Total No. of Questions : 12]

SEAT No. :

PA-674

[Total No. of Pages : 2

[5928]-120

M.E. (Computer) (Data Science)

BIG DATA ANALYTICS

(2017 Pattern) (Semester - I) (510303)

Time : 3 Hours] *[Max. Marks : 50*

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, Q.9 or Q.10, Q.11 or Q.12.***
2) ***Neat diagrams must be drawn wherever necessary.***
3) ***Figures to the right indicate full marks.***
4) ***Assume suitable data, if necessary.***
5) ***Use of logarithmic tables slide rule. Mollier charts, electronics pocket calculator and stream tables is allowed.***

Q1) Define Big Data Analytics? What are different Tools used for Big Data Analytics? Explain with suitable example. **[9]**

OR

Q2) List the key roles for a successful analytics project. Explain the following roles in details with example. **[9]**

a) Data Engineer

b) Data Scientist

Q3) Define Distributed Computing. Explain the working of a distributed computing environment with appropriate diagram. **[9]**

OR

Q4) Explain the working of APache Hadoop with suitable components and architecture diagram. **[9]**

P.T.O.

Q5) Compare Pig and Hive. Explain different execution modes available in Pig. Explain in brief the applications of PIG. **[8]**

OR

Q6) Describe Apache Zookeeper. Explain the working of Apache Zookeeper in brief and List the benefits of the same in Big Data problems. **[8]**

Q7) Explain how Spark runs applications with the help of its architecture and suitable example. **[8]**

OR

Q8) What is Apache Spark RDD? Why do we need RDD in Spark? Explain different Features of Spark RDD in detail. **[8]**

Q9) Describe role of PySpark in Big Data Analytics. Explain the various PySpark persistence levels in details. **[8]**

OR

Q10) What different cluster managers does Apache Spark offer? Briefly explain each. **[8]**

Q11) List the major challenges of Data Visualization. Visualization is a fantastic tool for understanding, analyzing, and sharing information. Justify the statement. **[8]**

OR

Q12) Describe Big Data Visualization? What is the Need for advanced visualization techniques in big data? **[8]**