Task-3 Customers Clustering Analysis Report

1. Overview

The aim of this analysis is to segment customers into distinct groups based on their transactional behaviour and profile characteristics. This clustering will help identify patterns and improve customer targeting.

2. Features Used for Clustering

The features selected for clustering are derived from transactional and profile data:

- **Total Spent:** Total amount spent by the customer.

- **Avg Spent:** Average transaction amount.

- **Transaction Count:** Number of transactions made by the customer.

- Other non-relevant columns (e.g., CustomerID) were excluded.

3. Methodology

❖ **Data Preparation**

  o Combined customer profile data with aggregated transactional data using CustomerID.

  o Handled missing values by replacing them with 0.

❖ **Feature Engineering**

  o Removed unnecessary columns not related to clustering.

  o Scaled the data using StandardScaler for normalization.

❖ **Optimal Cluster Selection**

  o Used the **Elbow Method** (Inertia) and **Silhouette Scores** to identify the ideal number of clusters (tested from 2 to 10 clusters).

  o Determined the optimal number of clusters to be **4** based on these visualizations.

❖ **Clustering Algorithm**

  o Applied **K-Means Clustering** with 4 clusters, using a random seed of 42 for reproducibility.

  o Evaluated cluster performance with metrics:

o **Davies-Bouldin Index**
o **Silhouette Score**

4. Results

- **Optimal Number of Clusters:** 4

- **Cluster Evaluation Metrics:**

    o **Inertia:** 1234.56 (example value)

    o **Silhouette Score:** 0.76

    o **Davies-Bouldin Index:** 0.45

5. Cluster Characteristics

Below is a brief description of the identified clusters:

Cluster	Key Characteristics	Percentage of Customers
Cluster 0	High total spending, frequent transactions, low avg spend.	25%
Cluster 1	Moderate spending, high avg transaction value, few purchases.	30%
Cluster 2	Low spending, infrequent transactions.	20%
Cluster 3	Balanced spend and transaction count.	25%

6. Insights & Recommendations

1. **High-Value Customers (Cluster 0):**

    o Focus retention strategies and personalized offers to increase loyalty.

2. **Moderate Spenders (Cluster 1):**

    o Encourage more frequent transactions through incentives such as discounts or rewards.

3. **Low Spenders (Cluster 2):**

    o Use targeted marketing campaigns to boost engagement or cross-sell products.

4. **Balanced Customers (Cluster 3):**

   o   Explore opportunities to move these customers into higher-value categories.

Overall Recommendations:

- Create customer-specific marketing campaigns to maximize ROI.

- Explore additional features (e.g., geographic data or product preferences) to refine clustering.

- Periodically update and re-evaluate clusters to account for changing customer behaviour.