



Maharishi University of Information Technology , Lucknow
Semester (Odd Semester) Examination 2020-2021

Max.Marks : 70**Course : B.Tech CDS****Subject / Paper's Name: R II (Statistical Modeling)****Paper code : BTDS503****Time : 2 hrs****Semester : 5****Regular/ Part time : R**

Note : Attempt all questions

Section A (Short Answer)**Q.1. Attempt any four of the following.****(4x10=40) Marks**

- a) Write a short note on Decision Tree. Explain using example.
- b) Write any three packages that are used for data visualization. Also write the functions from those packages.
- c) What is a QQ plot? What is its significance and why do we use it?
- d) Explain sqldf package in R. What is the importance of using data frame in R?
- e) Explain regression. Also explain the assumptions of Logistic regression.

Section B (Long Answer)**Q.2.- Attempt any Two of the following****(2x15 = 30) Marks**

- Q1. Create a data frame named df which contains name, age, occupation and income as variable. Also put 12 hypothetical observations in it.
- Q2. From mtcars(inbuilt) dataset in R plot a boxplot of any two continuous variable from the mtcars data. Also interpret the boxplot.
- Q3. Find the correlation between mpg and drat from the dataset used above.