

2056
B.Sc. Data Analytics (FYUP)
Second Semester
BDA-203 - Basics of Statistics

Time allowed: 3 Hours

Max. Marks: 90

NOTE: Attempt five questions in all, including Question No. 9 (Section- E) which is compulsory and selecting one question from each Section A-D. Simple calculator without programming is allowed. Log tables must be provided.

x-x-x

Section-A

1. a) Define inferential statistics. How it is different from descriptive statistics? Explain the types of descriptive statics.
- b) What is Geometric mean? Discuss its significance. The following is the distribution of marks obtained by 109 students in a subject in an institution. Find the Geometric mean.

Marks	Number of Students
4-8	6
8-12	10
12-16	18
16-20	30
20-24	15
24-28	12
28-32	10
32-36	6
36-40	2

(2x9)

2. Explain the arithmetic mean, geometric mean and harmonic mean with reference to significance, purpose, merits and demerits of each. Write the formula for each of the mean methods considering there are two data points - a and b. (12+6)

P.T.O.

(2)

Section - B

3. Write short notes on:

- a) Quartile
- b) Percentile
- c) Range
- d) Mode
- e) Median
- f) Septile

(6x3)

4. What is variance? How is it different from standard deviation? Why do we calculate variance? Find the variance of following data:

Length of Service (x)	Frequency (f)
0-10	30
10-20	42
20-30	23
30-40	13
40-50	8

(18)

Section - C

5. What is correlation? Discuss the types of correlation. What is the significance of correlation? How correlation and causation are different? Explain the steps to calculate rank correlation. (18)

6. Calculate the Karl Pearson's correlation coefficient on following data:

A doctor collected data from 8 patients to study the relationship between fasting insulin level and diabetes severity score.

Contd.....P/3



(3)

Patient	Insulin Level (X)	Diabetes Score (Y)
1	8	42
2	10	48
3	12	52
4	9	45
5	14	60
6	11	50
7	15	65
8	13	58

(18)

Section - D

7. a) Differentiate between correlation and regression.

b) A researcher studied the relationship between Hours of Study (X) and Marks Obtained (Y) for students. The data is given in grouped (frequency) form.

Calculate the regression equation of Y on X.

Hours of Study (X)	Marks (Y)	Frequency (f)
1	40	2
2	45	3
3	50	5
4	55	4
5	65	3
6	70	2

(8,10)

P.T.O.

(4)

8. a) What is regression? Discuss its types and significance. Enumerate the properties of regression coefficient.
- b) A psychologist studied the relationship between **Stress Score (X)** and **Blood Sugar Level (Y)** of patients.

Calculate the **regression coefficients** of Y on X and X on Y

Stress Score (X)	Blood Sugar (Y)
10	85
12	88
14	92
16	95
18	98
20	105

(6,12)

Section-E

9. Attempt all questions given below:
- a) Find the mean and mode of following data: 2,3,4,3,4,3,6 . (4)
- b) What is primary data? (2)
- c) What are hexailes? Give the formula for first hexaile. (4)
- d) What is cumulative frequency? (2)
- e) What does a correlation of 1, 0 and -1 indicate? Assume correlation lies between interval $[1,-1]$. (2)
- f) What is the significance of regression? (2)
- g) What is probable error? (2)

x-x-x