



Register No:

Date: 19-11-2020

ST. JOSEPH'S COLLEGE (AUTONOMOUS), BANGALORE

III SEMESTER EXAMINATION, NOVEMBER 2020

M.SC IN BIG DATA ANALYTICS

BDA 3218: ENABLING TECHNOLOGIES FOR DATA SCIENCE

TIME 2.5 HRS

MAX MARKS 70

This Question Paper Contains Two Parts

PART I

Answer Any Six Questions

(6*5=30)

- 1) How a secondary name node differs from the name node in HDFS?
- 2) Define the three key design principles of pig Latin.
- 3) How to create a table by using HIVEQL.
- 4) How to improve HIVE query performance?
- 5) Write all steps of dynamic partition table creation.
- 6) Explain about the implementation of map reduce concept with a small example?
- 7) Write syntax for incremental updates using Sqoop.
- 8) Explain the architecture of a pig with a neat sketch?
- 9) Drawback with 1.X and how to overcome

PART II

Answer Any Four Questions

(4*10=40)

- 1) Explain the role of driver code, mapper code and reducer code within a map reduce program model by a suitable example.
- 2) What are the components of YARN? Explain its component functionality with neat diagram.
- 3) Explain about the processing of a job in Hadoop with a neat sketch.
- 4) Explain the operators supported by pig with respect to data access, transformations and debugging operations.
- 5) With a neat diagram explain the complete architecture of HIVE and explain each component in detail.

BDA3218-A-20