



QP CODE: 25021049



25021049

Reg No :

Name :

**B.Sc DEGREE (CBCS) REGULAR / REAPPEARANCE / MERCY CHANCE
EXAMINATIONS, FEBRUARY 2025**

Sixth Semester

B.Sc Computer Science Model III

CORE COURSE - CC6CRT07 - BIG DATA : ANALYTICS

2017 Admission Onwards

4EFABFDA

Time: 3 Hours

Max. Marks : 80

Part A

*Answer any **ten** questions.*

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

1. What do you mean by intelligent data analysis?
2. Define mean, median and mode.
3. What do you mean by prediction error?
4. List the applications for filtering streams.
5. State the relevance of estimating moments.
6. How can we find the most popular element in a stream?
7. What is the need of Shuffle Phase?
8. Write about Sequence File Input Format.
9. How do Hadoop control scripts perform cluster wide operations?
10. Review the importance of metadata backups.
11. Define stream.
12. What is visual data analysis?

(10×2=20)

Part B

*Answer any **six** questions.*

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

13. Write a note on 'nature of big data'.





14. Explain the technique for estimating higher moments.
15. Explain the difference between InputSplit and DataBlock. What should be the appropriate size of InpuSplit and why?
16. Write about the Hadoop Java API.
17. How the Classic MapReduce Handle the Task Failure?
18. Restate in own words, how security is enforced in Hadoop through Kerberos.
19. Recall the steps involved in commisioning and decommissioning nodes in a Hadoop cluster.
20. Discuss the the advantages of Pig.
21. Describe the application of interactive data visualization.

(6×5=30)

Part C

*Answer any **two** questions.*

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

22. Explain DGIM algorithm.Explain query answering in DGIM algorithm.
23. Define MapReduce, Job and Task? Explain the working of a Classic MapReduce.
24. Describe the steps involved in setting up a Haddop cluster.
25. Explain various Pig latin data processsing operators with suitable examples.

(2×15=30)

