



Registration Number:

Date & session:

**ST JOSEPH'S UNIVERSITY, BENGALURU -27**  
**M.SC (BIG DATA ANALYTICS) – II SEMESTER**  
**SEMESTER EXAMINATION: APRIL 2024**  
**(Examination conducted in May / June 2024)**

**BDA2421 – ENABLING TECHNOLOGIES FOR DATA SCIENCE I**  
**(For current batch students only)**

**Time: 2 Hours**

**Max Marks: 50**

**This paper contains ONE printed page and THREE parts**

**PART - A**

**Answer ALL Questions**

**5 X 2 = 10**

1. What are the main components of Hadoop ecosystem?
2. Explain heartbeat
3. Explain Apache Ambari.
4. When is Pig preferred over Hive?
5. Write about Hive schema.

**PART – B**

**Answer ANY FIVE Questions**

**5 X 4 = 20**

6. What was the need for Hadoop to be developed?
7. How is a secondary name node different from a primary name node?
8. How is security prioritised in Hadoop?
9. Describe balancer.
10. Differentiate between the types of Hive tables.
11. Explain Sqoop.
12. When is using Hadoop a bad idea?

**PART - C**

**Answer ANY TWO Questions**

**2 X 10 = 20**

13. Explain the idea of data and metadata persistence in Hadoop 2x systems.
14. Explain Erasure coding in Hadoop 3x.
15. Explain map reduce with an example.