Task 3: Clustering Results Report

1. Number of Clusters Formed

Using the K-Means clustering algorithm and determining the optimal number of clusters through the Elbow Method and Silhouette Score analysis, we formed **3 clusters**. This number was chosen to maximize the cluster quality while maintaining interpretability.

2. Davies-Bouldin Index (DB Index)

The Davies-Bouldin Index for the clustering solution is **0.88**. The DB Index measures the compactness and separation of clusters, where lower values indicate better-defined clusters. This value signifies that the clusters are reasonably well-separated and compact.

3. Other Clustering Metrics

- **Silhouette Score**: The silhouette score for the clustering solution is **0.62**, indicating that the clusters are moderately well-separated. This metric assesses how similar data points within a cluster are compared to data points in other clusters.
- **Cluster Sizes**:
    - **Cluster 0**: 106 customers.
    - **Cluster 1**: 75 customers.
    - **Cluster 2**: 18 customers.
- **Cluster Characteristics**:
    - **Cluster 0**: Moderate spending and purchase frequency, with slightly recent transactions.
    - **Cluster 1**: High spending and purchase frequency, with very recent transactions.

- **Cluster 2**: Low spending and purchase frequency, with older transactions.

4. Visualizations

Clusters have been visualized using scatter plots to show relationships between key features such as:

- Total Spending vs. Recency
- Purchase Frequency vs. Total Spending

These visualizations provide a clear understanding of how customers are distributed across clusters and highlight their behavioural patterns.