Statistics Question Bank

First Paper

Abdullah Al Mahmud

www.statmania.info

Contents

Chapter 1 Statistics, Variable and Concepts of Different Symbols **1**
1.1 Creative Questions . . . 1
1.2 Short Questions . . . 6
Chapter 2 Collection, Presentation, and Organization of Data **9**
2.1 Creative Questions . . . 9
2.2 Short Questions . . . 11
Chapter 3 Measures of Central Tendency **12**
3.1 Creative Questions . . . 12
3.2 Short Questions . . . 19
3.2.1 General Questions . . . 19
3.2.2 Arithmetic Mean . . . 20
3.2.3 Geometric Mean . . . 21
3.2.4 Harmonic Mean . . . 21
3.2.5 Median . . . 21
3.2.6 Mode . . . 21
3.2.7 Quadratic Mean . . . 22
3.2.8 Partition Values . . . 22
Chapter 4 Measures of Dispersion **23**
4.1 Creative Questions . . . 23
4.2 Short Questions . . . 25
4.2.1 Variance . . . 26
Chapter 5 Moments, SKewness, and Kurtosis **27**
5.1 Creative Questions . . . 27
5.2 Short Questions . . . 30
5.3 Broad Questions . . . 31
Chapter 6 Correlation and Regression **32**
6.1 Creative Questions . . . 32
6.2 Short Questions . . . 33
Chapter 7 Time Series **34**
7.1 Creative Questions . . . 34
7.2 Short Questions . . . 36
Chapter 8 Published Statistics in Bangladesh **38**
8.1 Creative Questions . . . 38
8.2 Short Questions . . . 39
Conclusion **40**

Chapter 1

Statistics, Variable and Concepts of Different Symbols

1.1 Creative Questions

1. **A system analyst collected frequencies of a signal at different times. Then he realized due to some unknown noise, 0.5 units got added to all the values. The recorded values are given below:**

   10, 12, 15, 14, 12, 16, 20, 16, 18, 11

   (a) What is change of origin? 1

   (b) Does change of origin have an effect on median? 2

   (c) Find $\sum_{i=1}^{10}(X_i - 5)$. 3

   (d) Determine the summation of the values discarding the noise. 4

2. **The daily website visits (in thousands) for an online platform over seven days are recorded as 80, 85, 90, 95, 100, 105, and 110 (denoted by z). The platform analyst claimed that the square of the total visits is greater than the total of the squared visits.**

   (a) Calculate $\sum_{i=1}^{7}(z_i - 2z_i)^2$ using the provided data. 3

   (b) Verify whether the analyst's statement is accurate based on the data. 4

3. **The weekly sales (in hundreds of dollars) for a retail store over six weeks are recorded as 12, 15, 18, 20, 22, and 25 (denoted by y). The store manager claimed that the square of total sales exceeds the total of squared sales.**

   (a) Calculate $\sum_{i=1}^{6}(y_i - 3y_i)^2$ using the provided data. 3

   (b) Determine whether the manager's claim is correct based on the data. 4

4. **A software developer tracked the response times (in milliseconds) of a web application under test during peak usage hours. An unexpected delay of 1.2 ms was added to each recorded response time due to server lag. The recorded times are as follows:**

   45, 48, 52, 50, 47, 53, 60, 55, 58, 49

   (a) Calculate $\sum_{i=1}^{10}(X_i - 25)$. 3

   (b) Find the sum of the original response times before the lag was added. 4

5. **The monthly sales (in thousands of units) recorded by a store over five months are 120, 135, 150, 160, and 175. The store manager stated that the square of the total sales is greater than the total of the squared sales.**

   (a) Calculate $\sum_{i=1}^{5}(x_i - 1.5x_i)^2$ using the provided data. 3

(b) Assess whether the manager's statement is accurate based on the data. 4

6. **A nutritionist is analyzing the daily protein intake (in grams) of five athletes. The recorded values are:**

$$x_1 = 55, \quad x_2 = 60, \quad x_3 = 48, \quad x_4 = 62, \quad x_5 = 50$$

(a) Compute the value of $\sum_{i=1}^{5}(x_i - 52)^2$. 3

(b) Calculate $\sum_{i=1}^{5}(x_i^2 - 4x_i + 10)$ using both direct evaluation and by separating the summation terms. 4

7. **Marks obtained by five students in statisics out of 15 were 4, 6, 10, 12, and 15. The examiner said, the square of the sum of the marks is greater than the sum of the squares of the marks.**

(a) What is finite population? 1

(b) Explain quantitative variable with an example. 2

(c) In the light of the available data, find $\sum_{i=1}^{5}(x_i - 2x_i)^2$ 3

(d) Verify the comment of the examiner. 4

8. **The revenue and expenses (in million BDT) of different business sectors are shown below:**

Sector	1	2	3	4	5
Revenue (X)	45	32	50	60	40
Expenses (Y)	30	20	35	50	25

(a) Find the value of $\sum_{i=1}^{5}\sum_{j=1}^{5}(x_i - y_j)$. 3

(b) Examine the statement theoretically and empirically: $\sum_{i=1}^{5}(3x_i - 4y_j) = 3\sum_{i=1}^{5}x_i - 4\sum_{i=1}^{5}y_j$. 4

9. **The capital and profit (in million BDT) of some Bangladeshi industries are given below:**

Industry	1	2	3	4	5
Capital (X)	20	15	26	31	18
Profit (Y)	15	10	17	25	10

(a) What is finite population? 1

(b) What are the functions of statistics? 2

(c) Find the value of $\sum_{i=1}^{5}\sum_{j=1}^{5}(x_i - y_j)$ 3

(d) Analyze the statement theoretically and empirically: $\sum_{i=1}^{5}(4x_i - 6y_j) = 4\sum_{i=1}^{5}x_i - 6\sum_{i=1}^{5}y_j$ 4

10. **The revenue and expenditure (in million BDT) of five companies in Bangladesh are given below:**

Company	A	B	C	D	E
Revenue (X)	50	40	60	55	45
Expenditure (Y)	30	25	35	40	28

(a) Compute $\sum_{i=1}^{5}(x_i - y_i)$. Interpret the result in terms of net earnings. 3

(b) Show that $\sum_{i=1}^{5}(3x_i - 2y_i) = 3\sum_{i=1}^{5}x_i - 2\sum_{i=1}^{5}y_i$ theoretically and empirically. 4

Store	A	B	C	D	E
Sales (X)	50	65	40	70	55
Expenses (Y)	30	45	25	50	35

11. **The monthly sales and expenses (in thousand BDT) of five retail stores are given below:**

    (a) Calculate $\sum_{i=1}^{5}(x_i + y_i)$ 3

    (b) Verify whether the statement $\sum_{i=1}^{5}(3x_i - 2y_i) = 3\sum_{i=1}^{5} x_i - 2\sum_{i=1}^{5} y_i$ holds true. 4

12. **Goals scored by a footballer in 25 matches are summarized as shown below.**

Goals	0	1	2	3	4
Times	8	9	5	2	1

    (a) Is no. goals a discrete or continuous variable? 1

    (b) Verify theoretically: $\sum_{i=1}^{2} X_i Y_i = \sum_{i=1}^{2} X_i \times \sum_{i=1}^{2} Y_i$ 2

    (c) Find the total number of goals using a suitable notation. 3

    (d) If he scores two (2) goals in the next match, will the scoring rate increase? Explain logically and empirically 4

13. **Points scored by a basketball player in 20 games are summarized as shown below.**

Points	0	1	2	3	4
Games	4	6	5	3	2

    (a) Find the total number of points scored using appropriate notation. 3

    (b) If he scores five (5) points in the next game, will the average scoring rate increase? Justify your answer with calculations. 4

14. **A researcher conducting a study on economic indicators has gathered the following data points, representing the net income (in million BDT) of five different companies over a fiscal year:**

$$x_1 = 15, x_2 = -12, x_3 = 10, x_4 = 21, x_5 = 33$$

    (a) What is sample? 1

    (b) Briefly explain shift of origin and scale. 2

    (c) Compute the value of $\sum_{i=1}^{5}(x_i - 20)^2$ 3

    (d) Find the value of $\sum_{i=1}^{5}(3x_i^2 - 2x_i - 3)$ and examine its dependency on origin and scale. 4

15. **A health researcher recorded the daily steps (in thousands) of five individuals over a week:**

$$y_1 = 8,\ y_2 = 5,\ y_3 = 12,\ y_4 = 10,\ y_5 = 7$$

    (a) Compute $\sum_{i=1}^{5}(y_i - 10)^2$. 3

    (b) Evaluate $\sum_{i=1}^{5}(2y_i^2 - 3y_i + 4)$ and comment on how shifting the origin would affect this sum. 4

16. **A health specialist is analyzing the daily calorie intake (in hundreds of calories) of five individuals on a specific diet. The recorded values are:**

$$y_1 = 24, y_2 = 18, y_3 = 30, y_4 = 26, y_5 = 16$$

(a) Compute the value of $\sum_{i=1}^{5}(y_i - 20)^2$. 3

(b) Calculate the values of $\sum_{i=1}^{5}(3y_i^2 - 4y_i + 7)$ and $\sum_{i=1}^{3}(y_i^2 - 5y_i)$. 4

17. **A psychologist is studying the stress levels (measured on a scale of 1 to 50) experienced by five participants during a specific task. The observed values are:**

$$x_1 = 32, x_2 = 28, x_3 = 40, x_4 = 35, x_5 = 22$$

(a) Compute the value of $\sum_{i=1}^{5}(x_i - 30)^2$. 3

(b) Calculate $\sum_{i=1}^{5}(2x_i^2 - 5x_i + 4)$ using both a direct approach and by splitting the summation terms. 4

18. **A botanist is measuring the heights (in centimeters) of four seedlings after one week. The observed heights are:**

$$h_1 = 15, h_2 = 12, h_3 = 18, h_4 = 10$$

(a) Compute the value of $\sum_{i=1}^{4}(h_i - 14)^2$. 3

(b) Calculate $\sum_{i=1}^{4}(3h_i^2 - 2h_i + 1)$ using both a direct approach and by splitting the summation terms. Also demonstrate that they are mathematically equivalent. 4

19. **Height (in inches) of 10 cadets in a class are: 50, 60, 55, 65, 66, 70, 54, 64, 62, 72**

(a) What is population in statistics? 1

(b) Is height discrete or continuous? 2

(c) Find $\sum_{i=1}^{10} x_i^2$ 3

(d) Find the square of mean and mean of square. Are they equal? 4

20. **Marks of 10 students in Statistics in a class were found to be the following:**

99, 88, 98, 85, 97, 71, 87, 79, 70, 84

Later it was discovered that all marks should be 5 less than the recorded marks.

(a) What is change of origin? 1

(b) Does summation of a variable depend an change of origin? 2

(c) Considering the data in stem as X, find $\sum_{i=1}^{10} X_i$ and $\sum_{i=1}^{10}(X_i + 3)$ 3

(d) Find the arithmetic mean of the corrected values, employing the concept of shift of origin. 4

21. **Income and expenditure (both in thousands) of some individuals in four successive months are collected:**

Income (x)	20	30	25	10
Expenditure (y)	15	27	18	5

(a) What is a discrete variable? 1

(b) Can fractional numbers be discrete? Explain briefly. 2

(c) Find $\sum_{i=1}^{4}(x_i)(y_i - 3)$.

Income (x)	20	30	25	10	35	40	28	22
Expenditure (y)	15	27	18	5	30	33	25	20

(d) Are, in the stem, $\sum_{i=1}^{n}\sum_{i=1}^{n} x_i y_j = \sum_{i=1}^{n} x_i y_i$? Vindicate. 4

22. **Income and Expenditure (in thousand BDT) data for 8 individuals in a housing society are given below:**

(a) Using the data, prove $\sum_{i=1}^{10}\sum_{j=1}^{10} x_i y_j = (\sum_{i=1}^{10} x_i)(\sum_{j=1}^{10} y_j)$ 3

(b) Find the value of $\sum_{i=1}^{10}(x_i - y_i)^2$ using both direct evaluation and by separating the summation terms. 4

23. **Call duration of 6 calls in a customer care center are**

2, 2.5, 1.5, 5, 6, 3

(a) What is a sample? 1

(b) Are all quantitative variables continuous? 2

(c) Determine $\sum_{i=1}^{7}(x_i - 3)^3$ 3

(d) Find the values of $\sum_{i=1}^{7}(x_i - 5)^2$ and $\sum_{i=1}^{7} x_i^2 + 5$. 4
Explain mathematically why they are unequal.

24. $x_1 = 5, x_2 = 10, x_3 = 12, x_4 = 14, x_5 = 20$

(a) What is a sample? 1

(b) Find $\sum_{i=1}^{5}(x_i - 5)$ 2

(c) Determine $\sum_{i=1}^{5} 3(x_i - c)^2$, where c is the last digit of your cadet number. 3

(d) Find mean of square and square of mean from the data. When will they be equal?

25. **Goals scored by Karim Benzema in five seasons are recorded to be the following:**

Season	La Liga (x)	Uefa Champions League (y)
2017-18	5	5
2018-19	21	4
2019-20	21	5
2020-21	23	6
2021-22	27	15

(a) What is a quantitative variable? 1

(b) What is the notation to denote his total number of goals? 2

(c) Compute $\sum_{i=1}^{5}(y_i - 3)^2$

(d) Find total number of goals using two different notations and examine whether they match. 4

26. **Below are some information, with two variables x and y:**

x_i	3	4	1	0
y_i	1	5	0	2

(a) What is a qualitative variable? 1

(b) Find $\sum_{i=1}^{4} x_i^2$ 2

(c) Prove that $\sum_{i=1}^{4}(x_i + y_i) = \sum_{i=1}^{4} x_i + \sum_{i=1}^{4} y_i$ 3

(d) Find the value of $\sum_{i=1}^{4} x_i y_i - \sum_{i=1}^{4} x_i + 4$ 4

27. **An analyst obtains some data:**

$$x_1 = 15, x_2 = -12, x_3 = 17, x_4 = 11, x_5 = 23$$

(a) What is sample? 1

(b) Briefly explain shift or origin and scale. 2

(c) Compute the value of $\sum_{i=1}^{5}(x_i - 10)^2$ 3

(d) Find the value of $\sum_{i=1}^{5}(5x_i^2 - 4x_i - 3)$ and examine its dependency on origin and scale. 4

28. **Number of accidents occurred in two different cities A and B are denoted by two variables x and y, respectively.**

City	Day 1	Day 2	Day 3
A	3	4	2
B	5	3	2

An analyst observed that $\left(\sum_{i=1}^{3} x_i\right)^2 \neq \sum_{i=1}^{3} x_i^2$ and $\left(\sum_{i=1}^{3} y_i\right)^2 \neq \sum_{i=1}^{3} y_i^2$

(a) Show that

$$\sum_{i=1}^{3}(x_i + y_i)^2 = \sum_{i=1}^{3} x_i^2 + 2\sum_{i=1}^{3} x_i y_i + \sum_{i=1}^{3} y_i^2$$

3

(b) Evaluate the analyst's comment. 4

1.2 Short Questions

1. $x_1 = 2, x_2 = -3, x_3 = 7, x_4 = 12$.

Find the values of the following: $4 \times 1.5 = 6$

i) $\sum_{i=1}^{3} x_i$ ii) $\sum_{i=1}^{4} x_i^2$ iii) $\sum_{i=1}^{4}(x_i + 3)$ iv) $\sum_{i=1}^{4}(x_i - 4)^2$

2. **Write down the scales of measurement of the following variables.** $8 \times 0.5 = 4$

Gender, Religion, Temperature, Income group (Lower class, Low, Middle, High)

Income, Distance of stars, Radius of screws, Room no.

3. **Distinguish between the qualitative and quantitative variable.** $6 \times 0.5 = 3$

Diameter of trees, Color, Weight, Gender, Jersey Number, Family Size

4. **Distinguish between the discrete and continuous variable.** $8 \times 0.5 = 4$

Number of vote cast for a particular candidate, Time required to run 100 m, Years of schooling, Number of goals in a soccer match, Body temperature, Gravity of stars, Absolute humidity, Atomic Number

5. Find $\sum_{i=1}^{5} c$, where c is a constant. 1

6. What are the functions of statistics? 1

7. What is the origin of the word statistics? 1

8. What is finite population? 1

9. Give an example of an infinite population. 1
10. What is a variable? 1
11. What is a sample? 1
12. What is a qualitative variable? 1
13. Show that $\sum_{i=1}^{4}\sum_{j=1}^{4} x_i \times y_i = \neq \sum_{i=1}^{4} x_i \times y_i$
14. What is a quantitative variable? 1
15. What is a discrete variable? 1
16. What is a continuous variable? 1
17. Differentiate between discrete and continuous variable. 2
18. Is the weight of a person a discrete or continuous variable? 1
19. Is the number of children in a family a discrete or continuous variable? 1
20. Is the height of a building a discrete or continuous variable? 1
21. Is the time it takes for a car to complete a race a discrete or continuous variable? 1
22. Is the number of languages spoken by a person a discrete or continuous variable? 1
23. Is the distance between two cities a discrete or continuous variable? 1
24. Is the number of goals scored by a football team in a match a discrete or continuous variable? 1
25. Is the volume of water in a swimming pool a discrete or continuous variable? 1
26. Is the type of blood group (A, B, AB, O) a qualitative or quantitative variable? 1
27. Is the number of books in a library a qualitative or quantitative variable? 1
28. Is the brand of a smartphone (e.g., Apple, Samsung, etc.) a qualitative or quantitative variable? 1
29. Is the temperature measured in Celsius a qualitative or quantitative variable? 1
30. Is the color of a car a qualitative or quantitative variable? 1
31. What is univariate data? 1
32. What is bivariate data? 1
33. Differentiate between qualitative and quantitative variable. 2
34. What are the scales of measurement? 2
35. Differentiate between ratio and interval scale. 2
36. What is nominal scales of measurement? 1
37. What is ordinal data? 1
38. What is ratio data? 1
39. What is interval data? 1
40. Give one example of each scale of measurement. 2
41. Distinguish qualitative and quantitative variables. 2
42. Explain change of origin and scale with an example. 2
43. Differentiate between ratio and interval scale of measurement. 2
44. What is change of origin? 1
45. What is change of scale? 1
46. What is another way to write $\sum_{i=1}^{n} bx_i$? 1

47. After expansion, what does $\sum_{i=1}^{n}(ax_i - b)$ become?

48. What is the expansion of $\left(\sum_{i=1}^{n} x_i\right)^2$

49. If the scores of five students in a test are 78, 85, 92, 88, 95, find $\sum_{i=1}^{5}(x_i^2 - 2x_i + 3)$ 2

50. If the ages of a group of people are 22, 25, 28, 30, and 35, find $\sum_{i=1}^{5}(x_i^2 + 3x_i)$ 2

51. If the heights of five individuals are 150, 155, 160, 165, 170 cm, find $\sum_{i=1}^{5}(x_i^2 - 4x_i + 6)$ 2

52. If the monthly salaries of five employees are 2000, 2500, 3000, 3500, and 4000 dollars, find $\sum_{i=1}^{5}(x_i^2 + 2x_i - 1)$ 2

Chapter 2

Collection, Presentation, and Organization of Data

2.1 Creative Questions

1. **A teacher asked 30 students about their preferred mode of study. The responses were categorized into five types. The collected data are as follows:**

   Online Offline Mixed Offline Group Online Mixed Group Offline Mixed
   Online Group Offline Mixed Online Offline Online Group Mixed Online
   Group Online Mixed Offline Online Group Offline Mixed Online Group

   (a) Represent the above data using a Bar Diagram and explain your findings. 3

   (b) Would a Pie Chart be equally appropriate here? Discuss with reasoning. 4

2. **Given below are the names of the districts encountering road accidents during the first five days in January, 2025.**

Day	Districts
1	Narayanganj, Faridpur, Naogaon, Pabna, Satkhira, Mymensingh, Dinajpur, Rajshahi
2	Chuadanga, Khagrachhari, Cox's Bazar, Barishal, Chattogram,Munshiganj
3	Munshiganj, Narayanganj, Bogura, Gaibandha, Sunamganj, Chattogram
4	Dhaka, Patuakhali, Noakhali
5	Natore, Munshiganj, Mymensingh

   (a) On which day do you see the highest number of accidents? 1

   (b) Make a table of top 5 districts with highest number of accidents. 2

   (c) Make a Pie Chart using the table in (b). 3

   (d) Make a bar diagram showing the number of the districts affected daily, and interpret the trend. 4

3. **Some of the most-searched key words on the web by the employees of a social media research firm are:**

   Facebook, Google, YouTube, Facebook, Facebook, YouTube, Weather,
   YouTube, Gmail, YouTube, Facebook, YouTube, Google, YouTube,
   Gmail, Weather, Gmail, Gmail, YouTube, Facebook, YouTube,
   Gmail, Facebook, YouTube, Facebook

   (a) What is most popular key word? 1

   (b) Differentiate between primary and secondary data. 2

   (c) Draw a Bar Chart from the above data and explain. 3

   (d) Is a Pie Chart a better representation of this data? Justify. 4

4. **Favorite colors of 30 individuals are noted down. There are five different colors. The recorded colors are given below:**

   Brown Red Pink Green Green Green Brown Pink Brown Red
   Brown Red Green Pink White Red Brown Green White Brown
   White Brown Pink Red White Brown Green Red Pink Red

(a) What is nominal data? 1

(b) What are the ways to deal with categorical data? 2

(c) Draw a Pie Chart from the above data and explain. 3

(d) Is Bar Diagram a better representation of this data? Justify. 4

5. **The favorite fruits recorded from 30 individuals are as follows:**

Apple, Banana, Cherry, Banana, Mango, Apple, Cherry, Grape, Banana, Mango,
Cherry, Apple, Grape, Banana, Mango, Apple, Grape, Cherry, Apple, Banana,
Cherry, Mango, Apple, Banana, Grape, Cherry, Mango, Apple, Banana, Grape

(a) What kind of data is this? 1

(b) How can you present the data in this dataset? List three methods. 2

(c) Draw a Pie Chart from the above data and explain. 3

(d) Is Bar Diagram a better representation of this data? Justify. 4

6. **The following table tracks the number of individuals who sleep within specific hourly intervals.**

Hours of Sleep (per night)	4-5	5-6	6-7	7-8	8-9	9-10	10+
Number of Individuals	12	20	25	30	18	8	7

(a) Draw an Ogive from the data provided and explain. 3

(b) Write five useful insights about the data combining information from the Ogive and the table. 4

7. **Hourly wages of 100 workers in an idustry were collected by a market analyst. The analyst desires to mine a pattern and useful insights from the collected data about the industry. The obtained data are demonstrated below:**

Wage	51-55	56-60	61-65	66-70	71-75	76-80	81-85
Number of workers	7	11	18	36	15	8	5

(a) What is class interval? 1

(b) How does a frequency distribution help us to find pattern in data? 2

(c) Draw an Ogive from the data provided and explain. 3

(d) Write five useful insights about the data combining information from Ogive and the table. 4

8. **The heights of 20 adult males were collected and found to be as follows:**

170, 172, 168, 175, 180, 178, 169, 173, 176, 175
171, 174, 177, 165, 179, 167, 181, 164, 182, 172,

(a) What are discrete class intervals? 1

(b) What are the problems with discrete intervals? 2

(c) Create a Histogram from the data and explain. 3

(d) Create an Ogive and compare the focus of these two graphs. 4

9. **The test scores (out of 100) of 20 students in a mathematics exam were recorded as follows:**

65, 72, 78, 81, 55, 88, 92, 47, 69, 73
84, 91, 77, 63, 58, 71, 94, 86, 68, 79

(a) What are discrete class intervals? 1

(b) What are the limitations of using discrete class intervals with continuous data? 2

(c) Make a frequency distribution and explain. 3

(d) Construct a histogram for the given data and interpret its shape. 4

10. **A school health survey was conducted to understand the weight distribution among students. As part of this initiative, the weights (in kilograms) of 20 randomly selected students from different grades were measured and recorded to assess the general health status and identify patterns or anomalies in weight across the group. The collected data are as follows:**

48, 52, 55, 50, 60, 62, 53, 58, 51, 54
56, 59, 49, 47, 61, 63, 57, 46, 45, 50

(a) Construct a frequency distribution table using a suitable class interval. 3

(b) Create an Ogive and hence estimate the quartiles and interpret. 4

11. **The ages of 20 participants in a fitness program were recorded and found to be as follows:**

25, 30, 28, 35, 40, 38, 26, 32, 36, 31
27, 33, 29, 41, 42, 37, 34, 39, 43, 45

(a) Create a frequency distribution and interpret. 3

(b) Create a Histogram from the data and explain. If the no. of classes were fewer, how would the pattern of the distribution shift? 4

12. **The marks of 57 students in a village are given below.**

Class	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency	5	8	10	13	11	7	3

(a) Are the intervals discrete or continuous? 1

(b) Question 2

(c) Create an Ogive and interpret. 3

(d) Question 4

2.2 Short Questions

1. How is a bar diagram different from a pie chart? 1
2. If two class intervals of a frequency distribution are (10-30) and (30-50), what is the width of class interval? 1
3. If two class intervals of a frequency distribution are (20-30) and (30-40), what is the width of class interval? 1
4. What is an open-ended distribution? 1
5. Differentiate between primary and secondary data. 2
6. What are discrete class intervals? 1
7. What are the problems with discrete intervals 2
8. What are the two main types of variables? 1
9. How is a bar diagram different from a pie chart? 1
10. For a pie chart, how are the angles calculated? 1
11. What does a frequency distribution show? 1
12. How do you define a categorical variable? 1
13. What is the primary use of a histogram? 1
14. Explain the significance of the height of bars in a bar diagram. 1
15. What information can you gather from a pie chart? 1
16. How is data grouped in a frequency distribution? 1
17. What does the width of the bins represent in a histogram? 1
18. Why is it important to label the axes in a bar diagram? 1
19. Define nominal data. 1
20. Relate histogram and stem and leaf plot 2
21. What is the purpose of a frequency distribution? 1
22. Explain the difference between discrete and continuous data. 2
23. What is a histogram used for? 1

Chapter 3

Measures of Central Tendency

3.1 Creative Questions

1. **Three cars in a race cover different distances with different velocities as shown in the table below:**

Car	Distance (m)	Velocity (m/s)
A	20	5
B	28	7
C	36	9

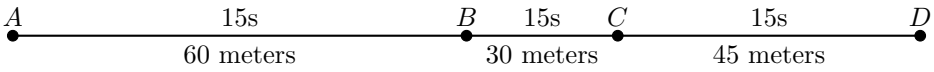
   (a) Question 1

   (b) Question 2

   (c) Question 3

   (d) Question 4

2. **A vehicle moves along the path ABCD, covering each segment at a uniform velocity.**



   (a) What is the formula of harmonic mean? 1

   (b) When should we use the harmonic mean? 2

   (c) Find the average velocity of the vehicle. 3

   (d) Compare the average speed estimated using the arithmetic and harmonic mean and verify their suitability in the given situation. 4

3. **While driving from city A to city B, a car got 22 miles per gallon and while returning on the same road, the car got 30 miles per gallon.** REVIEW THE QUESTION

   (a) How many measures of centrtal tendencies are there? 1

   (b) Which route takes more amount of fuel per mile? 2

   (c) If the total distance for the round trip was 300 miles, how many gallons of fuel were used for the entire journey? [1] 3

   (d) Find the car's average gas mileage for the entire trip, in miles per gallon. 4

4. **A shrimp producer wanted to get an insight into his shrimp production. To do so, he randomly collected weights of different shrimps in his farm.**

Weight of shrimp (gm)	10-20	20-30	30-40	40-50	50-60
Frequency	5	8	10	9	4

   (a) What is the primary goal of central tendency? 1

   (b) When is Median a better measure of central tendency than Arithmetic Mean and why? 2

[1] $$\frac{150}{22} + \frac{150}{30} \text{ gallons} = \frac{130}{11} \text{ gallons} = 11.82 \text{ gallons}$$

(c) From the stem, find 3rd quartile and explain. 3

(d) Find harmonic mean (HM) and compare with the arithmetic mean (AM) 4

5. **An arithmetic series is formed as follows:**

$$a, a+c, a+2c, \cdots, a+2nc$$

(a) What is change of origin and scale? 1

(b) Convert the series into a set of natural numbers. 2

(c) Find the arithmetic mean of the series with the use of change of origin and scale. 3

(d) Find the geometric mean of the series: $1, 2, 4, \cdots, 2^n$. 4

6. **Scores by Travis Head in the last two matches of ICC Men's Cricekt World - 2023 are given. In Cricket, Strike Rate (SR) is computed by dividing Runs by Balls and then multiplying the quotient by 100.**

Match	Runs	Balls
1	62	48
2	137	120

(a) How many averages do you know of? 1

(b) Give an example when arithmetic mean is appropriate instead of harmonic mean. 2

(c) When is Weighted Harmonic mean is used. Show a numerical example. 3

(d) Determine the average Strike Rate of the batter 4

7. **Average height of the four tallest towers in Dhaka is 153.25 meters. The heights of first three towers is 171, 153 and 152 meters, respectively. A new tower has been built with height 150 meters.**

(a) Write two primary uses of central tendency. 1

(b) Prove mathematically: $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$ 2

(c) Compute the height of the forth tower. 3

(d) After the addition of the new tower, will the average increase or decrease? Explain logically and empirically (using data). 4

8. **For two non-zero positive numbers, $GM = 4\sqrt{3}$ and $HM = 6$, where the quantities bear usual notations**

(a) When is Harmonic mean suitable? 1

(b) For two numbers, what is the relationship between AM, GM, and HM? 2

(c) What is the Arithmetic mean? 3

(d) Determine the numbers. 4

9. **For two positive non-zero numbers, if $GM = 5\sqrt{2}$ and $AM = 8$, where the symbols represent their usual meanings:**

(a) Find the Harmonic mean (HM). 3

(b) Identify the two numbers. 4

10. **For two positive non-zero numbers, if $GM = 6\sqrt{3}$ and $AM = 10$, where the symbols represent their usual meanings:**

(a) Determine the Harmonic Mean (HM) of the two numbers. 3

(b) Find the values of the two numbers. Comment on the obtained values. 4

11. **12 is deducted from each value of a variable and then divided by 3. The new arithmetic mean (AM) is found to be 4.**

(a) What is change of origin? 1

(b) Does AM depend on origin? Prove with an example. 2

(c) From the stem, find the original AM. 3

(d) Does the origin or the scale have greater impact on AM in this example? 4

12. **A statistics teacher gave a test to students worth 50 marks. After publishing the result, it was decided the marks would be converted to 100-marks scale. It was also noted that each student should get 2 marks more than the given marks.**

(a) In the stem, which concept is revealed? 1

(b) Does the arithmetic mean depend upon change of scale? 2

(c) Demonstrate the effect of change of origin and scale on arithmetic mean, deriving the formula for conversion between them. 3

(d) If the original arithmetic mean of the marks of the students is 44, what is the changed mean? Does the change of origin and scale always increase the value? Analyze. 4

13. **The arithmetic and geometric means of ages two boys Abir and Abid are 10 and 8.**

(a) What is arithmetic mean? 1

(b) When can we not calculate arithmetic mean? 2

(c) Determine the ages of Abir and Abid. 3

(d) Does the data comply with the theorem $AM \times = GM^2$? 4

14. **Income of 100 individuals in the city of Rajshahi were analyzed and found to produce the the following distribution:**

Income	40-50	50-60	60-70	70-80	80-90
Number of Individuals	15	20	35	20	10

(a) What is Median? 1

(b) Does median necessarily lie in the dataset? 2

(c) Estimate Median and explain the result. 3

(d) Find Arithmetic Mean and Mode. Which measure seems to be the best one? 4

15. **Amount of rainfall in some cities around the world for a month were obtained as follows:**

Rainfall (mm)	**Frequency**
20-30	5
30-40	6
40-50	4
50-60	3
60-70	5

(a) When is Short-Cut method for Arithmetic Mean useful? 1

(b) Derive the formula of Short-cut Method 2

(c) Compute the Arithmetic Mean using the Short-cut method. 3

(d) Compute the Arithmetic Mean with a different value of origin (a). Do both the methods give same result? 4

16. **The number of hours spent studying per week by students in a school were recorded as follows:**

Hours Studied	**Frequency**
0-5	8
5-10	12
10-15	10
15-20	6
20-25	4

(a) Relate short-cut method of arithmetic mean with change of origin and scale. 2

(b) Compute the Arithmetic Mean of the given data using the short-cut method. 3

(c) Compute the Arithmetic Mean with a different value of origin (a). Do both the methods give same result? What is the best choice of a? 4

Concentration (mol/L)	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6
Frequency	4	7	5	6	3

17. **Concentrations of a chemical solution (in mol/L) were recorded over several trials as follows:**

    (a) Compute the Arithmetic Mean using the Short-cut method. 3

    (b) Compute the Arithmetic Mean using a different assumed mean (A). Do both methods yield the same result? 4

18. **A sports analyst collected ages of athelets having ages between 10 and 35. He then presented his findings as below:**

Age	10-15	15-20	20-25	25-30	30-35
No. of Athlete	2	8	10	5	3

    (a) What is central tendency? 1

    (b) When is geometric mean iappropriate to measure? 2

    (c) Compute median from the stem. 3

    (d) Show that Arithmetic mean is greater than Harmonic mean. Which one of them is more suitable for this data? 4

19. **The monthly temperatures in the cities of Dhaka (in Degree Celsius) and New York (in degree Fahrenheit) are given in the following table:**

City	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Dhaka (^oC)	24.6	28.3	32.4	33.9	33.6	32.8	32.1	32.3	32.4	31.9	29.7	26.1
New York (^oF)	29	32	40	50	60	70	75	74	67	56	46	35

    (a) How is the Range defined? 1

    (b) Can the standard deviation exceed the variance? Prove. 2

    (c) Find the standard deviation of the above data and compare the cities. 3

    (d) Examine the efficacy the standard deviation in light of the given context. Is a better measure availabe? Analyze. 4

20. **Mean monthly salaries of employees of two companies A & B are tk. 65,000 and tk. 75,000. The combined arithmetic mean (AM) is tk. 71,000 and number of employees in the company A is 20.**

    (a) Write down the formula of combined AM for k groups. 1

    (b) What is the combined AM of two data sets with AM 35 and 45 and number of values equal? 2

    (c) How many employees are there in the company B? 3

    (d) Salary of an employee of company A was recorded as tk. 60,000 in place of 65,000. What is the new AM of company A. Also find the corrected combined AM. 4

21. **Average marks of two sections A & B in a statistics exam are 68 and 74 respectively. The overall average mark of both sections combined is 70. Section A has 25 students.**

    (a) How many students are there in section B? 3

    (b) Later, it was found that a student in section B was wrongly marked 80 instead of 90. Find the corrected average of section B and the new combined average. 4

22. **A departmental store records their sales. An analysis of products with prices less than tk. 30 generates the following table.**

Price	0-5	5-10	10-15	15-20	20-25	25-30
Frequency	1	0	2	3	8	12

    (a) What is relative frequency? 1

    (b) If $Y = a + bX$, $\bar{Y} =$? 2

    (c) Find 67th Percentile and 3rd Quartile and explain. 3

    (d) Is AM or Median more suitable for this data? Elucidate. 4

23. **Arithmetic (AM) and Harmonic Mean (HM) of two numbers are 25 and 9, respectively.**

(a) When is HM useful? 1

(b) Derive HM formula using the concept of average velocity. 2

(c) Find the two values from the stem. 3

(d) Show mathematically that $HM \leq AM$ (for n=2) 4

24. **Frequency distribution of marks in statistics of a college is given in the following table.**

Marks	Number of Students Group - A	Number of Students Group - B
25-30	11	10
30-35	18	16
35-40	21	22
40-45	26	28
45-50	14	9

(a) What is data? 1

(b) What are the disadvantages of secondary data? 2

(c) Calculate the arithmetic mean of Group - A 3

(d) Compute the combined mean. Is it greather than the arithmetic mean of Group - B? Explain the possible reason(s). 4

25. **The following table presents the distribution of monthly salaries (in thousand BDT) of employees in two different departments of a company.**

Salary Range (in 1000 BDT)	Number of Employees Department - X	Number of Employees Department - Y
20-25	8	6
25-30	14	12
30-35	19	21
35-40	24	26
40-45	15	10

(a) Determine the arithmetic mean salary of employees in Department - X. 3

(b) Compute the combined mean salary. Is it higher than the arithmetic mean of Department - Y? Justify your answer with a statistical explanation. 4

26. **In the test examination, marks of 11 students in statistics are: 90, 92, 93, 49, 44, 88, 80, 58, 83, 71, 76.**

(a) What is central tendency? 1

(b) When is median better than arithmetic mean? Explain with an example. 2

(c) Find the 3rd the quartile and 61^{st} percentile from the data and explain. 3

(d) Do quantiles depend on change of origin and scale. Prove using two examples. 4

27. **HSC exam result of two colleges in a village are given below:**

College	Students	Passed
1	10	8
2	20	5

(a) How many measures of central tendency are there? 1

(b) Which measure of central tendency is perfect? 2

(c) Find the pass rate of the colleges and, using them, of the village. 3

(d) Show an alternative method which gives the same result. 4

28. **Scores of a batsman in the last 20 innings are**

28, 30, 16, 48, 50, 86, 105, 20, 10, 36,
12, 25, 20, 35, 65, 12, 10, 76, 55, 32

(a) Write down the formula of weighted harmonic mean 1

(b) Can median be a better measure of central tendency than arithmetic mean for this data? 2

(c) Draw a stem and leaf plot from the data and explain. 3

(d) Make a frequency distribution from the data and also find and interpret cumulative frequencies and percentages. 4

29. **In ODI cricket, two top batsmen are (as of 2nd Sept, 2022) Babar Azam and Rassie van der Dussen. Their average (arithmetic mean) scores are 59.79 and 69.32, appearing in 90 (including being not out in 12 occassions) and 33 (including being not out in 11 occassions) matches, respectively.**

(a) When is arithmetic mean inappropriate to use? 1

(b) Is arithmetic mean always suitable for comparison? 2

(c) Find the combined arithmetic mean and explain. 3

(d) How to compare two sets of data having significantly distinct ranges? 4

30. **A fridge manufacturing company observe temperatures of newly developed 8 deep fridges. The observed temperatures (in degree celsius are:**

-10, -8, -2, -4, -4, -1, -12, -3, -13

(a) What is a Decile? 1

(b) How many Deciles does a data set have? Why? 2

(c) Compute the 8th Decile from the data and explain. 3

(d) What is the value below which 75% observations lie? Find in two distinct ways. 4

31. **Given below is a series of data.**

$$5, 7, 9, \cdots, 123$$

(a) What is the summation of natural numbers up to nth value? 1

(b) Find the arithmetic mean of natural numbers from 1 up to 20. 2

(c) Find the arithmetic mean of the given series. 3

(d) Prove that arithmetic mean is greater than gemetric mean theoretically and empricially. 4

32. **Grades of a an undergraduate student with major in statistics are given below:**
[Credits serve as weights]

Course	Probability	Simulation	Calculus	Linear Algebra	Econometrics	Programming
Grade	3.75	3.50	3.50	3.75	3.00	3.50
Credit	4	3	4	4	2	3

(a) Write down the formula of weighted mean. 1

(b) What is difference between weight and frequency? 2

(c) Determine the GPA of the student. 3

(d) Determine the geometric mean for the data and evaluate suitability. 4

33. **A student walks 3 hours at 5 km per hour (kph), 4 hours at 4 kph, and 2 hours at 3 kph.**

(a) When is harmonic mean suitable? 1

(b) Which mean could we use for the given data and why? 2

(c) Find the average speed using weighted harmonic mean. 3

(d) Find the correct and suitable average speed using another method and mathematically show they are equivalent. 4

34. **Scores of four athletes in different events at a track meet are recorded below:**

A coach believes that events with higher difficulty factors should contribute more to the overall ranking and suggested a new weighting where the weight for each event is the square of its difficulty factor.

Event	High Jump	Long Jump	Shot Put	Javelin Throw
Score	8.5	7.2	12.8	55.5
Difficulty Factor	2	3	2.5	1.5

(a) Write down the formula of weighted mean. 1

(b) What is difference between weight and frequency? 2

(c) Calculate the weighted average score of the athletes across all four events. 3

(d) If the coach's suggestion is implemented, would the mean be shifted upward or downward? Show mathematically and empricially. 4

35. **A biologist observes three different species of fish swimming at different speeds. The recorded swimming speeds are 2 m/s for 4 hours, 3 m/s for 3 hours, and 5 m/s for 2 hours.**

(a) Find the average swimming speed using the weighted harmonic mean. 3

(b) Confirm the average speed by applying another method and prove their mathematical equivalence. 4

36. **A biologist records the lengths (in cm) of a certain species of fish found in different water bodies as shown below:**

Category	5-10	10-15	15-20	20-25	25-30
Frequency	8	12	15	10	5

(a) Compute the Arithmetic Mean of the fish lengths. 3

(b) Suppose there is another set of data from a similar study with a mean of 17 cm and 20 observations. Find the combined mean of the two datasets and compare it with the mean found in (i). 4

37. **A meteorologist records the monthly rainfall (in mm) in different regions over a year as shown below:**

Rainfall (mm)	0-50	50-100	100-150	150-200	200-250
Frequency	6	10	14	8	7

(a) Which class contains the Mode? 1

(b) Find Δ_1 and Δ_2, where the symbols represent their usual meanings 2

(c) Find the Mode using the Direct formula. 3

(d) Find the Mode using histogram and compare with direct method. Which one do you think is more accurate? 4

38.

39. **An ecologist records the number of trees in different forest plots based on their heights (in meters) over a survey period as shown below:**

(a) How much data lie below the Median point? 1

(b) What is the relationship between the median and the quartiles? 2

(c) Find the median using the direct formula. 3

(d) Find the 1st and 3rd quartile from the table and interpret. 4

ADD QUESTIONS ASKING MODE AND MEDIAN AND OTHER QUARTILES FROM GRAPH

40. **A botanist measures the heights (in cm) of plants from a sample as shown below:**

(a) Find the median height of the plants and interpret. 3

(b) Determine the first (Q1) and third (Q3) quartiles of the plant heights. Explain the significance of these quartiles in understanding the data distribution. 4

41. **A passer-by walks 3 hours at 5 km per hour (kph), another 3 hours at 4 kph, and another 3 hours at 3 kph.**

(a) When is harmonic mean suitable? 1

(b) Which mean could we use for the given data and why? 2

(c) Find the average speed of the passer-by usingt he proper method. 3

Tree Height (m)	0-5	5-10	10-15	15-20	20-25
Frequency	5	12	18	9	6

Height (cm)	20-30	30-40	40-50	50-60	60-70
Frequency	6	10	12	8	4

(d) Find the correct and suitable average speed using another method and mathematically show they are equivalent. 4

42. **A cyclist moves around a square-shaped lake with the speeds 20, 25, 30, and 16 km per hour.**

(a) What is grouped data? 1

(b) Is arithmetic mean suitable for this data? 2

(c) Find the average speed of the cyclist. 3

(d) Can we use some other formula for finding the same value of the average? Demonstrate. 4

43. **There are only two students in a class IX in a college. In half-yearly exam, the arithmetic and the geometric mean of the marks of those two studnets are 25 and 15, respectively.**

(a) Write down the formula of Geometric Mean for grouped data. 1

(b) Prove with an example: $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$ 2

(c) Determine the marks of the students. 3

(d) Is 20 a possible value of the harmonic mean of this data? Explain theoretically and empirically. 4

Rainfall (mm)	20-30	30-40	40-50	50-60	60-70
Frequency	5	6	4	3	5

3.2 Short Questions

3.2.1 General Questions

1. What is the primary purpose of central tendency? 1
2. When is WAM equal to WHM? Show mathematically. 2
3. When is the equality true: $AM = GM = HM$? Prove mathematically and empirically. 3
4. When is Median a better measure of central tendency than Arithmetic Mean? 1
5. When is Harmonic Mean more suitable than Arithmetic Mean? 1
6. Write two primary uses of central tendency. 1
7. For two distinct non-zero values, what is the relationship among AM, GM, and HM? 1
8. what is the relationship among AM, GM, and HM? 1
9. What are the criteria of a good measure of central tendency? 1
10. For two non-zero positive numbers, prove $AM \geq GM \geq HM$ 3
11. Find the AM, GM, and HM: $1, 2, 4, \cdots, 2^n$ 4
12. A data set has the following values: 5, 10, 20, 30. Find AM, GM, HM, Median, and Mode and order the results in ascending order. 4

3.2.2 Arithmetic Mean

13. Does Arithmetic Mean depend on change of origin and scale? Prove mathematically and with an example. 3

14. If $\bar{X} = 3,$ and $n = 10$, what is $\sum X_i$? 2

15. Derive the formula of combined mean using logic and an example. 2

16. Prove with an example: $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$ 2

17. Prove mathematically: $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$ 2

18. Derive the formula of Arithmetic Mean for short-cut method. 2

19. Derive the formula of combined Arithmetic Mean for n nnumber of observations.

20. Find Arithmetic Mean of first n natural numbers. 2

21. If $u_i = x_i + y_i$, find $\bar{x}$ in terms of u. 1

22. For two numbers, $AM = 25$ and $GM = 15$. $HM =$? 1

23. $\bar{X} = 25$ and $Y_i = 5X_i + 20$. $\bar{Y} =$?

24. Find Arithmetic Mean: $11, 13, 15, \cdots, 57$

25. Find Arithmetic Mean: $115, 120, 125, \cdots, 225$

26. Find Arithmetic Mean: $14, 18, 22, \cdots, 70$ 2

27. Find Arithmetic Mean: $0, 5, 10, \cdots, 70$ 2

28. Arithmetic Mean $(\bar{X})$ of five numbers is 40, and of three of them is 30. What is $\bar{X}$ of the rest two?

29. AM of 200 values is found to be 50. Later it was seen two values were recorded as 92 and 8 in place of 192 and 88, respectively. What is the correct AM?

30. AM of 8 values of 20. If the 9th value is 0, what is the new AM?

31. AM of Income of City A is 1500 and of City B is 1200. If a person moves from city A to city B, can AM of both cities decrease?

32. Calculate Arithmetic Mean:

Renumeration (X)	**No. of Employees (f)**
30	5
35	8
40	10

33. Calculate Arithmetic Mean using short-cut method:

Rainfall (mm)	**Frequency**
20-30	5
30-40	6
40-50	4
50-60	3
60-70	5

34. What is the relationship between changing origin & scale and short-cut method of Arithmetic Mean?

3.2.3 Geometric Mean

35. Write down the formula of Geometric Mean for grouped data.
36. Find Geometric Mean for these values: $2, 4, 8$ 1
37. Derive the formula of Geometric Mean using logarithm. 1
38. When is Geometric Mean not calculable? 1
39. When is Geometric Mean appropriate? 1
40. Determine the formula of combined Geometric Mean when $n_1 = n_2 = n$ 1
41. $Y_i = 3X_i$. If $G_y = 9, G_x =$? [G stands for Geometric Mean] 2
42. $n_1 = 15, G_1 = 75, n_2 = 10, and G_2 = 80$ Find combined GM.

3.2.4 Harmonic Mean

43. Write down the formula of Weighted Harmonic Mean. 1
44. When is Weighted Harmonic Mean used instead of Unweighted one? 1
45. Calculate Harmonic Mean: $10, 15, 20$ 1
46. Show mathematically that Harmonic Mean of velocity for a fixed distance is equal to average velocity.
47. Find the average speed:

Path	Distance (km)	Speed (km/h)
Path 1	3	8
Path 2	2	9
Path 3	2	2

3.2.5 Median

48. Does Median depend on origin and scale? Prove with an example?
49. How much data lie below the Median point? 1
50. What is the relationship between the median and the quartiles? 2
51. Median score of 50 students is 70. What does it mean? 1
52. Write down the formula of Median for even number of observations.
53. Write down the formula of Median for grouped data.
54. Find median: $5, 10, 15, 20, 25, 30$
55. Find median: $7, 2, 4, 5, 6, 10$ and explain the value 2
56. Is Median affected by outliers?
57. Does Median depend on origin and scale? Prove. 2
58. What is the greatest disadvantage of Median?
59. Does Median always lie in the data set from which it is calculated? 1

3.2.6 Mode

60. What is the formula of Mode for grouped data?
61. Find the mode: $2, 2, 3, 3, 4, 4, 2, 2, 7$
62. What is an Unimodal dataset?

3.2.7 Quadratic Mean

63. What is the formula of Quadratic Mean for grouped and ungrouped data?
64. When is Quadratic Mean used?

3.2.8 Partition Values

65. If a data is divided into four parts, how many partition values are created?
66. Write down the formula of Median for even and odd number of observations.
67. Derive the general formula of Quartiles using the concept of Median?
68. Third Quartile of score of 50 students is 76. What does it mean?
69. Which Quartile is equal to Median?
70. Which Percentile is equal to 3rd Quartile?
71. Find all the Quartiles and interpret: $2, 1, 0, 5, -6, 7, -4$ 3

Chapter 4

Measures of Dispersion

4.1 Creative Questions

1. **Rainfall measurements (in mm) of two cities for six days are as below:**

Day	1	2	3	4	5	6
City X	15	12	10	18	14	11
City Y	20	8	25	5	22	10

   (a) What is a measure of dispersion? 1

   (b) Explain with an example why measuring dispersion is important in data analysis. 2

   (c) Calculate the variance of rainfall for City X. 3

   (d) Which city shows more consistent rainfall? Justify. 4

2. **Temperatures of two cold regions for five days are as below:**

City A: 2, 1, -1, 0, 3

City B: 3, 0, -2, 2, 3

   (a) What is dispersion? 1

   (b) Illustrate the necessity of dispersion with an example. 2

   (c) Find the variance of temperature of city A. 3

   (d) Which city has more consistent temperature? Analyze. 4

3. **Temperatures of two cold regions for five days are as below:**

City A: 2, 1, -1, 0, 3

City B: 3, 0, -2, 2, 3

   (a) What is standard deviation? 1

   (b) Is standard deviation of a set of negative values negative? Justify mathematically. 2

   (c) Find Mean Deviation about mean of the values of city A. 3

   (d) Which city has more consistent weather? Verify statistically. 4

4. **Rainfall (in mm) recorded in two towns over five days is as follows:**

Day	1	2	3	4	5
Town X	12	15	10	18	14
Town Y	20	22	18	25	21

   (a) Find the Mean Deviation about the median for the rainfall data of Town X. 3

   (b) Which town shows greater variability in rainfall? Support your answer with statistical measures. 4

5. **Scores of two students in five tests are as follows:**

   (a) Find the Mean Deviation about the mean for the scores of Student A. 3

Test	1	2	3	4	5
Student A	78	85	90	88	76
Student B	80	82	78	86	84

Month	1	2	3	4	5
Store P	50	55	48	52	49
Store Q	60	58	65	63	61

(b) Which student has more consistent performance? Justify using statistical measures. 4

6. **Monthly sales (in thousand dollars) of two stores over five months are given below:**

(a) Calculate the variance of Store P's sales data. 3

(b) Compare the sales stability of both stores using an appropriate statistical measure. 4

7. **Two companies A and B pay their workers on a weekly basis. The summary of wages paid by them is shown below:**

Factory	Mean Wage (BDT)	Standard Deviation	Number of workers
A	1560	90	200
B	1580	70	160

(a) What is dispersion? 1

(b) Is variance always greater than stanard deviation? Justify. 2

(c) Which company is more consistent with their wages? 3

(d) Compare the wages with variance alone. Does it produce the same conclusion. Justify the finding. 4

8. **Two schools, X and Y, conducted the same mathematics examination for their students. The summary of the test scores (out of 100) is shown below:**

(a) Determine the combined standard deviation of the scores of the schools. 3

(b) Which school has more consistent scores? Analyze. 4

9. **Mean and Standard Deviation of 200 items are found to be 60 and 20. Later it was found that two items were recorded as 3 and 67 in place of 13 and 17.**

(a) Does standard deviation depend on change of origin? 1

(b) Prove $\sigma^2 = \frac{\sum x^2}{n} - (\frac{\sum x}{n})^2$ from original formula. 2

(c) Should the correct mean be smaller or greater? Also find it and compare. 3

(d) Find the correct standard deviation. 4

10. **The monthly temperatures in the cities of Dhaka (in Degree Celsius) and New York (in degree Fahrenheit) are given in the following table:**

City	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Dhaka (^oC)	24	28	32	33	33	32	32	32	32	32	29	26
New York (^oF)	29	32	40	50	60	70	75	74	67	56	46	35

(a) How is the Range defined? 1

(b) Can the standard deviation exceed the variance? Prove. 2

(c) Find the standard deviation of the above data and compare the cities. 3

(d) Examine the efficacy the standard deviation in light of the given context. Is a better measure availabe? Analyze. 4

11. **A company recorded the monthly sales (in lakhs of) of two salespersons X and Y for the first half of the year:**

(a) Calculate the **mean deviation about median** for salesperson X. 3

(b) Which salesperson has **more consistent sales**? Justify using an appropriate relative measure of dispersion. 4

School	Mean Score	Standard Deviation	Number of Students
X	72.5	8.2	150
Y	74.0	6.5	120

Month	Jan	Feb	Mar	Apr	May	Jun
Salesperson X	42	38	45	40	43	41
Salesperson Y	50	28	60	35	55	30

12. **The table below presents the number of goals scored by two footballers, A and B, over five consecutive seasons**

    (a) Find Mean Deviation about mean and median of the footballer A. 3

    (b) Which footballer should be hired a club? Determine with the help of a suitable relative measure of dispersion. 4

4.2 Short Questions

1. Which measure of dispersion is suitable for an open-ended distribution? 1
2. What does dispersion measure? 1
3. What are the absolute measures of dispersion?
4. What are the relative measures of dispersion?
5. What is the formula of Range? 1
6. Does Range consider all values? 1
7. Is Range influenced by extreme values or outliers? 1
8. Write down the formula of Mean Deviation ($MD(\bar{x})$.
9. Write down the formula of Mean Deviation around median ($MD(Me)$.
10. Write down the formula of Mean Deviation for grouped data ($MD(\bar{X})$.
11. Does mean deviation depend on all values of a dataset? 1
12. Does Mean Deviation depend on change of origin and scale? Verify. 2
13. When is variance less than standard deviation? 1
14. Write down the formula of Quartile Deviation. 1
15. Which is the most important measure of dispersion?
16. Is $\sum |x_i - \bar{x}|$ always greater than $\sum(x_i - \bar{x})$? Prove mathematically. 2
17. If $\bar{X} = 25, CV = 50\%, \sigma^2 =$? 1
18. Arithmetic mean of a variable is 16 and variance is 9. What is the value of CV? 1
19. What is the variance of first 5 natural numbers?
20. Find the sum of squares: 12, 10, 11, 15, 20 1
21. Find the square of summation: 2, 5, 7, 8, 4 1
22. Find the sum of squared deviation from the mean of these values: 2, 5, 6, 7, 3 1
23. Find the sum of squares: 8, 6, 9, 12, 14 1
24. Find the square of summation: 3, 7, 5, 9, 6 1
25. Find the sum of squared deviation from the mean of these values: 4, 8, 6, 10, 7 1
26. Compute the sum of squares for the following values: 5, 9, 12, 14, 11 1
27. Compute the square of summation for: 1, 4, 6, 7, 2 1
28. Find the sum of squared deviations from the mean: 3, 6, 8, 10, 5 1
29. Find the sum of squares of differences from 10: 6, 8, 10, 12, 14 1

Season	1	2	3	4	5
Footballer A	10	15	12	9	20
Footballer B	25	5	10	15	6

4.2.1 Variance

30. If $n = 10, \sum x_i = 120, \sum x_i^2 = 2000, CV =$?

31. Is coefficient of variation a pure number? Explain 2

Chapter 5

Moments, SKewness, and Kurtosis

5.1 Creative Questions

1. **The first four moments around 4 of productions of a company over four years are -1.5, 17, -30, and 108.**

   (a) Can the second central moment be negative? 2

   (b) Determine the second and third central moments. 3

   (c) What kind of kurtosis do the data have? 4

2. **Duration of stays of a spy in foreign countries are obtained by a researcher. As part of an analysis, s/he starts with the following summary.**

Duration	1-10	11-20	21-30	31-40	41-50	51-60
Frequency	4	3	3	2	5	2

   (a) What is symmetry? 1

   (b) What is implied by the value of coefficient of skewness 0.8 2

   (c) Estimate the median of the data and interpret. 3

   (d) Obtain coefficient of skewness from data and comment on the life of the spy based on it. 4

3. **The kurtosis and the variance of a symmetrical distribution are $\gamma_2 = -2.3672$ and $\sigma^2 = 4096$, respectively.**

   (a) Find the fourth central moment. 3

   (b) Determine the value of the first and the third raw moments around 5. 4

4. **A human resources department recorded the number of years employees have been with the company. The grouped frequency distribution below summarizes the data for a department.**

Years of Service	0–2	3–5	6–8	9–11	12–14	15–17
Frequency	4	6	10	5	3	2

   (a) Calculate the coefficient of skewness using the grouped data. What does the result indicate about employee tenure? 3

   (b) Compute the coefficient of kurtosis. What does it reveal about the distribution of tenure lengths? 4

5. **A financial analyst is studying the annual returns (in percentage) of a set of investment portfolios. The following table summarizes the data.**

Annual Return (%)	-5 to 0	1-5	6-10	11-15	16-20	21-25
Frequency	3	5	7	6	4	3

   (a) Compute the skewness of the data and interpret the nature of the investment returns. 3

   (b) Determine the kurtosis and explain. 4

6. **There has been an increase in average lifetime of people of Bangladesh. To get more insight on this, a research was conducted, in which ages of retired government employees were recorded. A sample of 10 people is given below:**

75, 62, 63, 72, 66, 76, 59, 77, 70, 79

(a) What is the 2nd central moment equal to? 1

(b) Show that the first central moment is zero. 2

(c) Find the variance of the data. 3

(d) Are the data symmetric? Justify. 4

7. **A study was conducted to track the daily water consumption (in liters) of 10 individuals over a week. The recorded values are as follows:**

2.5, 3.1, 2.8, 3.5, 2.7, 3.3, 2.6, 3.0, 2.9, 3.4

(a) Determine the variance of the data set. 3

(b) Assess whether the data distribution appears to be symmetric with the help two different methods. 4

8. **A researcher analyzed the daily screen time (in hours) of 12 students over a month. The recorded values are as follows:**

4.2, 5.1, 3.8, 6.5, 4.7, 5.3, 4.6

(a) Compute the five-number summary of the data. 3

(b) Analyze the skewness of the dataset using both graphical and numerical approaches. 4

9. **The arithmetic and geometric means of the first and third quartiles of a distribution are 10 and 8, respectively. The second quartile is 10.**

(a) What is the formula suggested by Pearson to find skewness? 1

(b) Which moments are useful in measuring central tendency and dispersion? 2

(c) Find skewness from the stem using a suitable formula. 3

(d) Which method of finding skewness od you think is the best and why? 4

10. **For a particular data set, Median = 120, Mode = 110, Standard Deviation = 4, and Coefficient of Variation (CV) = 3.2**

(a) Why is CV used? 1

(b) Find arithmetic mean.. 2

(c) Find skewness according to Pearson's method (SK_P) 3

(d) Does (SK_P) convey the proper idea about the data as to the given information? Justify. 4

11. **For a given data set representing the monthly salaries (in thousands) of employees in a company, the following statistics are provided: Median = 60, Mode = 55, Standard Deviation = 5, and Coefficient of Variation (CV) = 8.3.**

(a) Calculate the skewness using Pearson's method (SK_P). 3

(b) Does the value of (SK_P) accurately reflect the nature of the data based on the given statistics? Justify your answer. 4

12. **US Dollar exchange (to taka) in Bangladesh since 1980 to 2005 (after each 5 years) were: 16, 31, 36, 40, 52, 64**

(a) What are moments? 1

(b) Which moment is equal to the variance? Show mathematically. 2

(c) Find, from the stem, the first and second raw moments about 1. 3

(d) Find skewness and kurtosis of and explain. 4

13. **A farmer in Dinajpur district produces seasonal crops. First four moments around 9 of his daily earnings are computed as -1, 8, -16, and 25.**

(a) What is the Box and Whisker plot? 1

(b) Can Box and Whisker plot suggest symmetry? 2

(c) Find the arithmetic mean and variance of the farmer's earnings. 3

(d) Do the earnings produce a symmetric data? Analyze. 4

14. **The first four moments about 3 of a distribution are -1, 5, -10, and 120.**

    (a) What are moments used for? 1

    (b) Can the second central moment be greater than the third central moment? 2

    (c) Find the second and third moments about arithmetic mean of the distribution. 3

    (d) Find skewness and kurtosis and comment on the values. 4

15. **The first four moments of a distribution around 5 are 2, 20, 40, and 50, respectively.**

    (a) Draw the shape of a left-skewed distribution. 1

    (b) Derive the value of thew first central moment. [1] 2

    (c) Obtain the first four central moments. 3

    (d) Estimate and comment on the skewness and kurtosis. 4

16. **A data set represents the test scores of 10 students in a recent exam, recorded as follows:**

    56, 62, 68, 71, 65, 59, 74, 67, 70, 63

    (a) Calculate the first four moments [2] about 3. 3

    (b) Compute the variance and kurtosis of the data using converted central moments. Explain what the kurtosis indicates about the distribution. 4

17. **Marks obtained by a student in 7 subjects are**

    70, 66, 55, 45, 80, 30, 82

    (a) What is negative skewness? 1

    (b) Draw graphs of positive and negative skewness showing the locations of mean and median. 2

    (c) Determine the five number summary from the stem and explain. 3

    (d) Are the data symmetric? If not, comment on the pattern of data. 4

18. **United Nations Children's Fund (UNICEF) is an agency of the United Nations responsible for providing humanitarian and developmental aid to children worldwide. A UNICEF researcher collected heights (in feet) of 7 children for a project, and the heights are**

    2.2, 2.15, 1.9, 3.1, 2.7, 3.0, 3.5

    (a) Which value are central moments estimated around? 1

    (b) Moments around origin (0) are central moments - Comment. 2

    (c) Find the first three central moments. 3

    (d) Find the skewness of the data and interpret. 4

19. **A researcher wants to compare average life time of people in Bangladesh and other countries. He collected life time of 10 people in Bangladesh.**

    75, 62, 63, 72, 66, 76, 59, 77, 70, 79

    (a) What is symmetry? 1

    (b) Mathematically show the theoretical value of the first central moment. 2

    (c) Compute the 2nd, 3rd, and 4th central moments of the data. 3

    (d) Estimate skewness and kurtosis and explain. 4

20. **Exam marks of two students were summarized for the purpose of comnparison. The summary is given below:**

    (a) What is kurtosis? 1

    (b) How much data are contained within Interquartile range? 2

    (c) For student A, estimate the Bowley's Coefficient of skewness and explain. 3

[1] $\frac{\sum(x_i-\bar{x}}{n} = \bar{x} - \bar{x}) = 0$

[2] 62.5, 3934.5, 249407.5, 15915069

Measure	Student X	Student Y
First Quartile	28	27
Second Quartile	60	60
Third Quartile	75	73
Minimum	16	14
Maximum	89	86

(d) On the basis of skewness (and hence shape of the data), compare the students. 4

21. **The first four moments around 2 of a dataset were the following:**

-1, 5, 20, 90

(a) What is raw moment? 1

(b) What is the standard deviation of the data in the stem? 2

(c) Determine the third central moment. 3

(d) Comment on the kurtosis of the given data. 4

22. **The heights of the trees of a certain species that were planted at around the same time in a park were examined by an analyst, hired by the park authority to check for any abnormality. He randomly observed 10 trees; the values (in cm) obtained are given below:**

200, 190, 185, 210, 220, 200, 205, 207, 230, 195

(a) What is five number summary? 1

(b) Which measures are shown on a Box & Whisker plot? 2

(c) Display the data on a box plot. 3

(d) Taking a look at the drawn box plot, comment on the symmetry of the data. 4

23. **The weights of newborn babies (in kilograms) in a hospital over a week were recorded by a pediatrician to monitor their health. The weights of 10 randomly selected babies are given below:**

3.2, 2.9, 3.5, 3.1, 3.4, 3.0, 3.3, 3.6, 2.8, 3.7

(a) Represent the data using a Box & Whisker plot. 3

(b) Compare the plot with the five-number summary and comment on the distribution of the data. 4

5.2 Short Questions

1. What are moments? 1
2. How many types of moments are there? 1
3. What is central moment? 1
4. What is raw moment? 1
5. What is the formula of raw moment for grouped data? 1
6. Derive the value of the first central moment. 2
7. Derive the relationship between the first central moment and the origin. 2
8. What is the second central moment equal to? 1
9. Write 3 uses of moments. 2
10. Can moments be negative? Analyze. 2
11. What is skewness? 1
12. What is a symmetric distribution? 1
13. In a symmetric distribution, what is the relationship between Mean, Median, and Mode? 1
14. What is negative skewness? 1

15. What is positive skewness? 1
16. What is the pattern of in a left-skewed didstribution? 2
17. What is the pattern of in a right-skewed didstribution? 2
18. In a left-skewed distribution, what is the relationship between Mean, Median, and Mode? 2
19. In a right-skewed distribution, how [3] are Mean, Median, and Mode related? 2
20. Draw the shape of a left-skewed distribution and show locations of mean, median, and mode. 2
21. Draw the shape of a right-skewed distribution and show locations of mean, median, and mode. 2
22. Write down the formula first moment around 3. 1
23. Draw a symmetrical distribution. 1
24. Draw the shape of a right-skewed distribution. 1
25. What is Pearson's measure of skewness? 1
26. How did Bowley measure skewness? 1
27. What is Kelly's method of coefficient of skewness? 1
28. How are moments used to mesure skewness?
29. What is the relationship between β_1 and γ_1 1
30. What does $\gamma_1 > 0$ imply 1
31. What is kurtosis? 1
32. How many types of kurtosis are there? 1
33. Draw the curve of a Mesokurtic distribution 1
34. Draw the curve of a Platykurtic distribution 1
35. Draw the curve of a Leptokurtic distribution 1
36. Explain the Platykurtic distribution. 2
37. How can you use moments to estimate kurtosis?
38. What is the relationship between β_2 and γ_2
39. What does $\gamma_2 > 0$ imply? 1
40. What is five number summary?
41. Find five number summary: 2, 1, 0, 5, -6, 7, -4 3
42. Write three uses of five number summary 2
43. Which measures are shown on a Box & Whisker plot?
44. How is inner fence found for a Box & Whisker plot
45. How is outer fence found for a Box & Whisker plot
46. What is IQR and how is it calculated? 2
47. Write four uses of the box plot. 2

5.3 Broad Questions

1. Analyze the effect of shift of origin & scale on central moments. 3
2. Prove that β_1 and β_2 do not depend on shift of origin and scale. 4
3. Express the 3rd central moment in terms of raw moments.
4. The first four moments of a distribution around 5 are 2, 20, 40, and 50, respectively. Obtain the first four central moments. Estimate and comment on the skewness and kurtosis.

[3] $Mo < Me < \bar{X}$

Chapter 6

Correlation and Regression

6.1 Creative Questions

1. **A real estate agent wants to model the relationship between the size of a house (in square feet) and its selling price (in lakh). Data for 5 houses are:**

Size (X, sq. ft.)	800	1000	1200	1500	1800
Price (Y, lakh)	30	38	45	55	65

(a) Estimate the regression equation of price on size. 3

(b) **The temperature (in ^{o} Celsius) and rainfall (mm) in the city of Mymensingh on the last 6 rainingd days are summarized as below (temperature is denoted by X and rainfall by Y:**

$$\sum x = 29, \sum y = 38, \sum xy = 205, \sum x^2 = 167, \sum y^2 = 282$$

i. What is covariance? 1

ii. Differentiate between cvariance and correlation. 2

iii. Analyze the linear association between temperature and rainfall. 3

iv. Find the degree of impact of temperature on rainfall. 4

(c) **The following table shows the exam scores of 10 students and the number of hours they studied for the exam. It is hypothesized that the exam score depends on the number of hours studied.**

Hours Studied (X)	2	3	4	5	6	7	8	9	10	11
Exam Score (Y)	65	70	72	78	80	85	88	92	95	98

i. Make a scatter plot using the data and state the potential direction and strength of the relationship between the variables. 3

ii. Find the correlation and interpret. 4

(d) **The following table shows the exam scores of 10 students and the number of hours they studied for the exam. It is hypothesized that the exam score depends on the number of hours studied.**

Hours Studied (X)	2	3	4	5	6	7	8	9	10	11
Exam Score (Y)	65	70	72	78	80	85	88	92	95	98

i. Estimate and interpret the regression coefficient of y on x. 3

ii. Using data, show $r = \sqrt{b_{yx} \times b_{xy}}$; How much variation is explained by the obtained model? 4

(e) **The following table shows the monthly sales (in thousand) of 8 salespeople and their years of experience. It is hypothesized that sales performance depends on experience.**

Experience (X, years)	1	2	3	4	5	6	7	8
Sales (Y, thousand)	25	32	38	45	50	58	63	70

i. Estimate and interpret the regression coefficient of y on x. 3

ii. Estimate the regression coefficient of x on y. Is this model realistic? Explain. 4

(f) **A company wants to study the relationship between its advertising expenditure (in thousand $) and monthly sales (in thousand $). Data for 6 months are given below:**

i. What is correlation? 1

Advertising (X, $'000)	2	4	6	8	10	12
Sales (Y, $'000)	48	56	62	70	78	84

ii. Make a scatter plot from the given data and propose the association of the variables. 2

iii. Find and interpret the correlation coefficient. 3

iv. Fit the regression line and predict the monthly sales when advertising expenditure is $9,000. 4

(g) **The table below displays the weekly hours of practice and corresponding tournament scores of 9 competitive tennis players. It is theorized that tournament score depends on practice hours.**

Practice Hours (X)	5	7	9	10	12	14	16	18	20
Tournament Score (Y)	45	52	60	68	75	82	88	93	98

i. Make a scatter plot using the data. 1

ii. Estimate and interpret the regression coefficient of y on x. 3

iii. Estimate the regression coefficient of x on y. Is this model realistic? Explain. 4

6.2 Short Questions

(a) What is scatter plot? 1

(b) Can a scatter plot measure the accurate relationship between two variables? 2

(c) Differentiate between the values of correlation $r_1 = +1$ and $r_2 = -1$ 2

(d) What is the difference between regression and correlation? 2

(e) What is the range of the Pearson's correlation coefficient (r)? 1

(f) $Y = a + bX$; What does each notation mean in regression context? 2

(g) What can we learn from the scatter plot? 2

(h) What is correlation? 1

(i) Write two purposes of regression analysis. 2

(j) What does $r = -1$ imply? 1

(k) Does the regression coefficient depend on scale? Verify mathematically. 2

(l) What is the range of the correlation coefficient? 1

(m) What is the range of the regression coefficient? 1

(n) What is the relationship between the regression coefficient and the correlation coefficient? 1

(o) What does $b_{yx} = 0.85$ imply? 1

(p) Two sets of variables have correlation $r_1 = 0.75$ and $r_2 = -0.82$. Which set has stonger linear association? 1

(q) How are correlation and regression different? 2

Chapter 7

Time Series

7.1 Creative Questions

(a) **The monthly sales (in units) of laptops at a major electronics store are recorded below:**

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Sales	120	150	130	170	160	180	200	210

i. Compute the trend using a three-monthly moving average method. 3

ii. Illustrate the trend graphically and estimate the expected laptop sales for September. 4

(b) **The yearly revenue (in hundred thousand) of shoe manufacturer company is given below**

Year	2005	2006	2007	2008	2009	2010	2011
Revenue	35	22	40	35	50	42	60

i. What is general trend? 1

ii. Which method of determining trend gives only two values? 2

iii. Determine the trend using three-yearly moving average method. 3

iv. Find the trend using graphical method and extrapolate the approximate revenue earned in 2012. 4

(c) **The annual production (in thousand tons) of a steel manufacturing plant over seven years is given below:**

Year	2015	2016	2017	2018	2019	2020	2021
Production	50	55	60	65	70	75	80

i. Compute the trend using the three-yearly moving average method. 3

ii. Estimate the approximate production for the year 2022 using both graphical and moving average methods. 4

(d) **The annual revenue (in million dollars) of a software company over eight years is given below:**

i. Compute the trend using the four-yearly moving average method. 3

ii. Estimate the approximate revenue for the year 2022 using both graphical and moving average methods. 4

(e) **The average monthly temperature (in degrees Celsius) recorded at a weather station over seven months is given below:**

Month	September	October	November	December	January	February	March
Temperature	22	20	18	15	12	16	20

i. Compute the trend using the three-monthly moving average method. 3

ii. Estimate the approximate average temperature for the month of April using both graphical and moving average methods, and compare the projections. 4

(f) **The monthly rainfall (in mm) recorded in a city over seven consecutive months is given below:**

i. Determine the trend using the three-month moving average method. 3

ii. Predict the expected rainfall for August using both graphical and moving average techniques. 4

(g) **The annual sales (in million dollars) of a tech company over eight years are provided below:**

Year	2014	2015	2016	2017	2018	2019	2020	2021
Revenue	120	135	150	165	180	195	210	225

Month	Jan	Feb	Mar	Apr	May	Jun	Jul
Rainfall (mm)	85	42	78	150	210	95	180

i. Calculate the trend using the three-yearly moving average method. 3
ii. Predict the approximate sales for the year 2022 using both graphical and moving average methods. 4

(h) **The monthly website traffic (in thousands of visits) for an online platform over the past 7 months is provided below:**

i. Determine the trend using the three-month moving average method. 3
ii. Plot the trend and forecast traffic for August using two methods and compare. 4

(i) **The quarterly revenue (in million) of a retail chain is given below:**

i. Calculate the trend using the moving average method for a 3-quarter period. 3
ii. Plot the trend line and predict the revenue for Q8 using two methods and compare. 4

(j) **The monthly sales (in thousands of units) of a product for the last 6 months are given below:**

i. Determine the trend using the three-month moving average method. 3
ii. Plot the trend and forecast the sales for July using two methods and compare. 4

(k) **The quarterly production data (in tons) for a factory is given below:**

i. Calculate the trend using the moving average method for a 3-quarter period. 3
ii. Plot the trend line and predict the production for Q8 using two methods and compare. 4

(l) **Bangladesh foreign debt has been increasing rapidly in recent years. The Bangladesh bank provides the follwoing data.**

i. Name the components of time series. 1
ii. What are linear and non-linear trends? 2
iii. Find 3-yearly moving avergae from the data and plot. 3
iv. Whaich components of time series may underlie the data? Analyze. 4

(m) **GDP (in bn. US$ PPP) of Bangladesh since 1980 to 1985 according to an estimate of International Monetary Fund: 41.2, 47.4, 52.0, 56.5, 61.0, 65.3**

i. What is time series data? 1
ii. What are the components of a time series model? 2
iii. Determine the 3-yearly moving average from the data. 3
iv. Find trend of the data using another method (other than (c)), plot both, and comment which is better. 4

(n) **Annual sales of company are as given in the following**

i. What is a trend? 1
ii. Do the data in the stem seem to have a trend? 2
iii. Find the trend using semi-average method. 3
iv. Find the trend using 2-yearly moving average method. Would it better if we used 3-yearly method? 4

(o) **Daily expense of a certain indvidual in the month of April is tracked by himself (sorted from the first day of the month through the last day). He intends to analyze and see which part of the month is more expensive for him.**

1430 777 4101 4840 3251 4035 2504 371 4326 2296
47 3207 1608 384 4705 3424 1168 1189 4276 4749
1117 156 572 1181 1031 4508 149 80 4475 2087

i. What is semi-average method? 1
ii. Make a line chart from the data? TO BE REVIEWED 2
iii. Find the trend of the data and explain. 3
iv. How can the person accomplish what he intends? 4

(p) **US Dollar to Taka Exchange rate since 2016 to 2023 (for each year, the rate in January is picked) is provided below.**

i. How many methods are there to measure the trend? 1

Year	2014	2015	2016	2017	2018	2019	2020	2021
Sales	120	150	140	160	180	200	220	240

Month	Jan	Feb	Mar	Apr	May	Jun	Jul
Visits	42	48	55	52	60	65	70

ii. Distinguish between seasonal and cyclic variation. 2

iii. Discuss the advantages and disadvantages of moving average method. 3

iv. What is the expected Exchange rate in 2024? Estimate using a suitable method. 4

(q) **Income of a freelancer in 6 successive months (from Jan to Jun) was found to be 46.0, 49.5, 51.5, 50.6, 56.5, and 60 (in thousands BDT.).**

i. What is time series data? 1

ii. What are the components of a time series model? 2

iii. Determine the 3-monthly moving average from the data. 3

iv. Draw the moving averages on a graph paper and interpret. 4

(r) **Average monthly temperatures (in ^oC) in the city of Sylhet are collected by an analyst. The analyst assumes the next month will not follow the current trend.**

i. What is seasonal variation? 1

ii. Differentiate between seasonal variation and cyclic variation. 2

iii. Find the general trend using semi-average method. 3

iv. Find the trend using moving average method and examine the assumption of the analyst. [the genuine next value is 31.2] 4

7.2 Short Questions

(a) What is a time series? 1

(b) Give an example of a time series data. 1

(c) What are the components of Time Series? 2

(d) What is trend? 1

(e) How many types of trends are there? 1

(f) What is linear trend? 1

(g) What is non-linear trend? 1

(h) What is seasonal variation? 1

(i) Give an example of seasonal variation. 1

(j) What is cyclic variation? 1

(k) Distinguish between seasonal and cyclic variation. 1

(l) Give an example of cyclic variation. 1

(m) What is irregular variation? 1

(n) Give an example of irregular variation? 1

(o) How many modelss are there for time series? 1

(p) Write down the additive model of time series. 1

(q) What is the mulplicative model? 1

(r) Mention the methods of measuring the trend. 2

(s) What are the weaknesses of graphical method of measuring the trend? 1

(t) How does semi-average method work? 2

(u) What are the limitations of semi-average method? 2

(v) How is moving average method used? 2

(w) Discuss the advantages and disadvantages of moving average method. 3

(x) When are moving averages centered? 2

Quarter	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Revenue	85	92	88	95	102	98	110

Month	January	February	March	April	May	June
Sales	80	95	110	105	120	115

Quarter	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Production	150	160	155	145	170	165	180

Fiscal Year	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22	2022-23	2023-24
Debt	41.17	45.81	56.01	62.63	68.55	81.62	95.45	98.94	∼130.00

Year	2010	2011	2012	2013	2014	2015	2016
Profit (million)	40	45	46	53	65	70	73

Year	2016	2017	2018	2019	2020	2021	2022	2023
USD Exchange Rate	78.35	79.49	82.87	83.26	84.60	84.37	85.80	106.70

Table 7.1: Source–Investing.com

Month	Jan	Feb	Mar	Apr	May	Jun	Jul
Temperature	25.2	27.1	30.4	30.8	30.9	30.9	31.6

Chapter 8

Published Statistics in Bangladesh

8.1 Creative Questions

(a) **In 2015, tens of thousands of Rohingya people were forcibly displaced from their villages and IDP camps in Rakhine state, Mynmar. Many of them fled to neighboring countries, including Bangladesh, Malaysia, Indonesia. Many national and internations agencies collect data on the issue.**

i. What is non-official statistics? 1
ii. Name five sources of official statistics. 2
iii. Shed some light on the limitations of official statistics. 3
iv. How can the quality of published statistics in Bangladesh be improved? 4

(b) **Climate change is an alarming problem throughout the world. To determine what to do to solve the problem, many government and non-government organizatios collcet and analyze data to come to a consistent solution.**

i. What is official statistics? 1
ii. What is the role of World Meteorological Organization? 2
iii. What are the limitations of published statistics in Bangladesh? 3
iv. How can the quality of published statistics in Bangladesh be improved? 4

(c) **Every country has one or more agencies to deal with statistics of the country for proper management of its assets and population. Bangladesh Bureau of Statistics (BBS) serves as the centralized official bureau in Bangladesh for collecting and disseminating statistics in Bangladesh. USA has several such agencies, like Census Bureau or Bureau of Labor Statistics.**

i. What is data? 1
ii. How is statistics important in planning? 2
iii. Differentiate between official and non-official statistics. 3
iv. Elucidate the classification of published statistics in Bangladesh. 4

(d) **In many industries, statistics plays a vital role in decision-making and strategy formation. A global retail company uses statistical analysis to forecast sales, manage inventory, and improve customer satisfaction. However, despite its benefits, misuse of statistics can lead to poor decision-making and financial losses.**

i. Discuss the role of statistics in business decision-making, highlighting its importance in forecasting and inventory management. 3
ii. Identify and explain the limitations of using statistics in decision-making and discuss a possible case of statistical misuse in business. 4

(e) **Government organizations rely on statistical data to create policies and allocate resources efficiently. In some countries, statistical data on health, education, and economic performance is used by policymakers to make decisions. However, the misuse of data can lead to ineffective or harmful policies.**

i. Define the scope of official statistics in policy-making and the role they play in resource allocation. 3
ii. Discuss the possible consequences of misusing statistical data in policymaking, providing an example from a real-world situation. 4

(f) **Globalization creates many opportuninities. Many government and non-government organizatios collcet and analyze data analyze different aspects of the trend.**

i. What is official statistics? 1
ii. What is the role of the United Nations? 2
iii. What are the limitations of published statistics in Bangladesh? 3
iv. Compare the statistical analysis in Bangladesh with international standard. 4

(g) **The role of official statistics in policy-making is crucial for economic and social development. Governments and international organizations rely on accurate data to formulate strategies and monitor progress.**

i. Discuss the sources of official statistics in Bangladesh and their importance in decision-making. 3
ii. Analyze the limitations of official statistics in Bangladesh and suggest ways to improve their reliability. 4

(h) **Non-official statistics, collected by private organizations and research institutions, play a significant role in complementing official data. These statistics often provide insights into areas not covered by government sources.**

i. Identify the major sources of non-official statistics in Bangladesh and their contributions to data analysis. 3
ii. Compare the strengths and weaknesses of non-official statistics with official statistics in Bangladesh. 4

(i) **The classification of published statistics is essential for organizing data into meaningful categories. This helps in better understanding and utilization of the data for research and policymaking.**

i. Explain the classification system used for published statistics in Bangladesh. 3
ii. Evaluate the effectiveness of this classification system in meeting the needs of researchers and policymakers. 4

8.2 Short Questions

(a) What does BBS stand for? 1
(b) Differentiate between official and non-official statistics. 2
(c) Explain the role of Bangladesh Bureau of Statistics. 2
(d) What is official statistics? 1
(e) Mention 5 sources of published statistics in Bangladesh. 2
(f) What is semi-official statistics? 1
(g) What is non-official statistics? 1
(h) Briefly mention what Bangladesh bank does. 2
(i) Classify the published statistics in Bangladesh. 2
(j) What are the limitations of official statistics? 2
(k) Mention 5 sources of non-official statistics. 2
(l) In Bangladesh, after an interval of how many years is a census conducted? 1

Conclusion

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Contents