Indian Maritime University

(A Central University, Govt of India)

End Semester Examinations – June 2023

Programme Name: B Tech (ME)

Semester: III

Subject Code: UG11T4305

Subject Name: Statistics and Data Analysis Using Python and R

Date: 17.05.2023 Max Marks: 70

Duration: 03 Hrs Pass Marks: 35

General Instructions

(i) All Sections (A, B & C) are to be attempted.

(ii) Options, if any, are specified in the respective section.

(iii) Scientific calculator is permitted.

Section A

MCQs –All Questions are Compulsory. (10×01 mark = 10 Marks)

1. Which one of the following variables is a nominal variable?
   a. Age of a person.
   b. Gender of a person: male or female.
   c. Weight of the person
   d. Height of the person
2. Which of the following declarations is incorrect in python language?
   a. xyzp = 5
   b. x y z p = 5 6 7 8
   c. x,y,z,p = 5, 6, 7, 8
   d. x_y_z_p = 5
3. Which of the following are modules/libraries in Python?
   a. Numpy
   b. Pandas
   c. Matplotlib
   d. All of the above
4. which of the following is used to define a block of code in Python Language?
   a. Indentation
   b. Key
   c. Brackets
   d. None of the above

5. Which of the following brackets are used to create a list in Python?
   a. []
   b. {}
   c. ()
   d. set()
6. Which of these is not a core data type?
   a. list
   b. dictionary
   c. Tuple
   d. class
7. The formula to calculate Arithmetic mean of the ungrouped data is
   a. $\frac{\Sigma x_i f_i}{\Sigma f_i}$
   b. $l+\frac{h}{f}(c.f.-f)$
   c. $l+\frac{h}{f}(c.f.+f)$
   d. None of the above
8. Given the following bivariate probability distribution,

Y \ X	-1	0	1
0	1/15	2/15	1/15
1	3/15	2/15	1/15
2	2/15	1/15	2/15

Then the marginal distributions of y is

a.

y	-1	0	1
P(y)	2/5	1/3	4/15

b.

y	0	1	2
P(y)	4/15	6/15	5/15

c.

y	-1	0	1
P(y)	3/15	2/15	1/15

d.

y	0	1	2
P(y)	2/15	9/15	4/15

9. Which of the following is true for a vector in R?
   a. It is a homogeneous 1-dimensional data structure
   b. It is a heterogeneous 1-dimensional data structure
   c. It is a homogeneous 2-dimensional data structure
   d. It is a heterogeneous 2-dimensional data structure

10. Which function is used to take input from the user in Python?
    a. scanf()
    b. input()
    c. cin>>
    d. None of the above

Section B

Answer all the Questions. **(05×02 marks=10 Marks)**

11. Explain if-else conditional statement with a flowchart.

12. What are lists and tuples? What is the key difference between the two?

13. The data given below show the marks of 15 students out of 100.

    40, 62, 25, 2, 50, 80,73,59, 59, 60,85, 99,49,50,39

    Find the Median.

14. Write a R program to print the output as a multiplication of two vectors

    x = (10, 20, 55) and y = (10, 5, 10).

15. Explain what is With () and By () function in R is used for?

Section C

Answer any 5 of the following 7 questions. **(05×10 marks=50 Marks)**

16. a) What is the need for data visualization? Explain any two data-visualizing libraries in python. (5)

    b) Explain key features of Pandas library. (5)

17. a) Explain the following terms in Python with suitable examples. (5)

    i) Comment
    ii) Reserve Words

    b) Write a python program for the creation of Data Frame using List.
    A list containing the values 'Python', 'Pandas', 'Seaborn', and 'NumPy'. (5)

18. a) Let X_1 and X_2 have the Joint probability distribution $f(x_1, x_2)$ of X_1 and X_2

		x_1		
		0	1	2
x_2	0	0.1	0.4	0.1

	1	0.2	0.2	0

Find $P(X_1+X_2 > 1)$ (5)

b) Explain read, write, and append modes in file handling with examples. (5)

19. Write a programme to calculate mean and standard deviation of continuous frequency distribution.

List1 contains observation and it is a list of list represented as [[10,20], [20,30], [30,40]].
List2 contains frequency and it is a list represented as [5, 7, 8].
List2[i] is the frequency of List1[i] observation. (10)

20. a) Explain for and while loop in Python. (5)

b) Write the output of the following R program. (5)

```
x = c(10, 20, 30)
y = c(20, 10, 40)
print("Original Vectors:")
print(x)
print(y)
print("After adding two Vectors:")
z = x + y
print(z)
```

21. (a) Explain different NumPy attributes. (3)

(b) Given two list of equal length find summation, subtraction, multiplication and division using NumPy. (4)

(c) Write the programme to create table from given dictionary.

```
data = {'product' : ['pen', 'pencil', 'notebook', 'marker'],
        'quantity': [10, 20, 12, 8]
        }
```

(3)

22. Find the median for the following data. (10)

Class Interval	110-120	120-130	130-140	140-150	150-160	160-170	170-180	180-190	190-200
Frequencies (f)	6	25	48	72	116	60	38	22	3