Task 3: Customer Segmentation / Clustering

Clustering Results:

1. **Number of Clusters**:
   Four clusters have been identified following the significance of interpretability and performance in K-Means Implementation. The clusters were determined by way of silhouette scores and Davies-Bouldin Index optimization.

2. **Clustering Metrics**:

   - **Davies-Bouldin Index (DB Index)**: Achieved a value of **1.21**, indicating moderately well-separated clusters.

   - Other metrics like inertia and silhouette score suggest a good cluster separation.

3. **Cluster Characteristics**:

   - **Cluster 1**: High-value customers with frequent purchases, ideal for loyalty programs.

   - **Cluster 2**: Customers with low frequency and low transaction values, requiring re-engagement strategies.

   - **Cluster 3**: Price-sensitive customers who buy in bulk during discounts.

   - **Cluster 4**: New customers with small transaction values, suitable for onboarding campaigns.

4. **Visualization**:

   - PCA-reduced scatterplots were used to visualize clusters in 2D, showcasing clear group separation.

   - Cluster-specific spending patterns and regions were also plotted for further analysis.