

22U471

(Pages: 2)

Name:

Reg.No:

FOURTH SEMESTER B.Voc. DEGREE EXAMINATION, APRIL 2024

(CBCSS - UG)

(Regular/Supplementary/Improvement)

CC21U SDC4 DS13 - DATA SCIENCE WITH R PROGRAMMING

(Information Technology - Skill Component Course)

(2021 Admission onwards)

Time : 2.00 Hours

Maximum : 60 Marks

Credit : 3

Part A (Short answer questions)

Answer *all* questions. Each question carries 2 marks.

1. What is an array in R?
2. What are the functions used for reading data into R?
3. What is the difference between `dput()` and `dump()`?
4. Explain how do you manage dates in R.
5. Give examples of “select” and “filter” functions from “dplyr” package.
6. What is the difference between next and break statements in R.
7. What is the lazy evaluation of function?
8. Define `mapply()`.
9. How do you deal whitespaces of a string using `stringr` package?
10. What is Exploratory data analysis?
11. Explain KMean clustering.
12. Explain random forest.

(Ceiling: 20 Marks)

Part B (Short essay questions - Paragraph)

Answer *all* questions. Each question carries 5 marks.

13. What are the advantages and disadvantages of R?
14. Explain the different types of operators in R.
15. Explain about the file connections in R.
16. Explain how subsetting is applied in vector and matrix.

17. Explain vectorized operations. Briefly explain vectorized matrix operations.
18. What is Debugging? Name some functions which can be used for debugging in R.
19. Differentiate linear and logistic regression in R.

(Ceiling: 30 Marks)

Part C (Essay questions)

Answer any *one* question. The question carries 10 marks.

20. What are the different control structures in R?
21. What is data visualization? Explain ggplot2 package.

(1 × 10 = 10 Marks)
