



QP CODE: 26019134



26019134

Reg No :

Name :

M.Sc DEGREE (CSS) EXAMINATION, MARCH 2026

Fourth Semester

M Sc COMPUTER SCIENCE (AIDED)

ELECTIVE - CA810402 - BIG DATA MANAGEMENT USING R

2019 Admission Onwards

9A02A593

Time: 3 Hours

Weightage: 30

Part A (Short Answer Questions)

*Answer any **eight** questions.*

Weight 1 each.

1. State the importance of the group "Data Savvy People" in new Big data ecosystem.
2. What is the role to be played by the Database Administrator in a successful Analytic project?
3. What is the importance of Resources in the Discovery phase of Data Analytic Lifecycle?
4. Name few Free or open tools used for the Model Planning phase of Data Analytic Lifecycle.
5. What is parsing?
6. What are Morphological features for Representing a Text?
7. Describe the output expected by the Data engineer / DBA, BI Analyst and Data Scientists at the end of the analytic project.
8. Name few open tools for Data Visualization.
9. Write R program to calculate mean, median and mode with these data. 63, 81, 56, 91, 47, 57, 76, 72, 62, 48.
10. Describe the uses of ggplot.

(8×1=8 weightage)

Part B (Short Essay/Problems)

*Answer any **six** questions.*

Weight 2 each.

11. Describe a short overview about Big data.
12. Describe the role of Data collectors in emerging Big data ecosystem.





13. Describe the relevance of Learning about Data in the Data Preparation phase of Data Analytic Lifecycle.
14. Briefly explain the relevance of Operationalize phase in Data Analytic Lifecycle.
15. Describe how to develop and execute a Mapreduce program.
16. Describe the usage considerations of Apache Hbase.
17. Describe the significance of providing Technical specifications and code in the final presentation of an analytical project.
18. Explain about different types of objects used in R for creating data.

(6×2=12 weightage)

Part C (Essay Type Questions)

*Answer any **two** questions.*

Weight 5 each.

19. Explain the differences between Business Intelligence and Data Science.
20. Explain the Model planning phase in Data Analytics Lifecycle.
21. Explain how to calculate Term Frequency-Inverse Document Frequency (TF-IDF) with an example.
22. Explain the role of Graphs in Final presentation and how to clean up the graphs.

(2×5=10 weightage)

