



QP CODE: 25050795



25050795

Reg No :

Name :

MSc DEGREE (CSS) EXAMINATION DECEMBER 2025

First Semester

M Sc Computer Science (Data Analytics)

Core Courses - CA030101 - STATISTICS FOR DATA ANALYTICS

2020 Admission Onwards

DCD26FE2

Time: 3 Hours

Weightage: 30

Part A (Short Answer Questions)

*Answer any **eight** questions.*

Weight 1 each.

1. What will be the sample space for the interval $[4, 10]$?
2. State the three axioms in probability theory.
3. What are the graphical representations of frequency distribution?
4. What is the significance of Kurtosis? Illustrate the three different types of curves with the help of a diagram.
5. Give an example for a discrete random variable.
6. Give an example for a continuous random variable.
7. Distinguish between positive correlation and negative correlation with example.
8. What is linear regression analysis?
9. Distinguish between continuous time series and discrete time series.
10. What are the utilities of a time plot?

(8×1=8 weightage)

Part B (Short Essay/Problems)

*Answer any **six** questions.*

Weight 2 each.

11. State and prove theorems on probability.
12. A dart is thrown at a circular dartboard such that it will land randomly over the area of the dartboard. What is the probability that it lands closer to the center than to the edge?
13. The following distribution gives a pattern of overtime work done by 100 employees in a company. Find the mean and median.





Overtime(Hrs)	10-15	15-20	20-25	25-30	30-35	35-40
No.of Employees	12	35	6	20	8	20

14. Calculate the range and quartile deviation of the following data:
20, 25, 29, 30, 35, 39, 41, 48, 51, 60 and 70
15. Define the distribution function of a random variable. What are the properties of the distribution function?
16. Define Sampling. Discuss any two methods of sampling.
17. Explain the direct method of finding correlation with an example.
18. Explain about the different methods used in the analysis of time series with a trend.

(6×2=12 weightage)

Part C (Essay Type Questions)

Answer any **two** questions.

Weight 5 each.

19. Calculate mean and standard deviation for the following table giving the age distribution of 542 members.

Age in years	20-30	30-40	40-50	50-60	60-70	70-80	80-90
Number of members	3	61	132	153	140	51	2

20. The probability that a bulb produced by the factory will fuse after 150 days of use is 0.05. Find the probability that out of five such bulbs,
a) None of them will fuse after 150 days of use.
b) not more than one will fuse after 150 days of use.
c) more than one will fuse after 150 days.
d) at least one will fuse after 150 days.
21. In a study of the relationship between X=mean daily temperature for the month and Y=monthly charges on electric bill, the following data was gathered.

Calculate the regression equation.

X	20	30	50	60	80	90
Y	125	110	95	90	110	130

22. Explain in detail about data analytics life cycle.

(2×5=10 weightage)

