



R17 Regulation

Subject code: 1E7EA

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

(Autonomous, Accredited by NAAC with 'A' Grade)

B.Tech IV Year I Semester Regular Examinations, February 2021

INTRODUCTION TO BIG DATA ANALYTICS

(Computer Science and Engineering)

Maximum Marks: 70

Date: 27.02.2021 Duration: 3 hours

- Note:**
1. This question paper contains two parts A and B.
 2. Part A is compulsory which carries 20 marks. Answer all questions in Part A.
 3. Part B consists of 5 Units. Answer any one full question from each unit.
 4. Each question carries 10 marks and may have a, b, c, d as sub questions.

Part-A

All the following questions carry equal marks

(10x2M=20 Marks)

- 1 What are the characteristics of big data?
- 2 When data becomes big data?
- 3 Explain MapReduce programming technique
- 4 Analyze the difficulties in ETL process
- 5 What are the issues in big data to meet business requirements?
- 6 Discuss the methods to eliminate outliers
- 7 List the categories of clustering methods
- 8 Define Hypothesis testing
- 9 Compare the visualization tools Tableau and Qlickview in terms of performance
- 10 Justify the importance of visualization in data analysis

Part-B

Answer All the following questions.

(10M X 5=50Marks)

- 11 a) List down the sources of live data. (5M)
b) Discuss the challenges in big data platform. (5M)
- OR
- 12 a) Explain any one cloud technology for big data handling. (5M)
b) Demonstrate the safe and secure environment with big data. (5M)
- 13 Compare the merits and demerits of Hadoop and spark system. (10M)
- OR
- 14 a) Define knowledge management. (5M)
b) Recommend a proper decision model for big data analysis. (5M)
- 15 Explain the different methods to analyze big data. (10M)
- OR
- 16 a) How description statistics are useful for effective analysis of huge data?(5M)
b) Write notes on Inter quartile range(IQR) with example. (5M)

- 17 Relate Machine learning with big data using COVID-19 dataset. (10M)
OR
- 18 a) List down data analytical methodologies. (5M)
b) Examine the purpose of regression and sentiment analysis. (5M)
- 19 a) Discuss the relation between visualization of data and product design. (5M)
b) What are the visualization techniques supports big data?(5M)
OR
- 20 a) How data analysis is done using Twitter data?(5M)
b) Explain the working principle of D3 tool. (5M)