

23U471

(Pages: 2)

Name :

Reg. No :

FOURTH SEMESTER B.Voc. DEGREE EXAMINATION, APRIL 2025

(CBCSS-UG)

(Regular/Supplementary/Improvement)

CC21U SDC4 DS13 - DATA SCIENCE WITH R PROGRAMMING

(Information Technology - Skill Component Course)

(2021 Admission onwards)

Time: 2 Hours

Maximum: 60 Marks

Credit: 3

Part A (Short answer questions)

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

1. What are the disadvantages of R Programming?
2. What are the functions which is used for merging of data frames horizontally and vertically in R?
3. What is the difference between dget() and source?
4. What are three operators that can be used to extract subsets of R objects?
5. Define mutate().
6. Explain the working of repeat loop in R.
7. What is ... Argument?
8. What is the difference between lapply and sapply?
9. Define object oriented programming in R.
10. What are the steps in exploratory data analysis?
11. What is classification?
12. Explain geoms in ggplot2.

(Ceiling: 20 Marks)

Part B (Short essay questions - Paragraph)

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

13. What are the different data objects in R?
14. Explain read.table() function.
15. Explain how do you create connections to a text file in R.
16. Explain vectorized operations. Briefly explain vectorized matrix operations.
17. Explain about dates and times in R.

18. What is the use of stringr package. Give some examples of the functions in Stringr.
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 clustering? What are the applications of clustering? Explain KMean clustering.

(1 × 10 = 10 Marks)
