

**Code No: 155AT****JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD****B. Tech III Year I Semester Examinations, August - 2022****DATA ANALYTICS****(Common to CSE, IT)****Time: 3 Hours****Max. Marks: 75**

**Answer any five questions**  
**All questions carry equal marks**

---

- 1.a) What is data? How to handle large collections of data?  
b) Discuss the various sources of data. [7+8]
- 2.a) How to identify data quality? What are the quality measures for data?  
b) Explain the need of data preprocessing with illustrations. [7+8]
- 3.a) Contrast nominal, ordinal and ratio-scaled data.  
b) Explain the applications of data modeling in business. [7+8]
- 4.a) Illustrate data imputations techniques.  
b) Describe the important features of Cloudera Impala. [7+8]
- 5.a) Explain ordinary least squares regression with an example.  
b) Illustrate Hosmer–Lemeshow test for goodness of fit of logistic regression. [7+8]
- 6.a) Discuss the significance of ROC analysis and ROC curve.  
b) Demonstrate variable rationalization in regression. [7+8]
- 7.a) Explain decision tree induction approach.  
b) How to extract features from time-series data? Explain with an example. [7+8]
- 8.a) Elucidate visualization of complex data and relations.  
b) Describe parallel coordinates and land scapes for geometric data visualization. [7+8]

---oo0oo---