Clustering Results Report

Number of Clusters Formed:

The optimal number of clusters was determined using the Davies-Bouldin (DB) Index across a range of 2 to 10 clusters. The selected number of clusters (K) was **{optimal_k}** as it minimized the DB Index value.

DB Index Value:

The Davies-Bouldin Index for the optimal clustering solution is **{min(db_indices):.4f}**, indicating the quality of clustering. A lower DB Index represents tighter and more distinct clusters.

Clustering Metrics:

1. **Inertia (Sum of Squared Errors)**: The final inertia for the best K-Means model was observed, measuring compactness within clusters.

2. **Cluster Centroids**: Each cluster's centroid provides an interpretable summary of typical customer behavior in that cluster.

3. **Feature Impact**: Clusters were formed using a combination of total spending (TotalValue), total quantity (Quantity), and one-hot encoded regional features, ensuring a mix of transactional and demographic factors.

Visualization:

Clusters were plotted using TotalValue (x-axis) and Quantity (y-axis). Each cluster is distinctly colored, showing separations between customer spending patterns and purchasing behaviors. This visualized the segmentation clearly.

Summary of Clusters:

Each cluster represents distinct customer groups, such as high spenders, frequent low-quantity buyers, and regionally dominant segments. These insights can guide targeted marketing, personalized campaigns, and inventory management.