Subhasmita Khuntia - Customer Segmentation Report

Clustering Algorithm: KMeans

Features Used:

- Total transaction value

- Quantity purchased

- Average product price

- Region (encoded)

- Category (encoded)

Preprocessing:

- Categorical encoding using pd.get_dummies()

- Standard scaling using StandardScaler()

Best Number of Clusters (K): 4

Lowest DB Index Achieved: 0.9231

PCA Visualization:

- A 2D scatter plot was generated using PCA to reduce feature space.

- Clear separation between customer segments is visible, indicating effective clustering.

Business Use:

The segments can be used for personalized marketing, product recommendations, and identifying high-value

customers.