

Roll No



**PRESIDENCY UNIVERSITY
BENGALURU**

SET A

**SCHOOL OF ENGINEERING
END TERM EXAMINATION - JAN 2024**

Semester : Semester III - 2022

Course Code : CSE2027

Course Name : Fundamentals of Data Analytics

Program : B.Tech.

Date : 09-JAN-2024

Time : 9:30AM - 12:30 PM

Max Marks : 100

Weightage : 50%

Instructions:

- (i) Read all questions carefully and answer accordingly.
- (ii) Question paper consists of 3 parts.
- (iii) Scientific and non-programmable calculator are permitted.
- (iv) Do not write any information on the question paper other than Roll Number.

PART A

ANSWER ALL THE QUESTIONS

5 X 2M = 10M

1. Explain the metrics for evaluation of a Regression Model.
(CO5) [Knowledge]
2. Explain the role of an enumerator in the collection of primary data and also mention which Data collection method uses enumerator.
(CO4,CO3) [Knowledge]
3. List the two basic types of descriptive statistics used in Data Analytics.
(CO3,CO2) [Knowledge]
4. List Four major differences between Questionnaires and Schedules.
(CO3) [Knowledge]
5. Explain the concept of Data Transformation.
(CO1) [Knowledge]

PART B

ANSWER ALL THE QUESTIONS

5 X 10M = 50M

6. Explain in detail about Building and Applying a Prediction Model.
(CO2,CO5) [Comprehension]
7. Explain in details about the rules for the good report preparation.
(CO4,CO5) [Comprehension]
8. Explain in details different types of interview method for data collection.
(CO4,CO3) [Comprehension]

9. Prepare a questionnaire to evaluate the effectiveness of teaching methods and assess student learning experiences within our educational institution. The questionnaire seeks to gather data on various aspects, including instructional techniques, course content relevance, and student engagement, to enhance the overall quality of education.

(CO3,CO2) [Comprehension]

10. Suppose a study of speeding violations and drivers who use cell phones produced the following fictional data

| | Speeding violation in the last year | No speeding violation in the last year | Total |
|-----------------------|-------------------------------------|--|-------|
| Cell Phone user | 25 | 280 | |
| Not a cell phone user | 45 | 405 | |
| Total | | | |

- Complete the contingency table
- Find $P(\text{Person is a car phone user})$
- Find $P(\text{person had no violation in the last year})$
- Find $P(\text{person had no violation in the last year AND was a cell phone user})$
- Find $P(\text{person is a cell phone user OR person had no violation in the last year})$

(CO3,CO2) [Comprehension]

PART C

ANSWER ALL THE QUESTIONS

2 X 20M = 40M

11. In a random sample of 100 tube lights produced by company A, the mean life time of tube lights is 1190 hrs with standard deviation of 90 hrs. Also, in a random sample of 75 tube lights from company B, the mean life time is 1230 hrs with a standard deviation of 120 hrs. Is there a difference between the mean life times of the two brands of tube lights at a significance level of (a) 0.05 and (b) 0.01 ?

| Level of significance α | Two tailed test | Right tailed test | Left tailed test |
|---|-----------------|-------------------|------------------|
| 90% confidence or $\alpha = 10\% = 0.1$ | $z = 1.645$ | $z = 1.28$ | $z = -1.28$ |
| 95% confidence or $\alpha = 5\% = 0.05$ | $z = 1.96$ | $z = 1.645$ | $z = -1.645$ |
| 99% confidence or $\alpha = 1\% = 0.01$ | $z = 2.58$ | $z = 2.33$ | $z = -2.33$ |

(CO2) [Application]

12. The results of the measurement of electric **resistance R** of a copper bar at various temperatures **t degree centigrade** are listed below

| | | | | | | | |
|-----|----|----|----|----|----|----|----|
| t | 19 | 25 | 30 | 36 | 40 | 45 | 50 |
| R | 76 | 77 | 79 | 80 | 82 | 83 | 85 |

Find a relation $R=a+ bt$, where a and b are constants to be determined. Also find the MSE, RMSE and MAE using the best fit regression line.

(CO5) [Application]