

Total No. of Questions : 8]

**PD4330**

SEAT No. :

[Total No. of Pages : 2

[6403] 128

**T.E. (Information Technology)**  
**DATA SCIENCE AND BIG DATA ANALYTICS**  
**(2019 Pattern) (Semester - VI) (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 GFS as solution for Big data Processing Challenges. [6]

b) Write short note on HIVE and HBASE. [8]

c) Explain any 3 Hadoop shell commands. [3]

OR

**Q2)** a) Explain HDFS with a neat diagram. [6]

b) Write the 4 types of NoSQL databases. [8]

c) State and explain CAP Theorem. [3]

**Q3)** a) Explain the 4 types of big data analytics along with the analytical purpose. [8]

b) Explain the different modes of data transformation in big data. [9]

OR

**Q4)** a) Explain the tasks in Data integration and Data Transformation phase. [8]

b) Explain the three ways of Data Normalization used in data analytics. [9]

*P.T.O.*

- Q5)** a) Explain different Data Visualization Techniques. [8]  
b) What are five purposes of using charts and state which type of charts used for each purpose? [10]

OR

- Q6)** a) Explain how data Visualization is beneficial for Business Analytics and Intelligence. [8]  
b) Write note on any three of the following DV tools. [10]  
i) Tableau  
ii) Infogram  
iii) Fusion Charts

- Q7)** a) Explain what is Social Media Data Analytics and state the advantages of SMA. [9]  
b) Explain the data analysis life cycle in big data with example. [9]

OR

- Q8)** a) What is Mobile analytics? Explain its importance with suitable example. [9]  
b) Explain Text mining architecture with suitable example. [9]

