



Reg. No. :

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

**Question Paper Code : 70164**

M.E./M.Tech. DEGREE EXAMINATION, APRIL/MAY 2018

Second Semester

M.E. Computer Science and Engineering

CP 5293 – BIG DATA ANALYTICS

(Common to : M.E. Mobile and Pervasive Computing/M.E. Software Engineering/M.Tech. Information Technology)

(Regulations 2017)

Time : Three Hours

Maximum : 100 Marks

Answer ALL questions

PART – A

(10×2=20 Marks)

1. Which conditions data is called by “big data” ?
2. Give the difference between data analysis and data reporting.
3. What is Big Table ?
4. Write down the two major functionalities of Hadoop YARN.
5. What are association rules ? Give an example.
6. Define predictive analytics.
7. Give the properties of stream data.
8. Is opinion mining and sentiment analysis are same. Justify your answer.
9. Why NoSQL is essential in handling big data ?
10. What is Pig grunt ?

PART – B

(5×13=65 Marks)

11. a) Explain the characteristics features and the structure of big data in detail. (13)

(OR)

- b) What is analytic sandbox ? What are the benefits of an analytic sandbox ?  
Explain its types in detail. (13)

70164



12. a) What is HDFS ? Illustrate the major blocks in HDFS architecture. (13)

(OR)

b) Explain how matrix multiplication is carried out using mapreduce algorithm. (13)

13. a) Explain multivariate regression analysis. Explain the pros and cons of statistical data analysis. (13)

(OR)

b) Explain the basic working methodology of density based clustering mechanism. (13)

14. a) What is a data stream model ? Explain data stream management system and its components in detail. (13)

(OR)

b) i) Explain the common procedure followed for extracting reliable sample from a stream. (7)

ii) What is RTAP ? Why it is important in timely decision making ? (6)

15. a) Explain Hbase architecture and its data model. (13)

(OR)

b) Write HiveQL queries to perform the following task with example.

● Create table

● Divide table into related parts (partitions)

● Joining tables. (13)

PART – C

(1×15=15 Marks)

16. a) Describe the technique which will you adapt to find the opinion of the buyer of on an electronic gadget. The output of the analysis should reveal whether the user has a positive or negative or neutral sentiment about that product. (15)

(OR)

b) Explain the procedure for installing R in your system. Describe the any three packages which support predictive analysis. (15)