

Seat No.	
-------------	--

**B.E. (Computer Science & Engg.) (Semester - VIII) (Revised)****Examination, November - 2017****DATA ANALYTICS****Sub. Code : 67824****Day and Date : Wednesday, 01 - 11 - 2017****Total Marks : 100****Time : 10.00 a.m. to 1.00 p.m.**

- Instructions :**
- 1) Figures to the right indicate full marks.
  - 2) Q.4 & Q.8 are compulsory.
  - 3) Attempt any two questions from Q.1, to Q.3 and from Q.5 to Q.7.

**Q1) a)** Explain the phases in the developments of Business Intelligence System with the help of neat diagram? **[8]**

b) Explain Star Schema, Snowflake Schema, Galaxy Schema with proper example? **[8]**

**Q2) a)** Explain different categories of mathematical models for decision making? **[8]**

b) Explain in detail process of univariate analysis? **[8]**

**Q3) a)** Explain HDFS Architecture and the working of Hadoop Heartbeat message in HDFS with proper figure? **[8]**

b) Explain data validation process in data preparation. **[8]**

**Q4) Write a note on (Attempt Any Three) :** **[18]**

a) Data Warehouse Architecture.

b) Hive, HBase, Pig & Pig Latin.

c) Multivariate Analysis.

d) Data mining tasks.

- Q5) a)** Explain simple linear regression model and structure of regression model. [8]  
b) Discuss the structure and phases of the learning process for a classification with a neat diagram. [8]
- Q6) a)** Explain in detail any one algorithm used for partition methods. [8]  
b) Explain the general association rules that is useful for range of applications. [8]
- Q7) a)** List and explain the different functions to handle the data in R workspace with an example. [8]  
b) List and explain the various types of R commands to import data. [8]
- Q8) Write a short note on (Any Three) :** [18]  
a) Bayesian methods.  
b) Exporting data from R.  
c) Hierarchical clustering methods.  
d) Apriori algorithm.

