Customer Segmentation / Clustering Report

Clustering Overview

Using the provided transactional and profile data, we performed clustering analysis to group customers into distinct segments based on their purchasing behavior and regional information.

Methodology:

1. **Features Used:**
   - Total Transactions: Number of transactions made by each customer.
   - Total Quantity: Total quantity of products purchased.
   - Total Value: Total monetary value of purchases.
   - Region: Encoded customer region.

2. **Clustering Algorithm:**
   - K-Means Clustering.

3. **Optimal Number of Clusters:**
   - Determined using the Elbow Method, resulting in **4 clusters.**

4. **Evaluation Metric:**
   - Davies-Bouldin Index (DB Index): **1.009**, indicating well-separated clusters.

Clustering Results

Cluster Summaries:

Cluster	Avg Transactions	Avg Quantity	Avg Total Value	Dominant Region
0	5.12	12.53	$3368.06	Asia
1	7.98	21.63	$6035.84	Asia
2	2.39	5.46	$1375.53	South America
3	4.89	12.43	$3412.17	South America

Key Observations:

1. **Cluster 1** represents high-value customers with significantly higher transactions, quantities, and total value. Marketing efforts targeting these customers could maximize revenue.
2. **Cluster 2** represents low-value customers, indicating an opportunity for engagement strategies to boost spending.
3. Customers from **Asia dominate Clusters 0 and 1**, while **South America dominates Clusters 2 and 3**.

Visualization

- A 2D visualization of the clusters was created using PCA (Principal Component Analysis). The clusters showed clear separations, confirming the effectiveness of segmentation.

Recommendations

1. **Target High-Value Customers:** Focus on Clusters 1 and 3 for premium offers and personalized services.

2. **Engage Low-Value Customers:** Develop engagement campaigns (e.g., discounts or loyalty programs) for Cluster 2 customers.

3. **Regional Customization:** Tailor campaigns to specific regional preferences, especially for South America and Asia.

Note: Clustering provides a data-driven approach for segmentation. Regular updates and further validation with business goals can refine these results.