FACULTY OF ENGINEERING

B. E. (CSE) VII – Semester (Supplementary) Examination, October 2021

Subject: Data Science Using R Program (E-II)

Time: 2 hours **Max. Marks: 70**

Missing data, if any may be suitably assumed

PART – A

Note: **Answers any five questions**. **(5x2=10 Marks)**

1. What do you mean by Data Science?
2. Define Eigen Vector.
3. Write the basic features of R.
4. Draw a box plot of the following observations
   28, 42, 25, 34, 37, 26, 33, 28, 36, 33, 22.
5. Write R code to return a complex object.
6. What is correlation analysis?
7. What is Regression?
8. Define classification.
9. Differentiate table and data frame in R.
10. Write the purpose of clustering.

PART - B

Note: Answers any four questions. **(4x15=60 Marks)**

11. How is linear algebra used in Data Science? Describe the objects that operate on Vectors and Matrices.

12. (a) Define hyperplanes. Demonstrate the usage of hyperplane in data science with an example.

    (b) Write R program to create Pie chart for the following data.

    Houseing-600, Food-300, Clothes-150, Entertainment-100, Others-200.

13. Describe the different types of Statistical Testing methods. Demonstrate T-test in R with an example.

14. (a) Write the steps in R to create a data frame containing name and income of father for 5 individuals using edit command.

…..2

(b) Write R-programming script to compute the sum of squares of N numbers. ($S_n=1^2+2^2+3^2+….n^2$)

15. Write the concept of normal distribution and explain an example with R code.

16.(a) Two teams, say the Cavs ant the Warriors, are playing a seven game championship series. The first to win four games, therefore, wins the series. The teams are equally good so that each have a 50-50 chance of winning each game. If the Cavs lose the first game, what is the probability that they win the series? Demonstrate with R..

(b) Describe an experiment of tossing a coin 80 times and prepare its frequency distribution.

17. Explain K-nearest neighbor technique and implementation of KNN in R.

******