

(2½ Hours)

[Total Marks: 60]

- N. B.: (1) All questions are compulsory.  
(2) Make suitable assumptions wherever necessary and state the assumptions made.  
(3) Answers to the same question must be written together.  
(4) Numbers to the right indicate marks.  
(5) Draw neat labeled diagrams wherever necessary.  
(6) Use of Non-programmable calculators is allowed.

1. Attempt any two of the following: 12
  - a. List and explain the characteristics of Big Data.
  - b. Explain the challenges with Big Data.
  - c. Explain the responsibilities of a Data Scientist.
  - d. What are the main phases of the Data Analytics? Explain using suitable diagram.
2. Attempt any two of the following: 12
  - a. Write a short note on K-Mean.
  - b. Explain Association Rules for unsupervised learning.
  - c. Write a short note on Regression Analysis using suitable example.
  - d. Explain any three applications of logistic regression model applied to situations in Government and the Private sector.
3. Attempt any two of the following: 12
  - a. Write a short note on Decision Tree.
  - b. Write a short note on Naïve Bayes Theorem.
  - c. List and explain the goals of time series analysis.
  - d. Explain the steps involved in text analysis.
4. Attempt any two of the following: 12
  - a. Explain the characteristics of Big Data pipeline.
  - b. Write a note on Hadoop Big Data Operating System.
  - c. Explain the Hadoop Architecture using suitable diagram.
  - d. Write a short note on HDFS.
5. Attempt any two of the following: 12
  - a. Explain the concept of Computing with Keys.
  - b. Write a short note on Design Patterns.
  - c. Explain the concept of structured queries with HIVE.
  - d. Write a short note on DATA INGESTION.