

(FOR THE CANDIDATES ADMITTED

22UDA412

DURING THE ACADEMIC YEAR 2022

ONLY)

REG.NO.

N.G.M.COLLEGE (AUTONOMOUS) : POLLACHI
END-OF-SEMESTER EXAMINATIONS : MAY 2024

BSC.CS WITH DA(SF)

MAXIMUM MARKS: 50

SEMESTER IV

TIME : 3 HOURS

PART - III

22UDA412-R PROGRAMMING IN DATA SCIENCE

SECTION – A

(10 X 1 = 10 MARKS)

ANSWER THE FOLLOWING QUESTIONS.

(K1)

1. What is R programming primarily used for?
(a) Web development (b) Statistical computing and graphics (c) Game development
(d) Database management
2. Which package is often used for reading Excel files in R?
(a) xlsx (b) excelread (c) xlsimport (d) readxl
3. Which function is used to subset a data frame based on specified conditions in R?
(a) subset() (b) filter() (c) extract() (d) slice()
4. In R, what is the role of the if-else statement?
(a) Iterative execution (b) Vectorized operations (c) Decision-making based on conditions
(d) Exception handling
5. Which function is used to set breakpoints in R for debugging purposes?
(a) stop() (b) breakpoint() (c) browser() (d) halt()

ANSWER THE FOLLOWING IN ONE (OR) TWO SENTENCES

(K2)

6. What symbol is used for assignment in R?
7. In R, what does the readLines() function do?
8. Which function is used to select columns from a data frame using dplyr in R?
9. What is the purpose of the return keyword in a function in R?
10. What is the purpose of the rnorm() function in R?

SECTION – B

(5 X 3 = 15 MARKS)

ANSWER EITHER (a) OR (b) IN EACH OF THE FOLLOWING QUESTIONS. (K3)

11. a) Describe the primary features that distinguish R from other programming languages.
(OR)
b) Explain the process of creating a numeric vector in R.

(CONT...2)

12. a) Discuss the various ways of importing data into R.
(OR)
b) Explain the purpose of the dput() function in R.
13. a) Discuss the role of row and column indices in subsetting a matrix.
(OR)
b) Interpret the role of the group_by function in dplyr.
14. a) Discuss the role of the While loop in R.
(OR)
b) Explain the concept of lazy evaluation in R.
15. a) Explain the significance of calculating column and row sums or means in data analysis.
(OR)
b) Discuss the importance of Timing Longer Expressions in R.

SECTION – C**(5 X 5 = 25 MARKS)****ANSWER EITHER (a) OR (b) IN EACH OF THE FOLLOWING QUESTIONS. (K4 (Or) K5)**

16. a) Explain the limitations of R as an open-source programming language and its role in the field of data science.
(OR)
b) Explain the concept of a data frame in R.
17. a) Discuss the factors that influence the memory requirements for R objects.
(OR)
b) Examine the process of reading data from a URL connection in R.
18. a) Explain the concept of vectorized matrix operations in R.
(OR)
b) Elaborate on the concept of piping (%>%) in dplyr. How does it enhance the readability and conciseness of data manipulation workflows?
19. a) Examine the else and else if clauses in R's if statements. How are they used to define alternative code blocks based on different conditions?
(OR)
b) Explain the concepts of lexical scoping and dynamic scoping in R. What is the fundamental difference between the two?
20. a) Explain the concept of splitting a data frame in R.
(OR)
b) Examine the concept of simple random sampling. How is it implemented in R using the sample() function,