Customer Segmentation Report

Objective

To segment customers using clustering techniques based on their transactional and demographic profiles, enabling targeted marketing strategies and improved customer engagement.

Clustering Approach

1. **Data Preprocessing:**
   - Aggregated transaction data at the customer level to calculate:
     - TotalValue: Total spending per customer.
     - Quantity: Total quantity purchased per customer.
   - Merged transactional data with customer profiles.
   - Encoded categorical features (e.g., Region).
   - Standardized numerical features (e.g., TotalValue, Quantity).
2. **Feature Selection:**
   - Selected key features for clustering by excluding non-informative attributes (e.g., CustomerID, CustomerName).
3. **Optimal Number of Clusters:**
   - Used the **Elbow Method** to identify the optimal number of clusters.
   - Based on the Elbow plot (Figure 1), **4 clusters** were chosen.
4. **Clustering Algorithm:**
   - Applied K-Means clustering with 4 clusters.
5. **Evaluation Metrics:**
   - **Davies-Bouldin Index:** 1.4337 (lower values indicate better-defined clusters).
   - **Silhouette Score:** 0.2241 (values closer to 1 indicate well-separated clusters).

Results

Cluster Characteristics:

1. **Cluster 0:**
   - Customers with average total spending and moderate transaction quantity.
   - Spread across all regions.
2. **Cluster 1:**
   - High-value customers with the highest spending and transaction volumes.
   - Likely to be targeted for loyalty programs or premium services.
3. **Cluster 2:**
   - Low-value customers with minimal spending and fewer transactions.
   - Could benefit from targeted promotions or engagement strategies.

4. **Cluster 3:**
    - Moderate-value customers with slightly higher-than-average spending.
    - Potential to upsell or cross-sell.

Visual Interpretation:

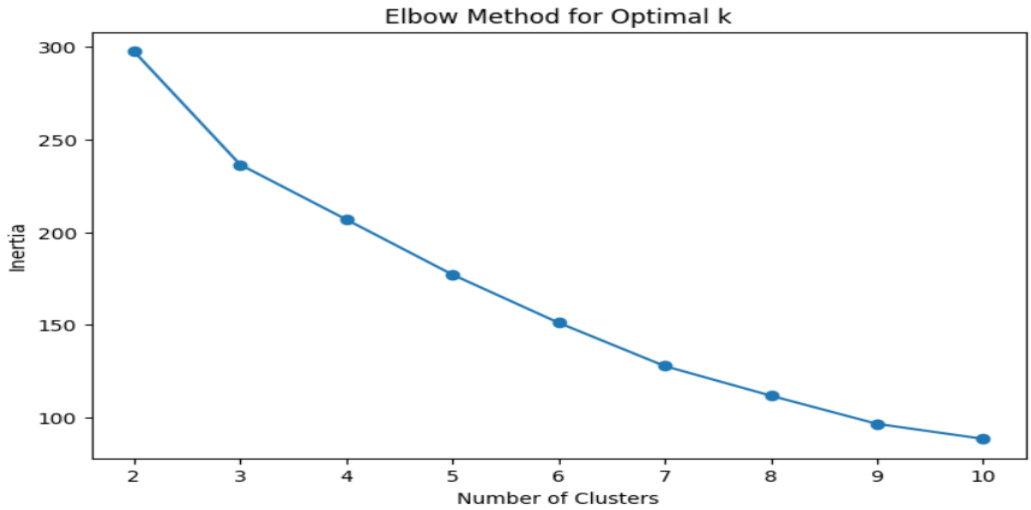
- **Figure 1: Elbow Plot**
    - Shows the decrease in inertia as the number of clusters increases, with an "elbow" at 4 clusters.
- **Figures 2 & 3: Pairplot Visualization**
    - Clear separation among clusters based on TotalValue, Quantity, and regional presence.
    - Cluster distributions highlight spending and transaction patterns.

Recommendations

1. **Cluster-Based Strategies:**
    - **Cluster 1:** Develop loyalty programs and premium services to retain high-value customers.
    - **Cluster 2:** Offer discounts and promotions to increase engagement.
    - **Cluster 3:** Use upselling techniques to boost spending.
2. **Regional Customization:**
    - Identify regional preferences for better alignment with customer needs.
3. **Monitor and Refine:**
    - Continuously evaluate cluster metrics (DB Index, Silhouette Score) to improve segmentation.

Figures:

Figure 1: Elbow Method for Optimal k



Figures 2 & 3: Pairplot Visualization of Clusters