C

Question Paper Code: 57203

B.E. / B.Tech. DEGREE EXAMINATION, APRIL 2019

Seventh Semester

Computer Science and Engineering

15UCS703- DATA SCIENCE

(Regulation 2015)

Duration: Three hours   Maximum: 100 Marks

Answer ALL Questions

PART A - (5x 5 = 5 Marks)

1. In the R code statement 'i' come under which data type " i<- -1"   CO1-R

    (a) Character variable   (b) Logical variable

    (c) Numeric variable   (d) None of the above

2. Which of the following is one of the key data science skills?   CO2-R

    (a) Statistics   (b) Machine Learning

    (c) Data Visualization   (d) All of the Mentioned

3. Hadoop-based platform is to deal with   CO3-R

    (a) Structured data   (b) Unstructured data

    (c) Semistructured data   (d) Both (b) and (c)

4. Which of the following step is performed by data scientist after   CO4-R
    acquiring the data?

    (a) Data Cleansing   (b) Data Integration

    (c) Data Replication   (d) All of the Mentioned

5. _____ creates one map task for each split and a set of   CO5-R
    reducer tasks

    (a) JobTracker   (b) TaskTracker   (c) Both (a) and (b)   (d) None of the above

PART – B (5 x 2= 10Marks)

6. Give the four highlighted window panel in R-GUI.   CO1-R

7. List out the principles of data conditioning.   CO2-R

8. What are the characteristics that define the Big data?   CO3-R

9. State about the Confidence.   CO4-R

10. How the MapReduce framework is useful in application development?   CO5-R

PART – C (5 x 16= 80Marks)

11. (a) Describe about the R Graphical User Interfaces in detail.   CO1- U   (16)

Or

    (b) Explain about the Data Exploration versus Presentation.   CO1- U   (16)

12. (a) Briefly explain the k-means algorithm with an example.   CO2- U   (16)

Or

    (b) Illustrate the method to find k clusters from a collection of M objects with n attributes using k-means clustering.   CO2- Ana   (16)

13. (a) Describe about the Big Data characteristics in detail   CO3-Ana   (16)

Or

    (b) Why Hadoop is called a Big Data technology? Explain the components and Application development of Hadoop   CO3 -Ana   (16)

14. (a) Describe HDFS concepts and its operations of HDFS.   CO4 -U   (16)

Or

    (b) What is Data flow Model? Discuss different Kinds of Data flow Models   CO4- Ana   (16)

15. (a) Describe the Hadoop Map Reduce Application.   CO5- U   (16)

Or

    (b) Implement word count using Map Reduce.   CO5- U   (16)

57203