

Code No: 155AT**JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD****B. Tech III Year I Semester Examinations, January/February - 2023****DATA ANALYTICS****(Common to CSE, IT)****Time: 3 Hours****Max. Marks: 75****Note:** i) Question paper consists of Part A, Part B.

ii) Part A is compulsory, which carries 25 marks. In Part A, Answer all questions.

iii) In Part B, Answer any one question from each unit. Each question carries 10 marks and may have a, b as sub questions.

PART – A**(25 Marks)**

- 1.a) Define Data Management. [2]
- b) Which source of Data is more reliable and Why? [3]
- c) What are the steps involved in Analysis of Data? [2]
- d) List out the features of Cloudera Impala. [3]
- e) Where can be Logistic Regression concepts applied? [2]
- f) Define Variable Rationalization. [3]
- g) Mention the applications of Time Series Analysis. [2]
- h) Name some Software to make Decision Tree. [3]
- i) What is Chernoff Faces technique? [2]
- j) What are the advantages of Data Visualization? [3]

PART – B**(50 Marks)**

- 2.a) What are the constraints and influences that will have an effect on Data Architecture Design? Explain.
- b) Discuss about the Data Reduction as a Data Preprocessing Step. [5+5]

OR

- 3.a) What is Secondary Data? Classify the Secondary Data Sources.
 - b) Discuss how data quality assessment can depend on the Intended Use of the data? [5+5]
- 4.a) List and explain the various Analytics techniques.
 - b) Explain the various methods to identify the Gaps in the Data and their handling mechanisms. [4+6]

OR

- 5.a) Discuss about the importance of Data Modeling in Business.
- b) What are the types of No SQL tools based on Data Models? Explain. [4+6]

- 6.a) What is meant by BLUE property? What are the blue properties of OLS method?
b) Discuss in detail about Multinomial Logistic Regression. [5+5]

OR

- 7.a) Explain the purpose of the Least Square Estimation in Regression with an example.
b) Discuss in short about the following Model Fit Statistics:
i) Hosmer Lemeshow Test ii) Error Matrix. [5+5]

- 8.a) What is Overfitting? How to Prevent Overfitting?
b) Compare and Contrast between ARMA and ARIMA. [5+5]

OR

- 9.a) Differentiate supervised and unsupervised learning.
b) Discuss the STL approach for Time Series Decomposition. [5+5]

- 10.a) Demonstrate Geometric projection visualization techniques.
b) List out various applications of data visualization. [5+5]

OR

- 11.a) Write down steps involved in Data Visualization In Tableau.
b) Explain about tag cloud visualization technique. [5+5]

---ooOoo---