

BCA [Honours]

[CBCS]

B.Sc. Fifth Semester End Examination-2023

(Regular & Supplementary Paper)

PAPER-C12T

Data Science

Full Marks: 40

Time: 02 Hrs

The figures in the right hand margin indicate marks

Candidates are required to give their answers in their own words as far as practicable

Illustrate the answers wherever necessary

Group A

1) Answer any FIVE questions of the following: 5x2= 10

- i) Why is relational database not sufficient for managing big data?
- ii) How does matrix differ from data frame?
- iii) How do we read a CSV file in R?
- iv) How does lapply () differ from sapply() function?
- v) What do you mean by inner join of two dataset?
- vi) Explain any two different ways to turn your data into actionable knowledge.
- vii) Mention two debugging functions in R.
- viii) What is multivariate analysis?

(2)

Group B

Answer any FOUR questions of the following: 4x5 = 20

- 2) Explain how to retrieve specific rows of a dataset using dplyr package.
- 3) Explain the roles of gather () and separate () functions in tidyr package.
- 4) Write a short note on github for version control.
- 5) Explain the mechanism to join two datasets (i) vertically and (ii) horizontally.
- 6) Write down the difference between business intelligence and data science.
- 7) What is exploratory data analysis? Why is exploratory data analysis important in data science?

Group C

Answer any ONE question of the following: 1x10 = 10

- 8) i) Write a R program to read two compatible datasets containing X, Y values and plot the Y values together with appropriate legend. 5
ii) Write a R program that does the following : load dplyr package, read three compatible datasets, bind these three datasets vertically, and display last eight rows of the resultant dataset.

(3)

- 9) What do you mean by lexical and dynamic scoping of R? What are the different types of functions of R? Explain each with a proper example. 4+6
