# Task 3: Customer Segmentation / Clustering

**Customer Segmentation Clustering Report**

**1. Number of Clusters Formed:**

Using the Elbow Method, we determined the optimal number of clusters. Based on the WCSS curve, we selected **4 clusters** for the K-Means algorithm.

**2. Clustering Metrics:**

- **Davies-Bouldin Index:** The calculated DB Index value is **Davies-Bouldin Index: 1.0041069554457889**. A lower DB Index indicates better clustering quality.
- **Silhouette Score:** The calculated silhouette score is **Silhouette Score: 0.3139253055723817**. A higher silhouette score (closer to 1) indicates well-separated clusters.

**3. Cluster Characteristics:**

Each cluster represents a different customer segment based on:

- **Total Spend** (amount spent by a customer across transactions)
- **Transaction Count** (number of transactions per customer)
- **Recency** (days since the last transaction)

| Cluster | Average Total Spend | Average Transaction Count | Average Recency |
|---|---|---|---|
| 0 | -0.756213 | -0.769854 | -0.111364 |
| 1 | 1.401367 | 1.387728 | -0.505548 |
| 2 | -1.402035 | -1.611763 | 2.744307 |
| 3 | 0.192260 | 0.248790 | -0.136226 |

**4. Visualization Summary:**

- **Elbow Method Plot** was used to determine the optimal number of clusters.
- **Pairplot Visualization** was created to observe the distribution of clusters across different attributes.
- **Cluster Scatter Plot** showed the separation of clusters in a 2D space.

**5. Conclusion:**

The clustering results successfully segmented customers into meaningful groups. These clusters can help businesses tailor marketing strategies, identify high-value customers, and enhance customer engagement strategies.