Assignment 1

Due on 2020-09-12, 23.59 IST.

This data set has been used for testing the assignment.

The data set contains the following columns:

- Age
- Sex
- PWD
- Income
- Education
- Occupation
- Region
- Marital Status
- Employment Status
- Spending Class
- Credit Score
- Current Employment
- Credit Limit
- Credit Utilization
- Credit History
- Credit Utilization Rate
- Credit Limit (FICO)
- Credit Utilization (FICO)
- Credit History (FICO)
- Credit Utilization Rate (FICO)

The objective is to predict whether a customer will default on their credit card.

1. Which command in Python, out of the following, will we use for accessing the second row of an array named "data" if we want to retrieve the values from columns 2 to 5?

   a) data[2].values()
   b) data[2:5].values()
   c) data[2].iloc[:, 2:5]
   d) data[2:5].iloc[:, 2]
   e) data[2:5].iloc[:, 2:5]

   Rs. the power is incorrect.
   Accepted Answers:
   - (3)
   - (5)

2. If def(f(data):
   Z = max(data) - min(data)
   return Z)

   The above defined data function in Python programming, will calculate the:

   a) range
   b) variance
   c) median
   d) sum of values
   e) maximum

   Rs. the power is incorrect.
   Accepted Answers:
   - (3)
   - (5)

3. Compile the following statements:
   Statement A: With this, we can pass in the negative values.
   Statement B: With this, we can pass in the negative values.
   A and B are correct
   Both are false
   A is correct B is false
   B is correct A is false

   Rs. the power is incorrect.
   Accepted Answers:
   - (1)
   - (3)

4. For getting list of all 4th row of a table named "df" in Python programming, we can write:
   a) df[3].iloc[:, 4]
   b) df.iloc[3, 4]
   c) df.iloc[4, 3]
   d) df.iloc[3, 4].values
   e) df.iloc[3, 4].values

   Rs. the power is incorrect.
   Accepted Answers:
   - (1)
   - (3)
   - (5)

5. Box plots are used for classification of:
   a) Categorical data
   b) Continuous data
   c) Numerical data
   d) None of the above

   Rs. the power is incorrect.
   Accepted Answers:
   - (2)
   - (3)

6. The frequency Polygon and bar Chart both are used for:
   a) Categorical data
   b) Continuous data
   c) Both a & b
   d) None of the above

   Rs. the power is incorrect.
   Accepted Answers:
   - (1)
   - (2)

7. If two groups of numbers have the same mean, then:
   a) Their standard deviations must be equal
   b) Their standard deviations must also be equal
   c) Their mode must also be equal
   d) None of the above

   Rs. the power is incorrect.
   Accepted Answers:
   - (1)

8. Median is not applicable to:
   a) Males
   b) Females
   c) Both
   d) None of the above

   Rs. the power is incorrect.
   Accepted Answers:
   - (3)

9. In relevant data, the least measure of central tendency is:
   a) Mean
   b) Median
   c) Mode
   d) None of the above

   Rs. the power is incorrect.
   Accepted Answers:
   - (1)

10. The 5th percentile of a normally distributed data set will represent its:
    a) Mean
    b) Mode
    c) Both
    d) None of the above

    Rs. the power is incorrect.
    Accepted Answers:
    - (2)