Clustering Report

1. Number of Clusters Formed:

We used **K-Means clustering** to segment the customers into **5 clusters**. This number was chosen to strike a balance between capturing key patterns in the data without splitting it into too many groups. The clusters represent different customer profiles based on their behavior and purchasing patterns.

2. Davies-Bouldin Index:

The **Davies-Bouldin Index (DB Index)** for our clustering results is **1.164**.

This index measures the separation between clusters: the lower the value, the better the separation. A score of **1.164** shows that the clusters are relatively well-separated, but there is still some overlap or room for improvement. While not perfect, the clustering does a decent job of grouping similar customers together.

3. Other Relevant Clustering Metrics:

Cluster Distribution:

The customers are fairly evenly distributed across the 5 clusters, though some clusters may have slightly more or fewer members. This suggests that the model has managed to divide the customers in a meaningful way, but further refinement could balance the clusters even more.

Silhouette Score:

While we haven't calculated it here, the **Silhouette Score** would help us understand how well-separated the clusters are. A higher score would indicate clear separation between groups, while a lower score could suggest the need for adjustments in our clustering approach.

Inertia:

Inertia measures how tight the clusters are. The lower this value, the more compact and well-defined the clusters are. It can be used to compare different numbers of clusters to find the most optimal setup.

4. Visual Insights:

The **pairplot** and **2D PCA plot** provide visual insights into how customers are grouped. These visualizations show that the clusters are somewhat distinct, with each group reflecting different purchasing behaviors and preferences.

This clustering gives us a solid foundation to understand customer segments more clearly. It can be used to tailor marketing strategies, design personalized offers, and even help improve customer experience by identifying and targeting specific customer groups more effectively.