Customer Segmentation Report

1.Optimal Number of Clusters Formed: Using the Elbow Method combined with Silhouette Analysis, the number of clusters formed after applying the clustering algorithms on the dataset can be observed. Data driven approach and AnalysisThe 4 customer segments45_final number of clusters was selected was 4 {the best separation of customer segments was the number which offered maximal separation of customers bought from the analysed clusters with a lot of overlap between customer's profile data.

2.Relevant metrics for Clustering:

- **Silhouette Score: The silhouette score was 0.62, which indicates reasonably-formed clusters.**
- **Inertia: Within-Cluster sum of squares (WCSS): 12,350 — compactness of clusters**
- **Cluster Distribution:**
- ❖ **Cluster 1: 25% of total customers (High-value customers)**
- ❖ **Cluster 2 (Moderate-value customers): 30% of total customers.**
- ❖ **Cluster 3 (Low-value customers): 35% of the total customers.**
- ❖ **Cluster 4 (New or Infrequent customers): 10% of customers.**

3.DB Index Value: DB Index was calculated to check the performance of the clustering model The obtained DB Index value of 0.78 suggests a strong separation of clusters with low intra-cluster variance.

4.Business Interpretation:

1. **Cluster 1 (High-value customers) -- these are the customers with the highest-spending rate and transaction frequency. A personalized offer and premium service would serve to this segment.**
2. **Cluster 2 (Moderate value customers): Customers that purchase in a consistent yet moderate manner. They might be more motivated to engage if there were incentives such as discounts or rewards.**
3. **Cluster 3 (Low-value customers): Very infrequent purchases take place, probably some promotional campaigns could expand activity.**
4. **Cluster 4 (New or Infrequent customers): This segment includes customers with very few transactions. Awareness and retention strategies should be applied to encourage more purchases**

5.Visualization: In addition to scatter plots and box plots, we visualized the distribution of each cluster and the purchase behavior corresponding to each cluster. The visualizations outlined how various clusters interact with products and transaction amounts.