

Indian Maritime University
(A Central University, Govt of India)
Supplementary Examinations – March / April 2024
Programme Name: B Tech (ME)
Semester: Three
Subject Code: UG11T4305

Subject Name: Statistics and Data Analysis using Python and R

Date: 13.04.2024

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.

Section A

Ten MCQs/Fill in the Blanks of 01 Mark each – Choose the correct answer as applicable.

1. What is the primary use of R in data analysis and statistics?
 - a. Creating data visualizations
 - b. Writing Python code
 - c. Data cleaning
 - d. Data analysis, statistical modelling, and visualization
2. Which keyword is used for function in Python language?
 - a. function
 - b. def
 - c. fun
 - d. define
3. To add a new element to a list we use which Python command?
 - a. list.add('element')
 - b. list.addend('element')
 - c. list.append('element')
 - d. list.addlast('element')
 - e.
4. Which of the following is a Python tuple?
 - a. {1,2,3}
 - b. (1,2,3)
 - c. [1,2,3]
 - d. None of the above

5. Find the mode of the call received on 7 consecutive day 11, 13, 13, 17, 19, 23, 25.
 - a. 11
 - b. 13
 - c. 17
 - d. 23

6. Standard deviation equal to
 - a. Square root of variance
 - b. Square of variance
 - c. Cube root of variance
 - d. Cube of variance

7. What is the return type of input()?
 - a. Integer (int)
 - b. String (str)
 - c. None
 - d. Undefined

8. What is the wrong method of variable assignment in R?
 - a. `x -> 5`
 - b. `x <- 5`
 - c. `x = 5`
 - d. None of the above

9. The most important object defined in NumPy is an N-dimensional array type called
 - a. ndarray
 - b. narray
 - c. nd_array
 - d. darray

10. What is the purpose of indentation in Python?
 - a. To define code blocks and maintain the structure of the program.
 - b. To add comments to the code.
 - c. To declare variables.
 - d. To print output to the console.

Section B

Five Questions of 02 Marks each

11. Python is a dynamically typed language, Explain briefly.
12. What are lists and tuples? What is the key difference between the two?
13. What are the advantages and disadvantages of mean, median and mode?
14. What is the main difference between an array and a matrix?
15. What is recycling of elements in a vector? Give an example.

Section C

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

16. (a) Consider the following code. Find if there is any syntax error and correct the syntax. Write down the output. (6)

```
print("What is my name?")
x = 2
for(i = 0; i < 5; i++):
    print(x+i, end='@')
else:
    x = "hello"
    print("\n", x, " Good Morning.")
```

- (b) Explain python break, continue and pass statements. (4)
17. (a) Explain docstring in python with an example. (3)
(b) Explain default value arguments. (3)
(c) Consider the following code, write the output (4)

```
def add_3(num):
    return num + 3

def add_2(num):
    return num + 2

x = 50

def change_number(x):
    if x % 2 == 0:
        return add_2(x)
    else:
        return add_3(x)

print(x)
print(change_number(x))
print(x)
```

18. (a) Explain different modes of opening a file with syntax. (4)
(b) Consider a document named 'document.txt'. Some text is written in the document and they are as follows, the numbers are representing the line numbers (don't consider the numbers as a part of document.txt).
- 1 Hello
 - 2 Good morning
 - 3 Indian

Write the output of the following code (6)

```
file = open('document.txt', 'rt')  
print(file.readline())  
print(file.readlines())  
print(file.read())  
file.close()
```

19. (a) Explain different types of data (data measurements) with examples. (4)
(b) Calculate median and mode marks of the following students. (6)

| Marks | 0-10 | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | 60-70 |
|----------------|------|-------|-------|-------|-------|-------|-------|
| No of students | 5 | 12 | 15 | 25 | 8 | 3 | 2 |

20. (a) Calculate the mean, variance and standard deviation for the following distribution : (7)

| class | 30-40 | 40-50 | 50-60 | 60-70 | 70-80 | 80-90 | 90-100 |
|-----------|-------|-------|-------|-------|-------|-------|--------|
| frequency | 3 | 7 | 12 | 15 | 8 | 3 | 2 |

- (b) Briefly explain skewness. (3)
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. 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)