Exam.Code:5152
Sub. Code: 14363

2055
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) What is Geometric Mean? How it is related to Arithmetic Mean? Find the geometric and Arithmetic mean of the following data:

Weights of ear heads (g)	No of ear heads (f)
60-80	22
80-100	38
100-120	45
120-140	35
140-160	20
Total	160

b) What are measures of central tendency? Discuss and compare all its types. (2x9)

2. a) What is Harmonic Mean? Calculate the harmonic mean of the following data:

x	1	3	5	7	9	11
f	2	4	6	8	10	12

b) What are various types of statistical techniques? Discuss various statistical methods and their significance. (8,10)

Section - B

3. a) What is Median? Write the steps to find median for grouped data. How median is different from quartiles? Calculate the median and first quartile for the following data: (2+3+2+3+3)

Marks out of 50	Frequency
0-10	2
10-20	4
20-30	5
30-40	4
40-50	2

b) Write short note on inter-quartile range. (5)

P.T.O.

4. a) What are measures of dispersion? Describe all its types. (5)

   b) What is standard deviation? Write the steps to calculate standard deviation. Find the standard deviation of following data:

   The grouped frequency table shows the length of service in years of employees who have been working for a company for at least ten years. Calculate an estimate of the standard deviation of the length of service of these employees.

Length of Service (x)	Frequency (f)
$10 \le x < 15$	30
$15 \le x < 20$	42
$20 \le x < 25$	23
$25 \le x < 30$	13
$30 \le x < 40$	8
$40 \le x < 50$	4

(2+5+6)

Section - C

5. a) While calculating correlation coefficient (Pearson) between two variables x and y for 30 pairs of observation, the students observed the following results:

   $\Sigma x = 135$,
   $\Sigma x^2 = 680$,
   $\Sigma y = 110$,
   $\Sigma y^2 = 485$,
   and
   $\Sigma xy = 530$.

   On rechecking, it was found that he had wrongly copied one pair as: (5,8) whereas value is (7,9) calculate the correct correlation coefficient (pearson) between x and y.

   b) What is correlation? Discuss its significance. Explain the different types of correlations. (8,10)

6. a) What is a scatter plot? make a scatter plot for the following data:

No. of games	3	5	6	7	1	2
Scores	80	90	80	90	50	65

(8)

b) What is rank coefficient? Explain the difference between Rank Coefficient and Karl Pearson's coefficient of correlation. Calculate the rank correlation for data given below:

Distance from CAM (m)	Price of 50cl bottle (€)
50	1.80
175	1.20
270	2.00
375	1.00
425	1.00
580	1.20
710	0.80
790	0.60
890	1.00
980	0.85

(2+2+6)

Section - D

7. a) What is regression? How it is different from correlation? What are the limitations of regression? (2+3+4)

b) Find the equation of regression line and the value of y when x is 19.

Time, x (seconds)	5	7	12	16	20
Mass, y (grams)	40	120	180	210	240

(6+3)

8. **Write short notes on**

a) Regression lines
b) Regression equations
c) Standard error
d) Types of regression (4+4+4+6)

Section - E

9. Attempt the following:-

a) What is primary data?
b) What are hexailes?
c) How mode is different from mean?
d) What is the frequency of different numbers in following series:
2,3,3,4,4,5,2,3,4,5,2,2,3,4
e) What is correlation?
f) What does a correlation of 1, 0 and -1 indicate? Assume correlation lies between interval [1,-1] .
g) What is the significance of Standard deviation?
h) What is probable error?
i) Give example of a continuous series. (9x2)

x-x-x