

Total No. of Questions : 8]

SEAT No. :

**PB3866**

[Total No. of Pages : 2

**[6262]-130**

**T.E. (Information Technology)**

**DATA SCIENCE AND BIG DATA ANALYTICS**

**(2019 Pattern) (Semester - II) (314452)**

*Time : 2½ Hours]*

*[Max. Marks : 70*

*Instructions to the candidates:*

- 1) *All questions are compulsory.*
- 2) *Figures to the right indicate full marks.*

**Q1) a)** Explain all the steps for writing a file in HDFS with neat diagram. **[8]**

b) Describe the various types of NoSQL Databases with example and also compare them. **[10]**

OR

**Q2) a)** Why map reduce is required in Hadoop? Explain the stages involved in map reduce task with a suitable example? **[9]**

b) What is Hadoop Distributed system? What is the advantage of heart bit message in Hadoop. **[9]**

**Q3) a)** Compare HBASE and HIVE with suitable parameters. **[8]**

b) How missing values are filled in Pandas Data Frame with zeros? Assume suitable data. **[3]**

c) Explain Min-max scaling. For the following dataset carry out min-max Scaling, X = 24,28,53,30,40,18,15,21. **[6]**

OR

**Q4) a)** What is categorical variable? Why do you need categorical variable encoding? With an example, explain one-hot encoding? **[8]**

b) What is data wrangling? Why do you need it? Explain data wrangling methods? **[9]**

**P.T.O.**

- Q5)** a) How Data Visualization is important in Big Data? Explain challenges to big data visualization? [6]
- b) Explain various techniques for visual data representation. [6]
- c) Explain the following data visualization techniques. [6]
- i) Google Chart API
- ii) D3.js

OR

- Q6)** a) Explain data visualization with respect to 1-D, 2-D, 3-D data? [6]
- b) Explain various analytical techniques and tools used in data visualization. [6]
- c) Draw boxplot with a suitable example and explain its usage. [6]

- Q7)** a) What is text mining? Draw and explain text mining architecture and its use. [8]
- b) Explain primary activities of Michael Porters value chain. [5]
- c) How mobile analytic is different from social media analytic? [4]

OR

- Q8)** a) What is Porters valuation creation model? Explain porter's value chain analysis. [9]
- b) What is social media analytic? Explain the process of social media data analytic. [8]

x x x