Clustering Report

Overview

Clustering analysis was performed on customer transactional data to group customers with similar purchasing behaviors. The results provide insights into customer segmentation, which can guide targeted marketing efforts and operational strategies.

Results Summary

Number of Clusters Formed

- **4 clusters** were identified in the dataset.
  - Cluster assignments are indicated in the data by the `Cluster` column, ranging from 0 to 3.

Davies-Bouldin (DB) Index

- **DB Index Value: 0.8910**
  - The DB Index evaluates clustering compactness and separation. Lower values indicate better-defined clusters. A value of 0.8910 suggests reasonably well-separated and cohesive clusters.

Other Metrics and Data Insights

1. **Cluster Descriptive Statistics**:
   - **Cluster 0:** High purchasing volume, with customers having the highest `TotalValue` and `Quantity` metrics. Likely represents high-value customers.
   - **Cluster 3:** Moderate purchasing behavior, showing medium levels of `TotalValue` and `Quantity`.
   - **Cluster 1 & 2:** Represent lower purchasing tiers, possibly casual or infrequent buyers.
2. **Transaction and Product Trends**:
   - Popular products include items in the "Electronics" category, particularly high-value products like the "ComfortLiving Bluetooth Speaker."
   - Average purchase amount (`Price_x`): **$272.55**, with individual transactions ranging from **$16.08 to $497.76**.
3. **Customer Geographic Distribution**:
   - Significant representation from **Europe**, followed by **Asia** and **South America**.

Conclusion and Recommendations

1. **High-Value Customers (Cluster 0):**
   Focus on loyalty programs and exclusive offers to retain this segment.

2. **Moderate Customers (Cluster 3):**
   Offer promotions to increase their purchasing frequency and move them to the high-value segment.
3. **Low-Value Customers (Clusters 1 & 2):**
   Consider targeted campaigns to encourage repeat purchases or bundle deals to maximize order value.
4. **Future Work:**
   o Investigate temporal patterns in customer purchases to predict peak activity periods.
   o Assess customer churn risk using segmentation insights.