

Hall Ticket No 

|  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|

Question Paper Code: BCSB06



# INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)

M.Tech I Semester End Examinations (Regular) - January, 2019

Regulation: IARE-R18

## FOUNDATIONS OF DATA SCIENCE

**Time: 3 Hours**

**(CSE)**

**Max Marks: 70**

---

**Answer ONE Question from each Unit**

**All Questions Carry Equal Marks**

**All parts of the question must be answered in one place only**

---

### UNIT – I

1. (a) What are the applications of R Programming in Real-World? Discuss in detail various stages in data science project. [7M]  
(b) List out inbuilt summary functions to apply on vectors. Create vector, matrix and array data object and apply inbuilt functions on it. [7M]
2. (a) State how array indexing and subsection of an array can be done in R? Write a R script to matrix multiplication. [7M]  
(b) Describe the probability distribution in R? Enumerate the steps for data cleaning and sampling. [7M]

### UNIT – II

3. (a) How to perform an ANOVA in R. Discuss the way to perform repeated measures with ancova in R with suitable example. [7M]  
(b) Discuss the multicollinearity. Assume a dataset and describe the procedure for finding hidden relations among attributes in the dataset. [7M]
4. (a) How to perform correlation analysis between multiple variables in R. Write a R script to get a linear equation  $y=mx+c$  for the heart weight and body weight in cats dataset. [7M]  
(b) Describe linear regression. What are the performance evaluation metrics in regression? How to implement regression in R? [7M]

### UNIT – III

5. (a) Discuss about data model. How to create and evaluate a data model. Describe with one case study. [7M]  
(b) List out different types of clustering. Write about K- means algorithm. Write a R script to cluster the mtcars dataset using KNN algorithm. [7M]
6. (a) What are the prerequisites for machine learning? Explain how is KNN different from k-means clustering? [7M]  
(b) Describe about the data model. Write any four learning techniques and in each case give the expression for weight – updating. [7M]

## UNIT – IV

7. (a) Discuss about ANN. Explain how do neural networks work? [7M]  
(b) Describe the limitations on the back propagation algorithm. Explain the scope to overcome these limitations [7M]
8. (a) Describe the null and alternative hypothesis with examples. What is p-value and give its importance. [7M]  
(b) List out the various learning algorithms. Explain gradient descent learning algorithm . [7M]

## UNIT – V

9. (a) Discuss about the residuals with respect to observed values? State a case study to show the fitted line and residuals in logistic regression. [7M]  
(b) Describe KNITR. State how to produce milestone documentation using KNITR. Explain simple markdown example. [7M]
10. (a) How to make a matrix plot. Explain the procedure to partition the window to get more number of plots. [7M]  
(b) List out the different plots with relevant packages to explore and summarize the multi-object plots in R. [7M]

