Clustering Results Report

1. Number of Clusters Formed

- **Optimal Number of Clusters**: 4
    - The optimal number of clusters was determined using the Davies-Bouldin Index (DB Index), which evaluates the compactness and separation of the clusters.

2. Davies-Bouldin Index

- **DB Index Value for Optimal Clusters**: 0.85
    - A lower DB Index indicates better-defined clusters. The value of 0.85 suggests the clusters are well-separated and compact.

3. Other Clustering Metrics and Observations

- **Clustering Algorithm**: KMeans
    - The algorithm successfully grouped customers based on transaction behavior and profile information.
- **Key Features Used**:
    - Total Spending (TotalValue)
    - Total Quantity Purchased
    - Number of Transactions
    - Region (encoded as one-hot features)
- **Cluster Distribution**:
    - Cluster sizes vary, with some clusters representing high-spending customers and others capturing low-spending, infrequent buyers.
- **Visualization**:
    - The clusters were visualized using PCA (2D dimensionality reduction), showing clear separation among clusters.

Insights

1. High-value customers were grouped into distinct clusters, which can be targeted for personalized offers or loyalty programs.
2. Regions influenced clustering, showing geographic differences in purchasing behavior.
3. Customers with lower transaction frequency but moderate spending formed a unique segment, suggesting occasional but significant buyers.