14CS601

Hall Ticket Number:

III/IV B.Tech (Regular/Supplementary-Repeat Exam) DEGREE EXAMINATION

January, 2021 **Common for CSE & IT**

Sixth Semester **Introduction to Data Analytics**

Time: Three Hours **Maximum:**60 Marks

*Answer **All Question** from **Part – A.*** (1X12 = 12 Marks)

*Answer **ANY FOUR** questions from **Part – B.*** (4X12=48 Marks)

Part – A

1. Answer all questions (1X12=12 Marks)
   a) What are the advantages of R language?
   b) What is a R session?
   c) List data types in R.
   d) Give syntax for abline() function.
   e) Which function is used to sort data in R?
   f) Remove missing value from the vector (2,4,NA,8).
   g) Define Normal distribution.
   h) What is the purpose of sprintf() function?
   i) What is regular expression?
   j) What is cluster analysis?
   k) Define Decision Tree.
   l) What is classification?

Part - B

2. a) Explain basic Math functions in R with examples. 6M
   b) Explain scan() and read.table() functions in R. 6M

3. a) Create a vector with first 100 numbers and write the appropriate R code: 6M
      i. Get the length of above vector
      ii. Get the first 10 numbers from above vector
      iii. Get the elements between 20 and 30 from above vector.
   b) Explain Lists with an example 6M

4. a) Write R program to display a bar plot for the following data representing various modes of travelling utilized by student in your college. 6M

Mode of Transportation	Bus	Bike	Bicycle	Others
Number of Vehicles	1000	500	100	200

   b) What is a box plot? Explain box plot with an example. 6M

5. a) Write a program in R that checks whether a given number is Prime or not. 6M
   b) For the given dataset below, write suitable R code 6M

id	name	branch	gender
1	VISHNU	CSE	M
2	DHANVI	CSE	F
3	PRAVEEN	IT	M
4	GEETHESWA	MCA	M
5	BHAVYA	MCA	F
6	VENKATESH	CSE	M
7	GAYATHRI	IT	F
8	RAJESH	MCA	M
9	DHARANI	CSE	F
10	PRANEETH	IT	M

      i. Read above data from the file data.csv
      ii. Display number of students exists in each branch.
      iii. Display number of male and female students exists in each branch.

P.T.O.

6. a) Fit a binomial distribution to the following data 6M

X	0	1	2	3	4	5	6
Y	5	18	28	12	7	6	4

b) Discuss Multiple Linear Regression. 6M

7. What is T-test? Explain different types of T-test with examples. 12M

8. a) What is a cluster? Explain steps involved in cluster analysis. 6M
b) Explain Hierarchial clustering with an example. 6M

9. a) What is classification? Explain Logistic regression. 6M
b) Explain about Decision Tree. 6M