

Roll No: -----



UNIVERSITY OF PETROLEUM AND ENERGY STUDIES

End Semester Examination, December 2017

Program: BTech CSE + Big Data

Semester – I

Subject (Course): Big Data Overview

Max. Marks : 100

Course Code : CSBD1001

Duration : 3 Hrs

No. of page/s: 2

*This paper consists of 3 sections as **Section A, Section B** and **Section C** having **20, 40 and 40***

Section A (Marks: 20)

Attempt ALL questions. The marks for each question are written alongside with the question.

1. What are the Key Attributes of a Data Lake? [4M]
2. What is Data Deluge Problem? [4M]
3. What is Spark and draw architectural diagram of Apache Spark. [4M]
4. What is NoSQL database and what are its different types. [4M]
5. What are the master-slave architecture components in HDFS. [4M]

Section B (Marks: 40)

The marks for each question are written alongside with the question.

6. What are the different data storage mechanisms in Big Data System. [10M]
7. What is Big Data ecosystem and explain its characteristics? What are the practical use cases of Big Data eco-system. [10M]
8. Explain [2.5 x 4 = 10M]
 - a. Avro
 - b. Parquet
 - c. JSON
 - d. Sequence Files

9. Explain how Big Data is used in various application

[2.5 x 4 = 10M]

- a. Sports
- b. Recommendation System
- c. Travel and Tourism
- d. Security

Section C (Marks: 40)

The marks for each question are written alongside with the question.

10. What are the benefits of Big Data? Discuss challenges under Big Data. How Big Data Analytics can be useful in the development of smart cities. (Discuss one application)

[6+6+8 Marks]

11. Define Data Lake? How is Data in Data Lakes Organized and Managed? What's the Difference Between a Data Lake and Data Warehouse? Where Can You Use a Data Lake?

[2+6+8+4 Marks]

