

Total No. of Questions : 8]

SEAT No. :

P826

[5870] - 1149

[Total No. of Pages : 2

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) Answer Q.1, or Q.2, Q.3 or Q.4, Q.5 or Q.6, Q.7 or Q.8.
- 2) Neat diagrams must be drawn wherever necessary.
- 3) Figures to the right side indicate full marks.
- 4) Assume the suitable data, if necessary.

Q1) a) Explain Big data Ecosystem with suitable diagram. [7]

b) Explain anatomy of File read and write in HDFS. [7]

c) Write and explain any two Hadoop shell commands. [4]

OR

Q2) a) Explain Map Reduce with proper diagram for word count example. [7]

b) Explain Google file system. [7]

c) Explain ETL processing. [4]

Q3) a) Explain different steps in Data Analytics Project Life cycle [7]

b) Draw and explain Architecture of HIVE. [7]

c) Explain different data transformation techniques. [3]

OR

Q4) a) Explain different kinds of Big Data Analysis. [7]

b) How data can be ingested in python. Write syntax in python for the same. [7]

c) Explain role of visualization in big data analytics. [3]

Q5) a) Explain different techniques of Big Data visualization. [7]

b) Explain challenges in Big data visualization. [7]

c) Write two data visualization functions from matplotlib. [3]

OR

P.T.O.

- Q6)** a) Explain different tools for data visualization. [7]
b) Explain scatter plot, histogram and heat map with example. [7]
c) Write two data visualization functions from seaborn. [3]

- Q7)** a) How Social Media analytics helps in value creation? Explain with suitable examples. [7]
b) Explain in brief data analytics life cycle. [7]
c) Explain big data value terminology. [4]

OR

- Q8)** a) What is text mining? Draw and explain text mining architecture and its use. [7]
b) Explain Big data analytics in research. [7]
c) Explain big data impact on organizations. [4]