



23105817

QP CODE: 23105817

Reg No : .....

Name : .....

**B.Sc DEGREE (CBCS) REGULAR / REAPPEARANCE EXAMINATIONS,  
MARCH 2023  
Sixth Semester**

B.Sc Computer Science Model III

**CORE COURSE - CC6CRT07 - BIG DATA : ANALYTICS**

2017 Admission Onwards

OC2C80FA

Time: 3 Hours

Max. Marks : 80

**Part A**

*Answer any **ten** questions.*

*Each question carries **2** marks.*

1. Define the term 'Big Data'
2. What do you mean by intelligent data analysis?
3. Define mean, median and mode.
4. What is surprise number?
5. Define decaying windows.
6. Give examples for real time analytics.
7. Define MapReduce.
8. Explain the purpose of RecordReader.
9. List any four Hadoop configuration files.
10. Review the importance of java management extensions.
11. Which are the execution environments needed for running a Pig latin program?
12. Define Hbase.

(10×2=20)

**Part B**

*Answer any **six** questions.*

*Each question carries **5** marks.*





13. Define the important v's in big data.
14. How can we report the number of unique user in a website over the past month? Write the query for it.
15. With a neat diagram explain different Phases in MapReduce?
16. Explain the Architecture of HDFS.
17. Briefly explain the use of counters available in MapReduce ? List the different type of counters.
18. Discuss the setting up of a Hadoop cluster.
19. What is a context? Illustrate the use of various contexts to collect metrics in Hadoop.
20. Write a note on stream processing language.
21. Describe the application of interactive data visualization.

(6×5=30)

### Part C

*Answer any **two** questions.*

*Each question carries **15** marks.*

22. Explain Flajolet Martin Algorithm.
23. Define MapReduce, Job and Task? Explain the working of a Classic MapReduce?
24. Describe the steps involved in setting up a Hadoop cluster.
25. Explain IBM infosphere biginsights in detail.

(2×15=30)

