Kiddocare

Test Case: Customer Segmentation Analysis

Problem Statement

As a Business Analyst at Kiddocare, you have been tasked with analyzing customer data to identify distinct segments based on their behavior and purchase patterns. The goal of this segmentation is to help the marketing team create targeted campaigns tailored to each customer group, thereby improving engagement and increasing conversion rates.

Data Description:

The dataset provided contains the following columns:

- **Customer ID**: Unique identifier for each customer.
- **Age**: Age of the customer.
- **Gender**: Gender of the customer (M/F).
- **Annual Income**: Annual income of the customer in RM.
- **Spending Score**: A score assigned to each customer based on their purchasing behaviour, ranging from 0 to 100.

Dataset :

https://docs.google.com/spreadsheets/d/1QwXoxeq_MEQEAQYIArl80LAnu2kyc3GU/edit?gid=284143541#gid=284143541

Test Steps :

1. **Import the Dataset:**
   - Import the provided dataset into your preferred data analysis tool (e.g., Python, R, Excel, etc.).

2. **Data Exploration:**
   - Perform an initial exploration of the dataset:
     - ✓ Check for any missing values and handle them appropriately.
     - ✓ Examine the data types of each column.
     - ✓ Identify and address any anomalies or outliers.

3. Data Visualization:
- Create visualizations to understand the distribution of key variables:
  - ✓ **Age Distribution**: Use histograms or box plots to visualize the age distribution.
  - ✓ **Gender Distribution**: Create a bar chart to show the proportion of male and female customers.
  - ✓ **Annual Income**: Use a histogram to display the distribution of annual income.
  - ✓ **Spending Score**: Visualize the distribution of spending scores using a histogram or box *plot.*

4. Customer Segmentation (Clustering Analysis):
- Apply a suitable clustering algorithm to segment customers based on their behavior and purchase patterns:
  - ✓ Use **K-means clustering** or **Hierarchical clustering** to group customers.
  - ✓ Normalize the data if necessary to ensure fair clustering.

5. Determine Optimal Number of Clusters:
- Use methods like the **Elbow method** or **Silhouette score** to determine the optimal number of clusters.

6. Visualize Clusters:
- Create visualizations to analyze the clusters:
  - ✓ Use scatter plots or pair plots to visualize clusters based on combinations of Age, Annual Income, and Spending Score.
  - ✓ Consider using color coding to differentiate between clusters.

7. Interpret the Results:
- Analyze and interpret the characteristics of each identified customer segment:
  - ✓ Describe the key characteristics of each segment, such as age range, income level, spending behavior, and gender distribution.
  - ✓ Identify patterns in purchasing behavior and any unique preferences of each segment.

8. Provide Recommendations:
- Based on the analysis, provide actionable recommendations to the marketing team:
  - ✓ Suggest specific marketing strategies or campaigns for each segment.
  - ✓ Recommend personalized offers, promotions, or communication approaches that align with the preferences and behavior of each customer group.

Evaluation Criteria

- **Data Analysis Techniques**: Accuracy and quality of the techniques used in data cleaning, exploration, and clustering.
- **Findings**: Clarity and coherence in the analysis findings, including the interpretation of clusters and insights into customer behavior.
- **Communication**: Effectiveness in communicating insights and recommendations to stakeholders.
- **Visualization**: Proficiency in using data visualization tools and techniques to represent data and clusters.
- **Attention to Detail**: Accuracy in interpreting data, identifying patterns, and ensuring data quality.
- **Knowledge of Clustering Algorithms**: Understanding of clustering techniques and their application in customer segmentation.
- **Creativity and Thoughtfulness**: Innovation in providing targeted recommendations that align with the business goals.

Due date of submission: 28th August 2024, 3pm.