



21002269

QP CODE: 21002269

Reg No : .....

Name : .....

**M Sc DEGREE (CSS) EXAMINATION, NOVEMBER 2021**

**First Semester**

M.Sc.Computer Science (Data Analytics)

**CORE - CA030101 - STATISTICS FOR DATA ANALYTICS**

2019 ADMISSION ONWARDS

D495613B

Time: 3 Hours

Weightage: 30

**Part A (Short Answer Questions)**

*Answer any **eight** questions.*

*Weight 1 each.*

1. Define probability.
2. What are Equally Likely events? Give an example.
3. Calculate the harmonic average from the following data:  
1,0.5,10,45.0,175.0,0.01,4.0,11.2
4. Distinguish between coefficient of dispersion and coefficient of variation.
5. Define continuous uniform distribution.
6. In a class there are 96 students with Roll Nos. from 1 to 96. It is desired to take sample of 10 students. Use the systematic sampling method to determine the sample.
7. What is the purpose of identifying correlation between two or more variables?
8. What is linear regression analysis?
9. What are the main goals of analysing trends in time series?
10. Name any four tools used in data analytics.

(8×1=8 weightage)

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

*Answer any **six** questions.*

*Weight 2 each.*

11. State and prove theorems on probability.
12. State and explain Bayes' theorem.





13. Mean of 200 items was 50. Later it was found that the two items were misread as 92 and 8 instead of 192 and 88. Find the correct mean.
14. For a dataset difference between Q3 and Q2 is 100 and difference between Q2 and Q1 is 120. Find coefficient of skewness.
15. A pair of dice is thrown four times. If getting a doublet is considered as success, find the probabilities of two successes.
16. Define binomial distribution. Find the mean and variance of the binomial distribution.
17. Explain the direct method of finding correlation with an example.
18. Briefly explain about analysis of time series which contains seasonal variation.

(6×2=12 weightage)

**Part C (Essay Type Questions)**

*Answer any **two** questions.*

*Weight 5 each.*

19. **Find the weighted geometric mean from the following data:**

Group	Index Number	Weights
Food	260	46
Fuel & Lighting	180	10
Clothing	220	8
House & Rent	230	20
Education	120	12
Misc.	200	4

20. a) Define the probability density function of a random variable.  
b) Let X is a random variable which denotes the number of heads obtained when a coin is tossed 3 times. Find the probability distribution of the random variable X and mean and variance of the distribution.





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 the different types of variations in time series.

(2×5=10 weightage)

