DATA SCIENCE ASSIGNMENT :

TASK 3.

A report on your clustering results, including:

- The number of clusters formed.
- DB Index value.
- Other relevant clustering metrics.

This report details the clustering analysis performed using the provided customer and transaction data. The goal was to segment customers into distinct groups based on their transactional behaviour and profile information.

First task is to merge the data of customers and transactions , and later to i have picked k means clustering algorithm.

NUMBER OF CLUSTERS USED.

And the number of clusters is 2, which gives the low DB index.

DB INDEX VALUE:

And the value of db index is 0.95

METRICS INVOLVED:

Cluster Distribution: The number of customers in each cluster varied, highlighting differences in customer behavior.

Key Features Influencing Clustering: Total spending and transaction count were primary drivers of segmentation. Regional and demographic data also contributed to defining clusters

To reduce dimensionality, **Principal Component Analysis (PCA)** was applied. A scatterplot of clusters in 2D PCA space showed good

Separation between the groups.