TASK 3: Customer Clustering Results Report

1. Objective

The primary objective of this analysis is to segment customers based on their transaction patterns, demographic profiles, and spending behavior using clustering techniques. This allows businesses to better understand and target different customer segments for personalized marketing and strategic decision-making.

Clustering

1. **Algorithm:** K-Means Clustering
    - Number of clusters **(k)**: **4** (chosen based on domain knowledge and experimentation).
    - Random seed: 42 (for reproducibility).
2. **Cluster Labels:**
    - Each customer was assigned a cluster label from 0 to 3.
3. **Cluster Evaluation Metrics:**
    - **Davies-Bouldin Index:** 1.1926 (lower values indicate better-defined clusters).
    - **Silhouette Score:** 0.3197 (measures how similar a customer is to its own cluster compared to others).
4. **Dimensionality Reduction for Visualization:**
    - PCA (Principal Component Analysis) reduced the dataset to 2 components for effective visualization of clusters.

Number of Clusters Formed

The dataset was segmented into **4 distinct clusters**. Each cluster represents a group of customers with similar spending behaviors and demographics.

Cluster Evaluation

- **Davies-Bouldin Index:** 1.1926
    - This indicates moderately well-defined clusters.
- **Silhouette Score:** 0.3197
    - This score shows that the clusters are distinguishable but have scope for improvement in separability.

Key Characteristics of Clusters

Cluster	Total Customers	Key Characteristics
0	X	High spenders, frequent transactions, prefer specific regions
1	X	Moderate spenders, diverse purchase patterns
2	X	Low spenders, occasional buyers, limited transaction history
3	X	Region-specific buyers, average transaction values

Cluster Visualization

The PCA visualization below shows the clustering results in a two-dimensional space. Each point represents a customer, and colors indicate different clusters.