

## **TASK-3**

### **Customer Segmentation/Clustering**

#### **1. Number of Clusters Formed:**

- The clustering was performed with  $k=4$  (as specified  $k$  is between 2 and 10).
- Customers have been grouped into 4 distinct clusters based on the features Region, TotalSpending, and AvgTransactionValue.

#### **2. Clustering Evaluation:**

- The **Davies-Bouldin Index (DB Index)** is used to evaluate the quality of the clusters. Lower values indicate better clustering.
- Provide the DB Index value you obtained from the script. If you need help interpreting it, let me know!

#### **3. Insights from Clustering:**

- **Cluster 0 (Blue):** Represents a specific group of customers with unique characteristics. Further profiling is possible based on customer spending or transaction history.
- **Cluster 1 (Orange):** Another segment based on distinct spending and regional patterns.
- **Cluster 2 (Green):** Indicates customers with average spending levels.
- **Cluster 3 (Red):** May represent either low or high spenders, depending on the total features contributing to the cluster.

#### **4. Visualization:**

- The provided scatter plot (PCA reduced to 2D) demonstrates clear segmentation between the clusters.

- Each point represents a customer, and their position along the principal components illustrates the variability captured in the clustering.

## **5. Next Steps:**

- Based on the clusters, further business strategies can be defined, such as targeted marketing for high-value clusters or customized offers for low-spending regions.