Report on Customer Segmentation Using Clustering Techniques

Objective

This report explains how I used clustering techniques to segment customers using their customer information and transaction data.

The goal is to identify different types of customers so the company can create better marketing strategies, offer personalized deals, and improve customer relationships.

Dataset Description

I used two datasets:

- **Customers.csv**: Contains customer details like ID and profile information.
- **Transactions.csv**: Contains purchase history, including transaction amounts and frequency.

I combined these datasets using the CustomerID column. Then, I focused on key details like total spending, number of purchases, and quantity of items bought. I standardized the data to ensure that all features were on the same scale before clustering.

Clustering Approach

I tested different clustering methods like K-Means, Agglomerative Clustering, and DBSCAN. To find the best number of clusters, I used the **Elbow Method**, which suggested 4 clusters.

I decided to use **K-Means Clustering** because it is simple and works well for this kind of data. To help visualize the results, I used **Principal Component Analysis (PCA)** to reduce the data to two dimensions.

Evaluation Metrics

To measure how good the clustering was, I used:

- **Davies-Bouldin Index (DBI)**
- **Silhouette Score**

Results

- **Number of Clusters:** 4
- **Davies-Bouldin Index:** 0.8650620583623065
- **Silhouette Score:** 0.37356460556541055

Visualization of Clusters

I used PCA to create a scatter plot of the clusters and Elbow Method for find the optimal number of cluster. The visualization clearly showed four separate groups, helping to understand the differences between them.

Conclusion

This segmentation helps the business understand customer behavior and improve marketing strategies. By focusing on different customer needs, the company can boost sales and build better relationships with customers.