



Registration Number:

Date & session:

ST. JOSEPH'S UNIVERSITY, BENGALURU -27
M.Sc. (BIG DATA ANALYTICS) – III SEMESTER
SEMESTER EXAMINATION: OCTOBER 2023
(Examination Conducted in November/December 2023)
BDA 3221 – ENABLING TECHNOLOGIES FOR DATA SCIENCE II
(For current batch students only)

Time: 2 Hours

This paper contains ONE printed page and THREE parts

Max Marks: 50

PART A

Answer All questions

2 X 5 = 10

- 1 What is a data lake? 2
- 2 Why use a data warehouse? 2
- 3 Show an example of a fact table. 2
- 4 Can data mining be a security risk? Explain. 2
- 5 Give any two applications of Scala. 2

PART B

Answer ANY FIVE questions

5 X 4 = 20

- 6 Explain slowly changing dimension and Junk dimension. 4
- 7 Explain snowflake schema with example. 4
- 8 Explain a bitmap index with example. 4
- 9 What is ROLAP? 4
- 10 What are some of the techniques used in Data Mining? 4
- 11 Why is Scala said to be safe? 4
- 12 What are the advantages of Spark? 4

PART C

Answer ANY TWO questions

2 X 10 =20

- 13(a) Explain Spark's architecture 5
- 13(b) Explain DAG in Spark 5
- 14(a) Explain the operations of a data cube 6
- 14(b) What are the uses of data cube? 4
- 15 Explain CRISP-DM 10