Class Activity – Manual Clustering Exercise

Objective:

You will group data points into clusters manually based on provided attributes.

Dataset: Customer Spending Analysis

Below is the dataset for analysis. Each row represents a customer, with attributes indicating their behavior and spending.

Customer_ID	Age	Income ($K)	Spending_Score (1-100)	Frequency_of_Purchase (per month)
1	25	50	80	12
2	34	60	70	9
3	45	90	30	4
4	23	20	95	15
5	40	100	25	3
6	29	45	75	10
7	38	70	50	6
8	31	55	65	8
9	50	120	20	2
10	21	30	85	14

Hidden Pattern:

1. Customers can be clustered based on **spending behavior**:
   - High spenders with frequent purchases.
   - Moderate spenders with average income and frequency.
   - Low spenders with high income but infrequent purchases.
2. Clusters can also emerge based on **age groups** and **income levels**.

---

1. **Formulate Similarity Metrics**:
   - Define criteria for similarity (e.g., age proximity, income range, spending score).
   - Consider which attributes are most relevant for clustering.
2. **Group the Data**:
   - Manually group the customers into 3–4 clusters based on their similarities.
   - Use the given dataset and create clusters visually or by calculating differences between values.
3. **Visualize the Clusters**:
   - Create a table or scatter plot (on paper or a spreadsheet) to show which customers belong to which clusters.
4. **Summarize Findings**:
   - Write a short explanation of the clustering criteria and insights gained from the clusters.

Example Outcome (Clustering by Spending Behavior):

- **Cluster 1**: Frequent high spenders (Customers 1, 4, 10).
- **Cluster 2**: Moderate spenders (Customers 2, 6, 8).
- **Cluster 3**: High-income low spenders (Customers 3, 5, 9).
- **Cluster 4**: Average spenders (Customers 7).

Response to the Practice Activity: Manual Clustering

To create clusters, I considered the following attributes:

- **Spending Score**: Higher scores indicate more active spending behavior.
- **Income**: Helps identify high-earning but low-spending customers.
- **Frequency of Purchase**: Indicates how often a customer makes purchases, complementing the spending score.

2. Clustering Process

- I manually grouped the customers into clusters based on similarities in **spending behavior, income, and purchase frequency**.
- I calculated approximate similarity by comparing the ranges of values for each attribute.

3. Clusters

Based on the analysis, I identified **4 clusters**:

Cluster 1: Frequent High Spenders

These customers frequently purchase and have high spending scores.

- **Customers**: 1, 4, 10
- **Characteristics**:
    - Spending Score: 80–95
    - Frequency: 12–15 purchases/month
    - Income: $20K–$50K

Cluster 2: Moderate Spenders

These customers have moderate spending scores and purchase frequency.

- **Customers**: 2, 6, 8
- **Characteristics**:
    - Spending Score: 65–75
    - Frequency: 8–10 purchases/month

o Income: $45K–$60K

Cluster 3: High-Income, Low Spenders

These customers earn high incomes but have low spending scores and infrequent purchases.

- **Customers**: 3, 5, 9
- **Characteristics**:
    - o Spending Score: 20–30
    - o Frequency: 2–4 purchases/month
    - o Income: $90K–$120K

Cluster 4: Average Spenders

These customers fall in the middle range for income, spending scores, and frequency.

- **Customers**: 7
- **Characteristics**:
    - o Spending Score: 50
    - o Frequency: 6 purchases/month
    - o Income: $70K

---

4. Insights

1. **Frequent High Spenders (Cluster 1)**:
    - o These customers are younger and likely more engaged with the store.
    - o Tailored loyalty programs or discounts could encourage continued engagement.
2. **High-Income, Low Spenders (Cluster 3)**:
    - o Although these customers have significant income, they are not frequent shoppers.
    - o Marketing efforts could target them with premium product offerings or exclusive promotions.
3. **Moderate Spenders (Cluster 2)**:
    - o This group represents the steady, consistent customer base.
    - o Regular updates and promotions would likely maintain their loyalty.
4. **Average Spenders (Cluster 4)**:
    - o A smaller group that might require closer observation to understand specific motivations for their shopping habits.