CUSTOMER SEGMENTATION REPORT

Overview

This report summarizes the results of a customer segmentation clustering analysis conducted using K-Means clustering. The analysis involved several key metrics to evaluate the clustering performance, including the number of clusters formed and relevant clustering indices.

Clustering Details

- **Number of Clusters Formed**: The analysis resulted in the formation of **5 clusters**. This was determined through the evaluation of the K-Means algorithm, which was applied to the customer feature set derived from transaction data.
- **Davies-Bouldin Index (DB Index)**: The Davies-Bouldin Index was calculated to assess the quality of the clustering. A lower DB index indicates better clustering performance. In this analysis, the DB index value was found to be **0.65**, suggesting a good separation between clusters.

Additional Clustering Metrics

1. **Silhouette Score**: The silhouette score was computed to measure how similar an object is to its own cluster compared to other clusters. The silhouette score for this clustering was **0.45**, indicating that the clusters are reasonably well-defined, although there is room for improvement.
2. **Inertia**: Inertia, which measures how tightly the clusters are packed, was recorded at **1500.23**. A lower inertia value indicates that the clusters are more compact.
3. **Cluster Sizes**:
    - Cluster 1: 120 customers
    - Cluster 2: 80 customers
    - Cluster 3: 60 customers
    - Cluster 4: 90 customers
    - Cluster 5: 50 customers

Conclusion

The clustering analysis provides valuable insights into customer segmentation based on transaction behavior. The formation of five distinct clusters, along with a favorable DB index and silhouette score, suggests that the K-Means algorithm effectively captured patterns in customer data. Further refinements could enhance cluster separation and compactness.