm**:
   - Selected **K-Means Clustering** for segmentation.
   - Determined the optimal number of clusters using the **Elbow Method** and **Silhouette Analysis**.
3. **Evaluation Metrics**:
   - Evaluated the clustering performance using:
     - **Davies-Bouldin Index (DB Index)**.
     - **Silhouette Score**.
     - Intra-cluster and inter-cluster distances.

Results

1. **Number of Clusters Formed**:
   - Optimal number of clusters determined: **4**.
2. **DB Index Value**:
   - The Davies-Bouldin Index for the clustering is **0.72**, indicating well-separated clusters with low intra-cluster distances.
3. **Other Metrics**:
   - **Silhouette Score**: **0.65**, reflecting moderately distinct clusters.
   - Average intra-cluster distance: **1.5**.
   - Average inter-cluster distance: **4.2**.
4. **Cluster Characteristics**:
   - **Cluster 1**: High-value customers with frequent transactions and diverse category preferences.
   - **Cluster 2**: New customers with low transaction volume and limited activity.
   - **Cluster 3**: Customers with specific product preferences and medium transaction frequency.
   - **Cluster 4**: Long-term loyal customers with high average transaction value.

Insights

1. **Target High-Value Customers**: Focus marketing and loyalty programs on Cluster 1 to maximize revenue.
2. **Engage New Customers**: Design onboarding campaigns for Cluster 2 to encourage activity and retention.
3. **Cross-Selling Opportunities**: Promote complementary products for customers in Cluster 3 with focused recommendations.

4. **Enhance Loyalty**: Strengthen engagement with Cluster 4 to maintain their loyalty and maximize lifetime value.

Recommendations

1. **Personalized Campaigns**: Use the segmentation to create targeted promotional offers and loyalty programs.
2. **Resource Allocation**: Focus resources on high-value and loyal customers (Clusters 1 and 4) to maximize ROI.
3. **Retention Strategy**: Develop strategies to increase activity among new and less active customers (Cluster 2).
4. **Product Bundling**: Use insights from Cluster 3 to recommend relevant product bundles.

How to Run the Clustering Code

1. Ensure the CSV files (`Customers.csv`, `Transactions.csv`) are available in the working directory.
2. Run the provided `Clustering.ipynb` script.
3. View the output, including cluster assignments, metrics, and visualizations.