

VR20

Reg. No:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 0 | 8 | N | 1 | A | 1 | 2 | 9 | 9 |
|---|---|---|---|---|---|---|---|---|---|

VELAGAPUDI RAMAKRISHNA

SIDDHARTHA ENGINEERING COLLEGE

(AUTONOMOUS)

II/IV B.Tech. DEGREE EXAMINATION, JULY, 2022

Fourth Semester

INFORMATION TECHNOLOGY

20BS4101E STATISTICS WITH R

Time: 3 hours

Max. Marks: 70

Part-A is compulsory

Answer One Question from each Unit of Part - B

Answer to any single question or its part shall be written at one place only

PART-A

10 x 1 = 10M

1.
 - a. List the data types in R.
 - b. Give a command to install a package in R.
 - c. What are joins in R?
 - d. List the most common functions for joins in R.
 - e. Write the formula of Binomial distribution.
 - f. Compare covariance and correlation.
 - g. What is the use of regression analysis?
 - h. Define simple linear regression.
 - i. Write if-else statement syntax in R.
 - j. Define decision tree.

20BS4101E

PART-B

4 x 15 = 60M

UNIT-I

2. a. Explain the looping statements in R with suitable examples. 7M
- b. Write R program to generate first n terms of a Fibonacci series. 8M

(or)

3. a. Discuss how data.frames and lists are created in R. 8M
- b. What is a vector? How to create it? Create a vector A of elements 3, 5, -7, 6, 8, 15, 35, 14, 16 and from it create a vector Y containing elements of A <=5. 7M

UNIT-II

4. a. What are the apply family functions available in R? Explain with examples. 8M
- b. Discuss about the following with an example 7M
 - i) cbind
 - ii) rbind

(or)

5. a. Illustrate the Math functions that calculates probability, cumulative sums and products with example. 7M
- b. Discuss different set operations in detail. 8M

UNIT-III

6. a. Assume the dataset 'faithful' given below.

| | Eruption | Waiting |
|---|----------|---------|
| 1 | 3.600 | 79 |
| 2 | 1.800 | 54 |
| 3 | 3.333 | 74 |
| 4 | 2.283 | 62 |
| 5 | 4.533 | 85 |
| 6 | 2.883 | 55 |

Find the covariance of eruption duration and waiting time from the above data set 'faithful'. Whether there is any linear relationship exists between the two variables. Justify. **7M**

- b. Explain in detail about Poisson distribution. **8M**

(or)

7. a. Explain the use of following with respect to t-test:

- i) Two-sample t-test and
- ii) Paired two-sample t-test. **7M**

- b. How to implement ANOVA in R? Is it possible to develop a test to compare two twitter topics to see which one is most popular using ANOVA? **8M**

UNIT-IV

8. a. Explain in detail about Poisson regression with R code examples. **7M**

- b. Explain about non linear least square model in detail. **8M**

20BS4101E

(or)

9. a. Explain autoregressive moving average model in R programming. **7M**
- b. Discuss hierarchical clustering in detail. **8M**

* * *