Clustering Results Report

Objective: Perform customer segmentation using clustering algorithms on e-commerce data.

1. Number of Clusters Formed

We experimented with K-Means clustering for different numbers of clusters (from 2 to 10). The optimal number of clusters was selected based on the Davies-Bouldin Index and inertia values.

2. Davies-Bouldin Index Value

The Davies-Bouldin Index (DBI) evaluates the quality of clustering by measuring the similarity between clusters. Lower values indicate better clustering. We calculated the DBI for each number of clusters, as plotted in the following graph:

Graph 1: Davies-Bouldin Index for Different Cluster Sizes

- The lowest DBI value was observed at **X clusters** (replace "X" with the optimal number of clusters you found). This suggests that **X clusters** is the best number for forming well-separated and distinct customer segments.

3. Other Clustering Metrics

- **Inertia**: Inertia measures the within-cluster sum of squares, indicating the compactness of the clusters. A lower inertia value suggests better clustering performance. The inertia values across different cluster sizes are plotted below:

Graph 2: Inertia for Different Cluster Sizes

- As seen from the plot, the inertia continues to decrease with the increase in the number of clusters. This suggests that as we increase the number of clusters, the model improves in terms of within-cluster cohesion, though the optimal number of clusters still balances compactness and separation.

4. Insights and Actionable Strategies

Based on the clustering results, we can identify distinct customer segments, such as:

- High-spending, frequent customers

- Occasional shoppers with lower spending

- Infrequent shoppers with varied preferences