

Reg. No.

Question Paper Code

11635

B.E. / B.Tech. - DEGREE EXAMINATIONS, NOV/DEC 2022

Fifth Semester

Computer Science and Engineering

20CSEL504 - FOUNDATION OF DATA SCIENCE

(Regulations 2020)

Duration: 3 Hours

Max. Marks: 100

PART - A (10 × 2 = 20 Marks)

Answer ALL Questions

- | | <i>Marks,
K-Level, CO</i> |
|---|-------------------------------|
| 1. Define Data science Process. | 2,K1,CO1 |
| 2. Draw the lifecycle of a data science project. | 2,K1,CO1 |
| 3. Define confusion matrix. | 2,K1,CO2 |
| 4. What do you mean by Data Exploration? | 2,K1,CO2 |
| 5. Define Linear regression. | 2,K1,CO3 |
| 6. List the uses of hclust() function . | 2,K1,CO3 |
| 7. State the steps for Merging two lists. | 2,K1,CO4 |
| 8. Identify characteristics of a data frame. | 2,K2,CO4 |
| 9. Compare Distributed file system and traditional file system. | 2,K2,CO5 |
| 10. Identify the function of heartbeat message. | 2,K2,CO5 |

PART - B (5 × 13 = 65 Marks)

Answer ALL Questions

11. a) Explain the roles and stages in Data science Project. 13,K2,CO1
- OR**
- b) With suitable example explain the steps for Loading data into R from files. 13,K2,CO1
12. a) Discuss in detail about Exploring data. 13,K2,CO2
- OR**
- b) Explain the following R functions: class(), help(), summary(), dim(), ggplot(). 13,K2,CO2
13. a) Explain the K-means algorithm with a suitable example. 13,K2,CO3
- OR**
- b) Demonstrate the process of sampling for modeling and validation with an example. 13,K2,CO3

K1 – Remember; K2 – Understand; K3 – Apply; K4 – Analyze; K5 – Evaluate; K6 – Create

11635

14. a) Discuss in detail about Ordered and unordered factors. 13,K2,CO4

OR

b) Explain how Data Manipulation is performed in R with dplyr Package. 13,K2,CO4

15. a) Explain in detail about Map - Reduce Architecture. 13,K2,CO5

OR

b) Discuss Hadoop Distributed File System in detail. 13,K2,CO5

PART - C (1 × 15 = 15 Marks)

16. a) Discuss in detail about the types of Matrix Plot with an example. 15,K2,CO6

OR

b) Discuss in detail about creating effective presentation for end user. 15,K2,CO6