Indian Maritime University

(A Central University, Govt of India)

Supplementary Examinations – March/April 2025

Programme Name: B Tech (ME)

Semester: III

Subject Code: UG11T4305

Subject Name: Statistics and Data Analysis using Python and R

Date: 26.03.2025 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 respective section.
(iii) Scientific calculator is permitted.

Section A

MCQs -All questions are compulsory [10X01 mark = 10 marks]

1. Which module is used for data visualization in Python?

a) NumPy
b) Pandas
c) Matplotlib
d) Scikit-learn

2. List AL is defined as follows: AL=[1,2,3,4,5], which of the following statements doesn't remove the middle element 3 from it so that the list AL is equal to [1,2,4,5]?

a) del AL[2]
b) AL[2:3]=[ ]
c) AL[2:2]=[ ]
d) AL.remove(3)

3.Decimal values are referred as _____ data types in R.

a) Numeric
b) Character
c) Integer
d) Lists

4.Consider the code given below:

```
b=100
def test(a):
        _____#Missing statement
        b=b+a
        print(a,b)
test(10)
print(b)
```

Which of the following statements should be given in the blank for #Missing statement, if the output produced is 110?

a) global a
b) global b=100
c) global b
d) global a=100

5. Which of the following various inputs can be used to create pandas DataFrame.

a) Lists, dict
b) Series
c) Numpy ndarrays and Another DataFrame
d) All of the above mentioned

6.What is the output of the following code?

```
import numpy as np
a = np.array([1, 2, 3])
b = np.array([4, 5, 6])
c = a + b
print(c)
```

a) [1, 2, 3, 4, 5, 6]
b) [[1, 4], [2, 5], [3, 6]]
c) [5, 7, 9]
d) Error

7.In absolute Measure Of Dispersion, the Square root of variation is known as _____

a) Mean Deviation
b) Quartile Deviation
c) Variance
d) Standard Deviation

8.When we open file in append mode the file pointer is at the _____ of the file.

a) end
b) beginning
c) anywhere in between the file
d) second line of the file

9. In which scale of measurement, classification, order and equality of units are ensured?

a) Nominal
b) Ordinal
c) Ratio
d) Interval

10. Which of the following is not a valid set operation in python?

a) Union
b) Intersection
c) Difference
d) None of the above

Section B

Five Questions of 02 Marks each

11. Write a R program to print the output for any two mathematical operations of two vectors.

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

13. Explain dir() function in python.

14. The data given below show the number of customers visited a store in a fort- night.

40, 62, 25, 2, 50, 80,73,59, 59, 60,85, 150,49,50,39

Find the inter quartile range.

15. What are the features of NumPy?

Section C

Seven Questions of 10 Marks each of which any 05 questions to be answered.

16.a) Explain in detail while and for loops in python with suitable examples. 5 Marks

b) List down types of operators in Python. Explain relational operators. 5 Marks

17.a)Write a python program for the creation of 1-D & 2-D NumPy array.
1-D array containing the values 1,2,3,4,5
2-D array containing two arrays with the values 1,2,3 & 4,5,6 5 marks

b) Explain Dictionary data type in Python. How to create, access, and modify dictionary elements? 5 marks

18. a) Explain features of Python programming. 5 marks

b) Write an output of the following R program

```
print("New vector using seq() function-")

v = seq(3, 15, by= 3)

print("Original vector:")

 print(v)

print("Check which values are even:\n")

is_even <- v %% 2 == 0

print(is_even)
```

5 marks

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 marks

20. a) Explain key features of Pandas library. 5 Marks

b) The joint probability distribution of two random variables X and Y is given by

y \ x	-2	4
1	0.1	0.1
-3	0.2	0.4
5	0.1	0.1

(i) Evaluate the marginal distributions of y.
(ii) Examine whether X & Y are independent.
(iii) Find P(Y=5/X=4)

21. a) What do you mean by data visualization? Explain data visualizing libraries in python. 5 Marks

b) Briefly explain different types of data(data measurements) with examples. 5 marks

22. A newly developed bike is tested for its mileage (Km/Ltr petrol). The followingdata reveal the number of times the bike is giving the mileage.

Mileage (Km/Ltr petrol).	40-45	45-50	50-55	55-60	60-65	65-70	70-75
No. of times	3	5	5	7	4	1	1

Obtain mean mileage and standard deviation of mileage of bike? 10 marks