



QP CODE: 25022659



Reg No : .....

Name : .....

**M.Sc DEGREE (CSS) SPECIAL REAPPEARANCE EXAMINATION, APRIL 2025**

**Third Semester**

M.Sc COMPUTER SCIENCE ( Data Analytics)

**CORE - CA030302 - EXPLORATORY DATA ANALYTICS FOR NLP**

2019 ADMISSION ONWARDS

5AADA7E1

Time: 3 Hours

Weightage: 30

**Part A (Short Answer Questions)**

*Answer any **eight** questions.*

*Weight **1** each.*

1. What do you mean by relational perspectives in natural language processing?
2. What is a dispersion plot?
3. How is strings are used for NLP ?
4. List some of the reasons for using exploratory data analysis.
5. Create a two dimensional array myArray=[[2, 3, 4, 5], [2, 4, 6, 8], [1, 3, 5, 7]]. Write the code snippet for the following:
  - a)To print the number of myArray's dimension.
  - b)To print the number of myArray's elements.
6. What are the advantages of data transformation?
7. Explain Left merge and right merge.
8. Explain Drill down and roll up operation.
9. Define p-hacking.
10. What are the two categories of supervised learning algorithms?

(8×1=8 weightage)

**Part B (Short Essay/Problems)**

*Answer any **six** questions.*

*Weight **2** each.*

11. What are the different phases of natural language processing?
12. Explain NLP, NLU, and NLG in detail.
13. Describe the significance of EDA.





14. Briefly explain different types of measurement scales.
15. List and explain types of outliers with help of suitable diagrams.
16. Why data transformation is important and explains the methods.
17. Briefly explain multiple linear regression model.
18. What are the different stages in data preprocessing?

(6×2=12 weightage)

**Part C (Essay Type Questions)**

*Answer any **two** questions.*

*Weight **5** each.*

19. Explain the computing with language in python.
20. Discuss in detail about the different visual aids for EDA.
21. Explain missing data handling in NLP, give the example.
22. Explain in detail about Regression Analysis.

(2×5=10 weightage)

