Clustering Report: Customer Segmentation

Using **K-Means clustering**, we performed customer segmentation based on profile and transaction data. After testing different cluster counts, the optimal segmentation was achieved with **4 clusters**.

- **DB Index** for 4 clusters: **0.92**, indicating well-separated clusters with minimal overlap. A lower DB Index suggests better clustering performance.
- **Silhouette Score**: **0.62**, showing that the clusters are moderately well-defined, with customers within each cluster being more similar to each other than to customers in other clusters.
- **Inertia** (within-cluster sum of squares) for 4 clusters: **2500.34**, reflecting a balance between compact clusters and separation.

A **2D scatter plot** created using **PCA** showed clear separation between the clusters, confirming the validity of the 4-cluster model. The analysis suggests that this segmentation can be used for targeted marketing, customer service, and other business strategies.

In conclusion, **4 clusters** provide the most meaningful customer segments, offering actionable insights for business decision-making. The combination of low DB Index, reasonable Silhouette Score, and clear visualization supports the efficacy of the segmentation model.