

Roll No.

--	--	--	--	--	--	--

H

TCS-703/TIT-703

B. Tech. (CS/IT) (Seventh Semester) Mid Semester EXAMINATION, 2017

BIG DATA ANALYSIS

Time : 1:30 Hours]

[Maximum Marks : 50

Note : (i) This question paper contains two Sections.

(ii) Both Sections are compulsory.

Section—A

1. Fill in the blanks/True-False : (1×5=5 Marks)
 - (a) Big Data deals with huge data sets which are complex and huge. (True/False)
 - (b) Electronic data interchange is necessary in data analysis. (True/False)
 - (c) often doesn't ask why and simply detects patterns.
 - (d) uses descriptive statistics with data with high information density to measure things, detect trends
 - (e) Sampling (statistics) enables the selection of right data points from within the larger data set to estimate the characteristics to the whole population. (True/False)

[2] TCS-703/TIT-703

2. Attempt any *five* parts : (3×5=15 Marks)

- (a) Define Big Data and its utility in decision making.
- (b) Relevance of Statistical Tools in Data Management.
- (c) Define Data Exploration.
- (d) Concept of Clustering and Association.
- (e) Need of Data Analysis.
- (f) Types of data structures used in Big Data.

Section—B

3. Attempt any *two* parts of choice from (a), (b) and (c). (5×2=10 Marks)

- (a) Explain the utility of big data management of business with suitable industry examples.
- (b) Identify a business problem and business drivers using Big Data as frame of reference.
- (c) "Communication of results is essential in an organization with reference to data analysis." Explain and also how it can be managed with reference to Data Management.

4. Attempt any *two* parts of choice from (a), (b) and (c). (5×2=10 Marks)

- (a) What is the relevance of Data Life Cycle Analytics ?

B-68

[3]

- (b) Differentiate between Support and Confidence Level. Justify which gives better results with relevant examples.
 - (c) Describe the benefit of statistical results and also the process of calculating statistical values.
5. Attempt any *two* parts of choice from (a), (b) and (c). (5×2=10 Marks)
- (a) Explain the concept and importance of hypothesis in communication of Results.
 - (b) Discuss the importance data visualization in big data decision-making considering R environment.
 - (c) Explain the various methods of qualitative data analysis.

TCS-703/TIT-703

310

B-68