

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT
TEST -2 EXAMINATIONS- 2023

M.Tech-I Semester (CSE)

COURSE CODE (CREDITS): 22M11CI112(3)

MAX. MARKS: 25

COURSE NAME: Introduction to Data Science

COURSE INSTRUCTORS: Dr. Anita

MAX. TIME: 1 Hour 30

Minutes

Note: (a) All questions are compulsory.

(b) Marks are indicated against each question in square brackets.

(c) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems

Q1a. Differentiate between Data Science, Machine Learning and AI. Python or R - Which one would you prefer for text analytics? (1.5) CO1

Q1b. Write a function in Python or R programming that takes in two sorted lists and outputs a sorted list that is their union. (1.5) CO2

Q1c What Does P-Value Signify About The Statistical Data? Explain with code. (2) CO3

Q2a Write the code in Python or R programming to calculate the Factorial of a number using Recursion. (1.5) CO2

Q2b What is the difference between skewed and uniform distribution? Write code also (2) CO1

Q2c Differentiate between func and func() in Python (1) CO2

Q2d In a time interval of 15-minutes, the probability that you may see a shooting star or a bunch of them is 0.3. What is the percentage chance of you seeing at least one star shooting from the sky if you are under it for about an hour? (1.5) CO5

Q2e Correction/ output of the following code: (4) CO2

(A) 350 = Number

```
WHILE Number<=1000:
```

```
    if Number=>750
```

```
        print (Number)
```

```
    Number=Number+100
```

```

else
print( Number*2)
Number=Number+50 Print( x-y)

```

(B) Val = int(rawinput("Value:"))

```

Adder = 0
for C in range(1,Val,3)
Adder+=C
    if C%2==0:
        Print (C*10)
    Else:
        Print (C*)
print (Adder)

```

(C) library("plotrix")

n <- c(178,160,136,227,123, 23,46,67,1,235)

print(std.error(a))

(D) ggplot(mydata100, aes(gender, fill=workshop)) + geom_bar(position="fill")

3a Fill in the blanks (5) CO2

- 1.R comes with a _____ to help you optimize your code and improve its performance.
- 2._____ suspends the execution of a function wherever it is called and puts the function in debug mode
- 3.The _____ function takes a vector or other objects and splits it into groups determined by a factor or list of factors.
- 4._____ involves predicting a response with meaningful magnitude, such as quantity sold, stock price, or return on investment.
5. How Statistical analysis is performed in R language?

3b How Geo map and time series can be plotted in Python. (5) CO4