

Reg. No.																				
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Question Paper Code	12006
---------------------	-------

14 JUL 2023

**M.E. / M.Tech. - DEGREE EXAMINATIONS, APRIL/MAY 2023**

Second Semester

**M.E. - Computer Science and Engineering  
20PCSPC204 - BIG DATA ANALYTICS**

(Regulations 2020)

Duration: 3 Hours

Max. Marks: 100

**PART - A (10 × 2 = 20 Marks)**

Answer ALL Questions

- |   | <i>Marks,<br/>K-Level, CO</i> |
|---|-------------------------------|
| 1. Why big data is important?   | 2,K2,CO1                      |
| 2. How is traditional BI environment different from the big data environment?                                     | 2,K2,CO1                      |
| 3. State few challenges with big data.  | 2,K1,CO2                      |
| 4. Define Brewer's Theorem.   | 2,K1,CO2                      |
| 5. List Hadoop's configuration files.   | 2,K1,CO3                      |
| 6. What do you mean by Map Reduce?  | 2,K1,CO3                      |
| 7. Let (x1, x2, .....xn) are set elements in a cluster. Write equation to find Centroid and Radius of the cluster | 2,K2,CO4                      |
| 8. List out the various R objects.  | 2,K1,CO4                      |
| 9. Enumerate few features of MongoDB.   | 2,K2,CO5                      |
| 10. Write equivalent MongoDB query for Select * from employees order by salary.                                   | 2,K2,CO5                      |

**PART - B (5 × 13 = 65 Marks)**

Answer ALL Questions

- |   |           |
|---|-----------|
| 11. a) Explain the classification of data.                        | 13,K2,CO1 |
| <b>OR</b>   |           |
| b) Describe a typical Data warehouse and Hadoop Environment.      | 13,K2,CO1 |
| 12. a) Explain in detail about the classification of Analytics.   | 13,K2,CO2 |
| <b>OR</b>   |           |
| b) (i) Discuss the responsibilities of Data Scientist.            | 7,K2,CO2  |
| (ii) Write a short note on Soft state eventual consistency.       | 6,K2,CO2  |
| 13. a) Describe the MapReduce programming architecture in detail. | 13,K2,CO3 |

**OR**

- b) Discuss how the resources and applications are managed using Yarn architecture. *13,K2,CO3*
14. a) Explain in detail about the K-means clustering algorithm with an example. *13,K2,CO4*

**OR**

- b) Outline the basic working methodology of density based clustering mechanism. *13,K2,CO4*
15. a) How is a cursor implemented in MongoDB? Explain with a suitable example. *13,K2,CO5*

**OR**

- b) Interpret how Jasper Report is prepared using Jasper soft. *13,K2,CO6*

**PART - C (1 × 15 = 15 Marks)**

16. a) Sketch Apache Hbase architecture and explain the architecture in detail. *15,K2,CO6*

**OR**

- b) Discuss the various Interaction Techniques in detail. *15,K2,CO6*