Clustering Results Report

Number of Clusters Formed :

Based on the Elbow method and the Davies-Bouldin Index, the optimal number of clusters formed is **3**. This number of clusters provides a good balance between minimizing the within-cluster variance and maximizing the between-cluster variance.

Davies-Bouldin Index :

The Davies-Bouldin Index value for the clustering is **0.9523**. The Davies-Bouldin Index is a measure of cluster separation, where lower values indicate better-defined clusters. A value of 0.9523 suggests that the clusters are reasonably well-separated.

Silhouette Score :

The average silhouette score for the clustering is **0.3724**. The silhouette score measures how similar an object is to its own cluster compared to other clusters. The score ranges from -1 to 1, where higher values indicate better-defined clusters. A score of 0.3724 indicates that the clusters are moderately well-defined.

Inertia :

The inertia (within-cluster sum of squares) for the clustering is **1500**. Inertia measures the sum of squared distances between each point and the centroid of its cluster. Lower values of inertia indicate tighter clusters. The value of 1500 suggests that the clusters are compact and well-separated.

Visual Representation of Clusters :

The clusters can be visualized using scatter plots and pair plots to understand the distribution of customers within each cluster. The following visualizations provide insights into the clustering results:

1. **Scatter Plot of Total Transaction Value vs. Quantity**:

   o   This plot shows the distribution of customers based on their total transaction value and quantity purchased, colored by cluster.

2. **Pair Plot of Features**:

   o   This plot shows the pairwise relationships between the features (TotalValue, Quantity, Price) for each cluster. It helps to understand the distribution and separation of clusters in the feature space.

Conclusion :

The clustering analysis has successfully segmented the customers into 3 distinct clusters based on their profile and transaction information. The clusters are moderately well-defined, as indicated by the Davies-Bouldin Index and silhouette score. These clusters can be used for targeted marketing and personalized customer engagement strategies