BBA

Semester – II
Take Home Assignment
2023

Course: Data Analysis for Business-I

Maximum marks: 100
Weightage: 40%
A viva of 10 marks will be conducted upon submitting the assignment.

INSTRUCTIONS TO CANDIDATES

1. The exam is **Closed Book.**
2. **Attempt any 5 questions out of 6**
3. Calculator is allowed.

1. The following data present number of units of spare parts sold during last 50 days in a company: [5+5+5+5]

100	173	202	178	147	102	153	197	127	101
157	185	190	116	172	148	213	130	165	111
141	149	206	175	123	128	144	168	109	167
100	163	150	154	130	143	186	166	139	149
108	119	183	151	114	135	191	137	129	258

(a) Prepare a frequency distribution starting from 100 with class width 20. Draw a histogram for the frequency distribution.
(b) Draw a frequency polygon for the same frequency distribution.
(c) Calculate the cumulative frequencies for the frequency distribution in Q(a) and draw an ogive for the data.
(d) Calculate mean number of units sold for the above data.

2. The following data present scores of 100 students in one of their examination. [3+8+9]

Scores in an Exam	Less than 15	15-20	20-25	25-30	30-35	35-40	40-45	More than 45
number of students	3	13	16	17	28	13	7	3

(a) Can you calculate mean marks for the above data? If yes, calculate the mean value. Else state your reasons for not calculating mean.
(b) Can you calculate median and mode marks for the above data? If yes, calculate the median and mode values. Else state your reasons for not calculating them.
(c) Calculate 3rd quartile (Q_3), 6th decile (D_6) and 38th percentile (P_{38}) for the given data.

3. (a) What do you mean by measures of dispersion? Mention names and formulae of any five measures of dispersion. [6+8+6]
(b) The following data represent the time (in minutes) to process customer complaints in a national bank in its two branches, Main City Branch and Countryside Branch:

Main City Branch	1.48	1.75	0.78	2.85	0.52	1.6	4.2	3.97	1.48	3.1
Countryside Branch	7.55	3.75	0.1	1.1	0.6	0.52	3.3	2.1	0.58	4

Calculate standard deviation for both the branches and comment on which branch has less variation in processing the complaints.

(c) Data have been collected on three variables, viz., price of commodity (in Rs.), number of purchases (in units) and shelf space (in square feet) as presented below. Use an appropriate measure of dispersion and comment on which dataset has got highest level of variation.

|	sample size	mean	standard deviation
I. price of commodity (in Rs)	10	436	23.4
II. number of purchases (in units)	10	26	5
III. shelf space used for commodities (in square feet)	10	4	1.5

4. Answer the following: [16+2+2]
(a) Compare and contrast the three measures of central tendency, mean, median and mode with respect to the criteria of ideal measure of central tendency. Discuss at least 8 criteria.
(b) For a moderately symmetric distribution, if mean=43, median=40, find mode value.
(c) For a symmetric distribution, if mode=26, calculate mean and median.

5. What do you mean by correlation between two variables? For the following data calculate Karl Pearson's correlation coefficient. Comment on the nature of relationship between two variables. [20]

Marks in Statistics (out of 50)	45	39	12	49	34	33	23	45	28	39
Marks in Accountancy (out of 50)	40	42	20	43	38	19	34	40	30	46

6. (a) For the following scatter plots, give the appropriate value/range of correlation coefficient (r):   [10+10]

(i)   X

(ii)   X

(iii)   X

(iv)   X

Y=X

(v)   X
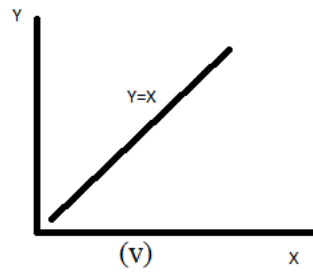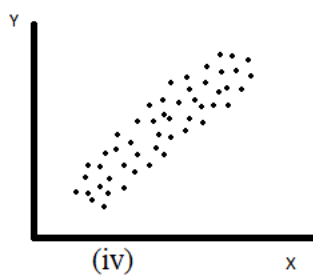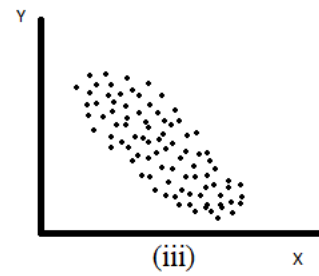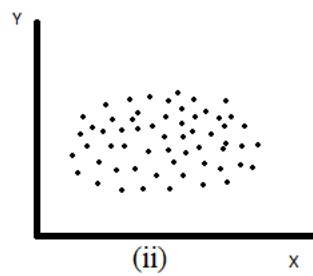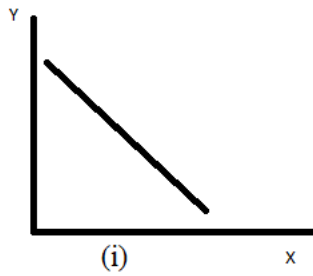
(b) Explain the following with proper examples:
(i) nominal scaled data
(ii) ordinal scaled data
(iii) interval scaled data
(iv) ratio scaled data