

S Name: \_\_\_\_\_

Reg. No.: \_\_\_\_\_

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|

**END TERM EXAMINATION – January-2023**

**SEMESTER – VII**  
**(B.Tech. CSE DS/A1)**

**Subject Code: CDS4001**

**Subject: Data Science**

**Duration: 3 hours**

**Max. Marks: 100**

**Instructions**

- All Questions are compulsory
- The Question paper consists of 2 sections - Part A contains 10 questions of 2 marks each. Part B consists of 5 questions of 16 marks each.
- There is no overall choice. Only Part B question include internal choice.

**PART – A**

**(2 \* 10 = 20 Marks)**

1. What is MapReduce in Hadoop?
2. What are the 5V's of Big Data?
3. What are the Features of Hive?
4. What is Brute Statistics?
5. Explain Data Node.
6. Explain Data Science.
7. Explain HDFS and It's Characteristics.
8. What is Veracity?
9. Explain the Types of Data.
10. What are the Model Requirements while transforming data?

**PART – B**

**(16 \* 5 = 80 Marks)**

- ✓ 11.a) Elaborate the term 'Big Data'. Paraphrase it's important characteristics in detail. (4+12)

**OR**

- b) Elaborate the term SQOOP? Quote it's functionality an application in the field of Big Data. (4+12)

- ✓ 12.a) Paraphrase the working of Map Reduce in Hadoop? Categorize it and explain the functionality of each phase and Quote the example of Word Count. (4+12)

**OR**

- b) Introduce HDFS? Justify Master Slave architecture of HDFS with clear representation. (4+12)

- 13.a) Introduce Software Development Life Cycle. Quote the phases of it. Paraphrase about each phase with a clear representation with a live scenario. (2+4+10)

**OR**

- ✓ b) Explain in detail the recent project developed by you in Data Science with clear abstract, chronology and prototype. (16)

- ✓ 14. a) Introduce Data science ? What is it's association with Machine Learning and Artificial Intelligence? Explain clearly the importance of each term in the field of Computer Science. (2+4+10)

**OR**

- b) What is Data visualization? And What are the common graphs used in Data visualization? (4+12)

**15. a) Elaborate Big SQL? Elaborate its characteristics in detail.**

**(4+12)**

**OR**

**b) Quote and Elaborate the term 'Hadoop'. Justify the importance of Hadoop framework in the field of DS.**

**(4+12)**