UNIT 6

BASIC STATISTICS

Statistics **(K.B)**

The branch of mathematics which deals with collection and manipulation of the numeric data.

Information Handling **(K.B)**

To present the information in a manageable way so that some useful conclusions can be drawn is called information handling.

Data **(K.B)**

The numerical figure obtained from any field of study is known as data.

Types of Data **(K.B)**

There are two types of data

(i) Primary data,

(ii) Secondary data

Primary Data **(K.B)**

The data directly collected from its source is called primary data.

For example:

Prices of fruits collected from market.

Secondary Data **(K.B)**

The data which have been passed through some statistical treatment at least once is called secondary data.

For example:

Office records, printed materials.

Constant **(K.B)**

Any quantity that has a single value is known as constant.

Example:

All real numbers can be taken as constant.

Variables **(K.B)**

Any quantity whose value is not fixed is called variable.

Examples:

Age of person, weight of person, height of person.

Types of Variable **(K.B)**

There are two types of variable.

(i) Discrete variable

(ii) Continuous variable.

Discrete Variable **(K.B + U.B)**

A variable which can take / assume some specific values in the given data is called discrete variable.

Note **(K.B + U.B)**

(i) A discrete variable usually takes the values which are integers or whole numbers
i.e. 0,1,2,3.

(ii) It represents the countable data.

Examples:

(1) Numbers of heads appeared in tossing 5 coins.

(2) No of children in a family.

(3) No of students in a class.

Continuous Variable **(K.B + U.B)**

A variable which can take / assure every possible value with in the given range or interval i.e. (*a* to *b*) is called continuous variable.

Note **(K.B + U.B)**

(1) It may be a whole figure or a fraction.

(2) It represents the measurable data

Examples:

(i) Age of an employee.

(ii) Temperature of a place.

(iii) Height / weight of an individual.

Ungrouped Data **(K.B)**

The numerical facts obtained on first hand and record as they stand are called ungrouped data.

Grouped Data **(K.B)**

The data which have gone / passed through some statistical process is called grouped

data. It may be classified into certain groups or into rows and columns.

Classification **(K.B)**

The process of arranging the data into certain groups or classes having similar characteristics is called classification.

Note about Number of Classes **(K.B)**

(i) More than 15 groups or classes are generally not advisable.

(ii) There is no hard and fast rule to find exact number of classes.

(iii) Groups / classes should be between 5 and 15.

(iv) Too small number of groups / classes results into loss of important information.

(v) Taking too many groups do not pay for the labour involved condensing the information.

Example: **(K.B)**

Marks obtained in maths out of 100.Form a frequency distribution.

60, 62, 65, 66, 67, 68, 66, 60, 73, 69, 70, 63, 64, 63, 67

Given Data:

Total number of observations = 15

The smallest number = 60,

The largest number = 73

Frequency Distribution

Groups	No of Students
60 ––––– 62	3
63 ––––– 65	4
66 ––––– 68	5
69 ––––– 71	2
72 ––––– 74	1
	$n = 15$

Class Limits **(K.B)**

Each class / group is defined by two values one small and other large these values are called class limits. The smaller one is called lower class limit.

The large one is called upper class limit.

Size of Class Interval **(K.B)**

The difference between lower or upper class limits of any two consecutive classes / groups is called size or length of class interval. It is denoted by h.

Groups	No. of Students	Class Marks (X)	Class Boundaries
60 ––––– 62	3	$\frac{60+62}{2} = \frac{122}{2} = 61$	59.5 - 62.5
63 ––––– 65	4	$\frac{63+65}{2} = \frac{128}{2} = 64$	62.5 - 65.5
66 ––––– 68	5	$\frac{66+68}{5} = \frac{134}{2} = 67$	65.5 - 68.5

Class Mark / Mid Point **(K.B)**

The average value of lower and upper class limits of any class interval is called class mark, it is the mid point of any class. Class marks are represented by X.

For C.I, 60 – 62, 63 – 65,....

Class mark = $(X) = \frac{60+62}{2} = \frac{122}{2} = 61$

Class Boundaries **(K.B)**

The values which describe true class limits of a class are called class boundaries.

The smaller value is called lower class boundary.

The larger value is called upper class boundary.

Class Frequency **(K.B)**

The number of occurrence of items corresponding to the class interval is called class frequency. It is denoted by f.

Frequency Distribution **(K.B)**

The tabular arrangement of data in which various items are arranged into classes and the number of items lying in each class is called frequency distribution.

Formation of Frequency Distribution **(K.B)**

By Tally Bar Method

By Direct Method

Cumulative Frequency (K.B)

The total of frequency upto an upper class limit or boundary is called the cumulative frequency.

Histogram (K.B)

A histogram is a graph of adjacent rectangle constructed on *XY*-plane. It is a graph of frequency distribution, on which class boundaries or class marks are marked on x-axis and frequency on y-axis.

Frequency Polygon (K.B)

Frequency polygon is a many sided closed figure on XY plane, on which class marks are marked on X-axis and frequencies on Y-axis.

Cumulative Frequency Polygon / Ogive (K.B)

A cumulative frequency polygon or ogive is a graph of many sided closed figure of less than cumulative frequency distribution, in which cumulative frequencies are marked on Y-axis and upper class boundaries on X-axis.

Exercise 6.1

Q.1 The following data shows the number of members in various families. Construct frequency distribution. Also find cumulative frequencies. 9, 11, 4, 5, 6, 8, 4, 3, 7, 8, 5, 5, 8, 3, 4, 9, 12, 8, 9, 10, 6, 7, 7, 11, 4, 4, 8, 4, 3, 2, 7, 9, 10, 9, 7, 6, 9, 5, 7. **(A.B)**

Solution:

Min. Value = 2, Max. Value = 12, Total values = 39

Discrete Frequency Distribution

No. of members	Tally bars	No. of families (*f*)	C.F
2	|	1	1
3	|||	3	1+3=4
4	𝍸 |	6	4+6=10
5	||||	4	10+4=14
6	|||	3	17
7	𝍸 |	6	23
8	𝍸	5	28
9	𝍸 |	6	34
10	||	2	36
11	||	2	38
12	|	1	39

Q.2 The following data has been obtained after weighing 40 students of class V. Make a frequency distribution taking class interval size as 5. Also find the class boundaries and midpoints. 34, 26, 33, 32, 24, 21, 37, 40, 41, 28, 28, 31, 33, 34, 37, 23, 27, 31, 31, 36, 29, 35, 36, 37, 38, 22, 27, 28, 29, 31, 35, 35, 40, 21, 32, 33, 27, 29, 30, 23. **(A.B)**

Also make a less than cumulative frequency distribution. (Hint: Make classes 20-24, 25-29…).

Solution:

Min. Value = 21, Max. Value = 41, Size of class interval = 5, Total values = 40

Frequency Distribution

C – I	Tally Bars	Frequency	C – B	Mid points
15-19		-	14.5-19.5	17
20-24	𝍸 |	6	19.5-24.5	22
25-29	𝍸 𝍸	10	24.5-29.5	27
30-34	𝍸 𝍸 ||	12	29.5-34.5	32
35-39	𝍸 |||	9	34.5-39.5	37
40-44	|||	3	39.5-44.5	42
Total		40		

Less than cumulative frequency distribution

C – B	C.f
Less than 19.5	0
Less than 24.5	6
Less than 29.5	16
Less than 34.5	28
Less than 39.5	37
Less than 44.5	40

Q.3 From the following data representing the salaries of 30 teachers of a school. Make a frequency distribution taking class interval size of Rs. 100, 450, 500, 550, 580, 670, 1200, 1150, 1120, 950, 1130, 1230, 890, 780, 760, 670, 880, 890, 1050, 980, 970, 1020, 1130, 1220, 760, 690, 710, 750, 1120, 760, 1240. **(A.B + U.B +K.B)**

(Hint: Make classes 450-549, 550-29…).

Solution:

Min. Value = 450, Max. Value = 1240, Size of class interval = 100 , total values = n = 30

Frequency Distribution

C – I	Tally Bars	Frequency (*f*)
450 – 549	||	2
550 – 649	||	2
650 – 749	||||	4
750 – 849	𝍸 |	5
850 – 949	|||	3
950 – 1049	||||	4
1050 – 1149	𝍸 |	5
1150 - 1249	𝍸 |	5
Total		30

Q.4 The following data shows the daily load shedding duration in hours in 31 localities of a certain city. Make a frequency distribution of the load shedding duration taking 2 hours as class interval size and answer the following questions. 6, 12, 5, 7, 8, 3, 6, 7, 10, 2, 14, 11, 12, 8, 6, 8, 9, 7, 11, 6, 9, 12, 13,10, 14, 7, 6, 10, 11, 14, 12**.**

(A.B + U.B +K.B)

(a) Find the most frequent load shedding hours?

(b) Find the least load shedding intervals?

(Hint: Make classes 2 - 3, 4 - 5, 6 - 7….)

Solution:

Min. Value =2, Max. Value = 14, Size of class interval = 2, total values = n = 31

Frequency Distribution (By Direct Method)

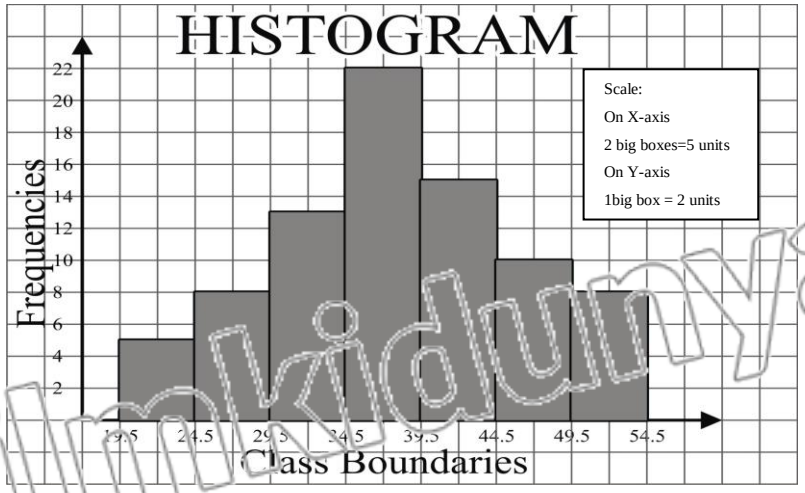
C – I	Values	Frequency (f)
2 – 3	3, 2	2
4 – 5	5	1
6 – 7	6, 7, 6, 7, 6, 7, 6, 7, 6	9
8 – 9	8, 8, 8, 9, 9,	5
10 – 11	10, 11, 11, 10, 10, 11	6
12 – 13	12, 12, 12, 13, 12	5
14 – 15	14, 14, 14	3
Total		31

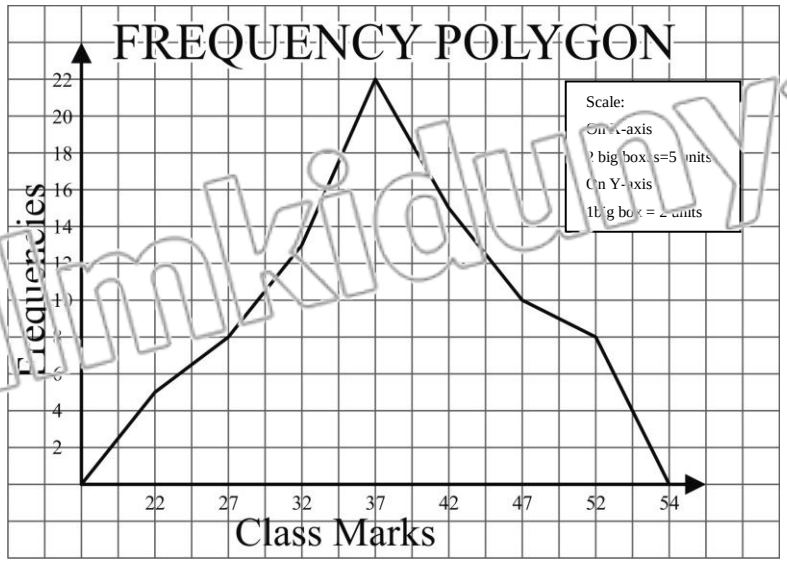
(a) The most frequent load shading hours = 6 – 7

(b) The least load shading intervals = 4 – 5

Q.5 Construct a Histogram and frequency Polygon for the following data showing weights of students in kg. **(A.B + U.B +K.B)**

Weights	Frequency / No. of Students	Mid Values	Class Boundaries
20-24	5	22	19.5-24.5
25-29	8	27	24.5-29.5
30-34	13	32	29.5-34.5
35-39	22	37	34.5-39.5
40-44	15	42	39.5-44.5
45-49	10	47	44.5-49.5
50-54	8	52	49.5-54.5





Average **(K.B)**

A single value which represents the data is called average.

Types of Average **(A.B + K.B)**

(i) Arithmetic mean
(ii) Median
(iii) Mode
(iv) Geometric mean
(v) Harmonic mean
(vi) Quartiles

Arithmetic Mean **(A.B + K.B)**

A value obtained by dividing the sum of all the observations by their number of observations is called arithmetic mean it is denoted by $\overline{X}$

Let X_1, X_2, X_3,——— ,X_n be the values of a data thus the A.M is defined as:

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \ldots\ldots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

where i= 1,2,3,4,…n

$$\Rightarrow \overline{X} = \frac{\sum X}{n} \quad \text{(For ungrouped data)}$$

$$\overline{X} = \frac{\sum fx}{\sum f} \quad \text{(For grouped data)}$$

Arithmetic Mean (By short cut method) **(A.B + K.B)**

$$\overline{X} = A + \frac{\sum D}{n} \quad \text{(For ungrouped data.)}$$

$$\overline{X} = A + \frac{\sum fD}{\sum f} \quad \text{(For grouped data)}$$

where D = X – A

A = Provisional Mean or Assumed Mean and X= Mid point of a class

Arithmetic Mean (By Coding Method) **(A.B + U.B)**

$$\overline{X} = A + \frac{\sum u}{n} \times h \quad \text{(For ungrouped data)}$$

where 'h' is common difference / size of C-I

$$\overline{X} = A + \frac{\sum fu}{\sum f} \times h \quad \text{(For grouped data)}$$

where $u = \frac{X - A}{h}$

Median **(U.B + K.B)**

A value which divides an arranged data into two equal parts (i.e., 50% data before the median and 50% after it is called median.

Or

Median is the middle most observation in an arranged data set. It divides the data set into two equal parts.

It is represented on $\tilde{x}$.

Formulas (K.B)

(For ungrouped data)

(i) If "n" is odd number

Median = $\frac{n+1^{th}}{2}$ term

If "n" is even number

Median = $\frac{1}{2}\left[\left(\frac{n}{2}\right)^{th} item + \left(\frac{n+2}{2}\right)^{th} item\right]$

(For grouped data)

Median = $l + \frac{h}{f}\left(\frac{n}{2} - c\right)$

where

l = lower class boundary of median class

h = Size of median class

f = Frequency of median class

c = Preceding commutative frequency of median class

$n/2$ is used to decide the median class

$n = \sum f$

Mode (K.B)

The most repeated value or most frequent value in the given data is called mode.

Formula:

For Ungrouped data

Mode = the most repeated value in given data. (No Formula)

For Grouped data

$$\text{Mode} = l + \frac{(f_m - f_1)}{2f_m - f_1 - f_2} \times h$$

Where

l = lower class boundary of modal class.

f_m = Frequency of mode class

f_1= Preceding frequency of modal frequency class.

f_2 = Following frequency of modal class.

h = size of modal class.

Empirical relation among Mean, Median and Mode (K.B)

Mode = 3 Median – 2 Means

Geometric Mean (K.B)

Geometric mean of a variable X is the n^{th} positive root of the product of the $x_1, x_2, x_3, ..., x_n$ observation. In symbols, we write

$$\text{G.M.} = (x_1, x_2, x_3, ..., x_n)^{\frac{1}{n}}$$

The above formula can also be written by using logarithm.

For Ungrouped data

$$\text{G.M.} = Anti\log\left(\frac{\sum \log x}{n}\right)$$

For Grouped data

$$\text{G.M.} = Anti\log\left(\frac{\sum f \log X}{\sum f}\right)$$

Harmonic Mean (K.B)

Harmonic Mean refers to the value obtained by reciprocating the mean of the reciprocal of $x_1, x_2, x_3, ..., x_n$ observations. In symbols,

For ungrouped data,

$$\text{H.M.} = \frac{n}{\sum \frac{1}{X}}$$

For Grouped data,

$$\text{H.M.} = \frac{\sum f}{\sum \frac{f}{X}}$$

Properties of Arithmetic Mean (K.B + U.B)

(i) Mean of a variable with similar observations say constant k is the constant k itself.

(ii) Mean is affected by change in origin.

(iii) Mean is affected by change in scale.

(iv) Sum of the deviations of the variable X from its mean is always zero.

The Weighted Arithmetic Mean (K.B + U.B)

The relative importance of a number is called its weight. When numbers $x_1, x_2, ..., x_n$ are not equally important, we associate them with certain weight, $w_1, w_2, w_3, ..., w_n$ depending on the importance or significance. Average of such data is called weighted arithmetic mean. It can be calculated by the formula

$$\bar{x}_w = \frac{\sum wx}{\sum w}$$

Moving Averages **(K.B + U.B)**

Moving averages are defined as the successive averages (arithmetic means) which are computed for a sequence of days/months/years at a time. Place the average of each period against the mid of the each period. Dropping first value and adding succeeding value to this group and continue the process until all values from first to last are used.

Quartiles **(K.B)**

It divides the data in four equal parts.

Formula: (for grouped data)

$$Q_i = l + \frac{h}{f}\left(\frac{in}{4} - c\right) \; ; i = 1,2,3$$

Note **(K.B)**

i) Q_2 = Median

ii) Q_1 = Lower Quartiles

iii) Q_3 = Upper Quartiles

Deciles **(K.B)**

It divides the data in ten equal parts.

Formula: (for grouped data)

$$D_i = l + \frac{h}{f}\left(\frac{in}{10} - c\right) \; ; \; i = 1,2,...,9$$

Note **(K.B)**

D_5 = Median

Example 1: (Page # 119) **(A.B)**

The marks of seven students in Mathematics are as follows. Calculate the Arithmetic Mean and interpret the result.

Student No	1	2	3	4	5	6	7
Marks	45	60	74	58	65	63	49

Solution:

Let X = marks of a students

$$\bar{X} = \frac{\sum X}{n} = \frac{x_1 + x_2 + x_3 + ... + x_7}{7}$$

Or $$\bar{X} = \frac{45+60+74+58+65+63+49}{7} = \frac{414}{7} = 59.14 \text{ marks}$$

Example 2 (Page # 120) **(GRW 2014, FSD 2017, D.G.K 2017)** **(A.B)**

The salaries of five teachers are as follows. Find the mean salary using direct and indirect methods and compare the results. 11500, 12400, 1500, 14500, 14800.

Solution:

We proceed as follows:

(a) Using Direct method

$$\bar{X} = \frac{\sum_{i=1}^{5} x_i}{5} = \frac{11500+12400+15000+14500+14800}{5}$$

$$= \frac{74000}{5} = 13640 \text{ Rupees.}$$

(b) Indirect methods:

We assume $A = 1300$, $D_i = (x_i - 13000)$, $h = 100$ and $u_i = \frac{(x_i - A)}{100}$, the computations are shown in the following table:

X	$D_i=(x_i-13000)$	$u_i=\frac{(x_i-A)}{100}$
11500	–1500	–15
12400	–600	–6
15000	2000	20
14500	1500	15
14800	1800	18
$\sum X_i=74000$	$\sum D_i=3200$	$\sum u_i=32$

(i) Short formula:

$$\bar{X}=13000+\frac{3200}{5}=13000+640=13640 \text{ Rupees}$$

(ii) Coding method:

$$\bar{X}=13000+\frac{32}{5}\times 100=13640 \text{ Rupees}$$

Example 5 (Page # 122) **(A.B)**

Find arithmetic mean using short formula taking $X=34.5$ as the provisional mean in example 4.

Solution:

We use the following formula

(i) $\bar{X}=A+\frac{\sum fD}{\sum f}$ **(ii)** $\bar{X}=A+\frac{\sum fu}{\sum f}\times h$

Given $A=34.5,$ we notes that the distribution has equal class interval size of 10. So we may take $h=10$ and make the following calculations:

Classes/groups	f	Midpoints x	$D=X-34.5$	$u=(X-A)/10$	fD	fu
0---9	2	4.5	–30	–3	–60	–6
10---19	10	14.5	–20	–2	–200	–20
20---29	5	24.5	–10	–1	–50	–5
30---39	9	34.5	0	0	0	0
40---49	6	44.5	10	1	60	6
50---59	7	54.5	20	2	140	14
60---69	1	64.5	30	3	30	3
Total	40		1300		–80	–8

Substituting the totals in the above formulae we get

(i) $\bar{X}=34.5+\frac{(-80)}{40}=34.5-2=32.5$ gm

(ii) $\bar{X}=34.5+\left(\frac{-8}{40}\right)\times 10=34.5-2=32.5$ gm

Example 1 (Page # 123) **(A.B)**

On 5 term test in mathematics, a student has made marks of 82,93,86,92 and 79. Find the median for the marks.

Solution:

By arranging the grades in ascending order, the arranged data is $79, 82, 86, 92, 93$

Since number of observation is odd i.e., $n = 5$.

$\tilde{x} =$ size of $\left(\frac{5+1}{2}\right)^{th}$ observation

$\tilde{x} =$ size of 3^{rd} observation

$\tilde{x} = 86$

Example 2 (Page # 123) **(LHR 2017, FSD 2017, 18, SGD 2014, MTN 2016)** **(A.B)**

The sugar contents for a random sample of 6 packs of juices of a certain brand are found to be 2.3, 2.7, 2.5, 2.9, 3.1 and 1.9 milligram. Find the median.

Solution:

Arranging the values by increasing order of magnitude

1.9, 2.3, 2.5, 2.7, 2.9, 3.1

Since the number of observations are even i.e., $n = 6$.

$$\tilde{x} = \frac{1}{2}\left[\text{size of } \left(\frac{6}{2}\text{th} + \frac{6+2}{2}\text{th}\right) \text{observations}\right]$$

$$\tilde{x} = \frac{1}{2}\left[\text{size of } (3\text{rd} + 4\text{th}) \text{ observations}\right]$$

$$\tilde{x} = \frac{2.5 + 2.7}{2} = 2.6 \text{ milligram.}$$

Example 4 (Page # 125) **(A.B)**

The following data is the time taken by 40 students to solve a problem is recorded. Find the median time taken by the students.

138	164	150	132	144	125	149	157
146	158	140	147	136	148	152	144
168	126	138	176	163	119	154	165
146	173	142	147	135	153	140	135
161	145	135	142	150	156	145	128

Solution:

Class Intervals	Frequency	Class boundaries	Cumulative Frequency
118 — 126	**3**	**117.5 – 126.5**	**3**
127 — 135	**5**	**126.5 – 135.5**	**8**
136 — 144	**9**	**135.5 – 144.5**	**19**
145 — 153	**12**	**144.5 – 153.5**	**29**
154 — 162	**5**	**153.5 – 162.5**	**34**
163 — 171	**4**	**162.5 – 171.5**	**38**
172 — 180	**2**	**171.5 – 180.5**	**40**
Total	$40 = \sum f$	—	—

Median class = class containing $\left(\frac{n}{2}\right)^{th}$ observation

Median class = class containing $\left(\frac{40}{2}\right)^{th} = 20^{th}$ observation

$$\tilde{x} = l + \frac{h}{f}\left(\frac{n}{2} - c\right) = 144.5 + \frac{9}{12}(20 - 17) = 146.8$$

Example 1 (Page # 127) **(SGD 2014)** **(A.B)**

Find the geometric mean of the observations 2,4,8 using

(i) basic formula and **(ii)** using logarithmic formula.

Solution:

Basic formula

$$G.M = (2\times 4\times 8)^{\frac{1}{3}} = (64)^{\frac{1}{3}} = 4$$

Using log method:

X	$\log X$
2	0.3010
4	0.6021
8	0.9031
Total	1.8062

$$\text{G.M} = \text{Antilog}\left(\frac{\sum \log X}{n}\right) = \text{Antilog}\left(\frac{1.8062}{3}\right)$$

$$= \text{Antilog}\ (0.6021) = 4.00003 = 4$$

Example 2: (Page # 130) **(A.B + U.B)**

A variable X take the following values $4,5,8,6,2$. Find the mean of X, also find the mean when (a) 5 is added to each observation (b) 10 is multiplied with each observation (c) Prove that sum of the deviation from mean is zero.

Solution:

Given the values of X,

X: 4 5 8 6 2.

We may introduce two new variables Y and Z under (a) and (b) respectively. So we are given that $(a) = Y = X + 5$ $(b) Z = 10X$. The following table shows the desired result:

X	$Y = X + 5$	$Z = 10X$	$X - \bar{X}$
4	9	40	–1
5	10	50	0
8	13	80	3
6	11	60	1
2	7	20	–3
25	50	250	$\sum(X - \bar{X}) = 0$

From the above table we get,

$$\bar{X} = \frac{\sum X}{n} = \frac{25}{5} = 5 \quad ; \quad \bar{Y} = \frac{\sum Y}{n} = \frac{50}{5} = 10 \quad ; \quad \bar{Z} = \frac{\sum Z}{n} = \frac{250}{5} = 50$$

Note **(A.B + U.B + K.B)**

(a) $\bar{Y} = 10 = 5 + 5 = \bar{X} + 5$

(b) $\bar{Z} - 50 = 10(5) = 10\bar{X}$

Which shows that mean is affected by change in origin and scale.

(c) From the last column of the table we note that $\sum(X - \bar{X}) = 0$, the sum of the deviations from mean is zero.

Example 1: (Page # 131) **(A.B)**

The following table gives the monthly earnings and the number of workers in a factory, compute the weighted average.

No. of employees	Monthly earnings. Rs.
4	800
22	45
20	100
30	30
80	35
300	15

Solution:

Number of employees are treated as a weight (w) and monthly earnings as variable (x)

No. of employees (w)	Monthly earning in Rs. (x)	(xw)
4	800	3200
22	45	990
20	100	2000
30	30	900
80	35	2800
300	15	4500
$\sum w = 456$	_____	$\sum xw = 14390$

$$\bar{X}_w = \frac{\sum xw}{\sum w} = \frac{14390}{456} = 31.5$$

Example 2: (Page # 131) **(A.B)**

Calculate three days moving average for the following record of attendance:

Week	Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	24	55	28	45	51	54	60

Solution:

Week and days	Attendance	3-days moving	
		Total	Average
Sun.	24	——	——
Mon.	55	107	107/3 = 35.67
Tue.	28	128	128/3 = 42.67
Wed.	45	124	124/3 = 41.33
Thu.	51	150	150/3 = 50.00
Fri.	54	168	168/3 = 56.00
Sat.	60	——	——

Example 1: (Page # 132) **(A.B)**

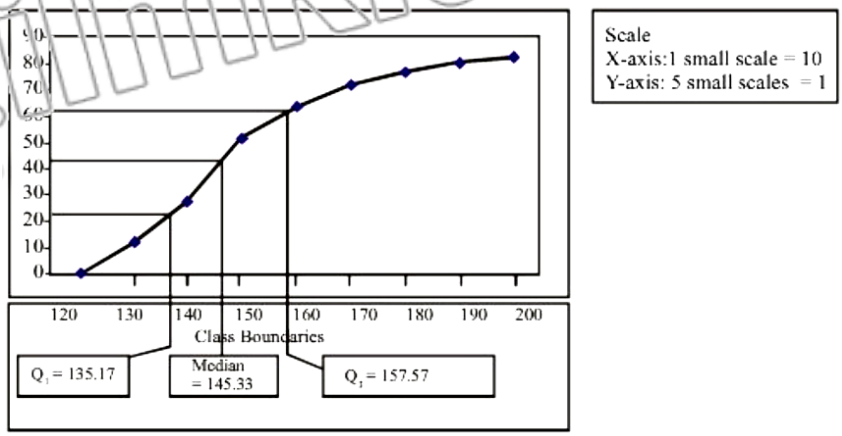
For the following distribution locate Median and Quartiles on graph.

Class boundaries	Cumulative frequency
Less than 120	0
Less than 130	12
Less than 140	27
Less than 150	51

Less than 160	64
Less than 170	71
Less than 180	76
Less than 190	80
Less than 200	82

Solution:

We locate median and Quartiles by using the following cumulative frequency polygon.



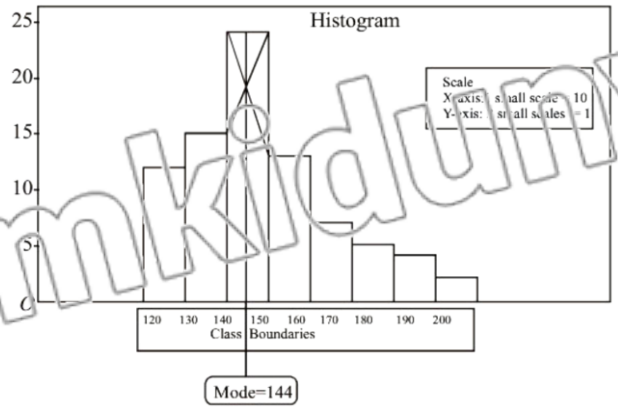
Example 2: (Page # 133) **(A.B + U.B + K.B)**

For the following distribution locate Mode on graph.

Salaries in Rupees	No. of teachers
120 — 130	12
130 — 140	15
140 — 150	24
150 — 160	13
160 — 170	7
170 — 180	5
180 — 190	4
190 — 200	2

Solution:

On Histogram, the mode is located on X-axis as shown below:



Exercise 6.2

Q.1 What do you understand by measures of central tendency? (A.B)

Ans: A single value which represents the data is called as average. As average tends to lie in the centre of the data, so to locate an average is called measure of central tendency.

Q.2 Define Arithmetic mean, Geometric mean, Harmonic mean, mode and median. (A.B)

Ans: See Definition page # 188, 189

Q.3 Find arithmetic mean by direct method for the following set of data: (A.B)

(i) 12, 14, 17, 20, 24, 29, 35, 45.

(ii) 200, 225, 350, 375, 270, 320, 290.

Ans:

(i) Given Data:

12, 14, 17, 20, 24, 29, 35, 45

(LHR 2014, 16, GRW 2017, FSD 2017, RWP 2015, 17, MTN 2016, D.G.K 2014, 16)

Required:

Arithmetic mean by direct method

Solution:

X	12	14	17	20	24	29	35	45	$\sum X = 196$

Formula:

$$\overline{X} = \frac{\sum X}{n}$$

$$= \frac{196}{8}$$

$$= 24.5$$

Result

$$\overline{X} = 24.5$$

(ii) Given Data:

200, 225, 350, 375, 270, 320, 290 **(GRW 2014, 16, D.G.K 2016)**

Required:

Arithmetic mean by direct method

Solution:

X	200	225	350	375	270	320	290	$\Sigma X = 2030$

We known that

$$\overline{X} = \frac{\sum X}{n} = \frac{2030}{7}$$

$$= 290$$

Result: $\overline{X} = 290$

Q.4 For each of the data in Q. No 3, compute arithmetic mean using indirect method. (A.B)

(i) Give Data:

12, 14, 17, 20, 24, 29, 35, 45

Required:

Arithmetic mean by indirect method

Solution:

X	D = X – 20
12	– 8
14	– 6
17	– 3
20	0
24	4
29	9
35	15
45	25
	ΣD = 36

We know that

$$\overline{X} = A + \frac{\Sigma D}{n}$$

Putting the values

$$= 20 + \frac{36}{8}$$

$$= 20 + 4.5 = 24.5$$

Result:

$\overline{X}$ **= 24.5**

(ii) Give Data: **(A.B)**

200, 225, 350, 375, 270, 320, 290

Required:

Arithmetic mean by indirect method

Solution: (By using short-cut and coding method)

X	$D = X - 200$	$U = \frac{X - 200}{5}$
200	0	0
225	25	5
270	70	14
290	90	18
320	120	24
350	150	30
375	175	35
	ΣD = 630	ΣU = 126

(H.C.F of the difference of class marks is 5, so *h* is 5)

Arithmetic mean by short-cut method

Formula:

$$\overline{X} = A + \frac{\Sigma D}{n}$$

$$= 200 + \frac{630}{7}$$

$$= 200 + 90$$

$$\overline{X} = 290$$

Arithmetic mean by coding method

Formula:

$$\bar{X}=A+\frac{\sum U}{n}\times h$$

$$=200+\frac{126}{7}\times 5$$

$$=200+\frac{630}{7}$$

$$=200+90=290$$

Result $\bar{X}$ **= 290**

Q.5 **The marks obtained by students of class XI in mathematics are given below. Compute arithmetic mean by direct and indirect methods.** **(LHR 2015)** **(A.B)**

Given Data:

Classes/Groups	Frequency
0 – 9	2
10 – 19	10
20 – 29	5
30 – 39	9
40 – 49	6
50 – 59	7
60 – 69	1

Required:

Arithmetic mean by direct and indirect method

Solution:

C – I	*f*	*X*	*fX*	*D=X- 34.5*	*fD*	$U=\frac{X-34.5}{10}$	*fU*
0 – 9	2	4.5	9	– 30	– 60	– 3	– 6
10 – 19	10	14.5	145	– 20	– 200	– 2	– 20
20 – 29	5	24.5	122.5	– 10	– 50	– 1	– 5
30 – 39	9	34.5	310.5	0	0	0	0
40 – 49	6	44.5	267	10	60	1	6
50 – 59	7	54.5	381.5	20	140	2	14
60 – 69	1	64.5	64.5	30	30	3	3
	$\sum f = 40$		$\sum fx=1300$		$\sum D = -80$		$\sum U = -8$

Arithmetic mean by direct method

Formula

$$\overline{X} = \frac{\sum fx}{\sum f}$$

$$= \frac{1300}{40}$$

$$= 32.5$$

Results

$\overline{X}$ **= 32.5**

Arithmetic mean by short cut method

Formula

$$\overline{X} = A + \frac{\sum fD}{\sum f}$$

Putting the values

$$\Rightarrow \overline{X} = 34.5 + \left(\frac{-80}{40}\right)$$

$$= 34.5 - 2$$

$$\overline{X} = 32.5$$

Results

$\overline{X}$ **= 32.5**

Arithmetic mean by coding method

Formula

$$\overline{X} = A + \frac{\sum fD}{\sum f} \times h$$

$$= 34.5 + \frac{-8}{40} \times 10$$

$$= 34.5 - 2$$

$$\overline{X} = 32.5$$

Results

$\overline{X}$ **= 32.5**

Q.6 The following data relates to the ages of children in a school. Compute the mean age by direct and short-cut method taking any Provisional mean. (Hint. Take A = 8)

(A.B)

Given Data:

Class limits	Frequency
4 – 6	10
7 – 9	20
10 – 12	13
13 – 15	7
Total	50

Required:

(i) Arithmetic mean by direct and short-cut method

(ii) G.M

(iii) H.M

Solution:

C – I	*f*	*X*	*fX*	*D = X– 8*	*fD*	log *X*	*f* log *X*	$\frac{f}{X}$
4 – 6	10	5	50	– 3	– 30	0.6990	6.99	2.0000
7 – 9	20	8	160	0	0	0.9031	18.062	2.5
10 – 12	13	11	143	3	39	1.0414	13.5382	1.1818
13 – 15	7	14	98	6	42	1.1461	8.0227	0.5
	$\sum f = 50$		$\sum fX = 451$		$\sum fD = 51$		$\sum f \log X = 46.6129$	$\sum \frac{f}{X} = 6.1818$

Arithmetic mean by direct method

$$\overline{X} = \frac{\sum fX}{\sum f}$$
$$= \frac{451}{50}$$
$$= 9.02$$

Arithmetic mean by short method

$$\overline{X} = A + \frac{\sum fD}{\sum f}$$
$$= 8 + \frac{51}{50}$$
$$= 8 + 1.02$$
$$= 9.02$$

Geometric Mean:

$$G.M = \text{anti log}\left(\frac{\sum f \log X}{\sum f}\right)$$
$$= \text{antilog}\left(\frac{46.6129}{50}\right)$$
$$= \text{antilog}\,(0.9323)$$
$$G.M = 8.5566$$

Harmonic Mean:

$$H.M = \frac{\sum f}{\sum\left(\frac{f}{X}\right)}$$
$$= \frac{50}{6.1818}$$
H.M=8.0883

Q.7 The following data shows the number of children in various families. Find mode and median. 9, 11, 4, 5, 6, 8, 4, 3, 7, 8, 5, 5, 8, 3, 4, 9, 12, 8, 9, 10, 6, 7, 7, 11, 4, 4, 8, 4, 3, 2, 7, 9, 10, 9, 7, 6, 9, 5. **(FSD 2015, SGD 2015)** **(A.B)**

Required:

(i) Mode

(ii) Median

Solution:

Min Value = 2 , Max. Value = 12

Frequency Distribution

X	Tally Bars	Frequency	C.f
2	|	1	1
3	|||	3	4
4	𝍸 |	6	10
5	||||	4	14
6	|||	3	17
7	𝍸	5	22
8	𝍸	5	27
9	𝍸 |	6	33
10	||	2	35
11	||	2	37
12	|	1	38

Mode **(A.B + U.B)**

Mode = 4 and 9 (Repeating maximum number of times)

Median **(A.B + U.B)**

Median= $\left(\frac{n}{2}\right)^{th}$ term

$= \left(\frac{38}{2}\right)^{th}$ term

$= 19^{th}$ term

Median = 7

Q.8 Find Modal number of heads for the following distribution showing the number of heads when 5 coins are tossed. Also determine median. (A.B + U.B + K.B)

Given Data:

X (No. of Heads)	Freq. (No. of times)
1	3
2	8
3	5
4	3
5	1

Required:

(i) Mode

(ii) Median

Solution:

X	f	C.f
1	3	3
2	8	11
3	5	16
4	3	19
5	1	20

Mode

Mode =2 (Repeating max. number of times)

Median

Median $= \left(\frac{n}{2}\right)^{th}$ term $= \left(\frac{20}{2}\right)^{th}$ term $= 10^{th}$ term $= 2$

Q.9 The following frequency distribution the weights of boys in kilogram. Compute mean, median, mode. (A.B + U.B + K.B)

Given Data:

Class intervals	Frequency
1 – 3	2
4 – 6	3
7 – 9	5
10 – 12	4
13 – 15	6
16 – 18	2
19 – 21	1

Required:

(i) Mean =? **(ii)** Median =? **(iii)** Mode =?

Solution:

C – I	f	X	fX	C – B	C.f
1 – 3	2	2	4	0.5 – 3.5	2
4 – 6	3	5	15	3.5 – 6.5	5
7 – 9	5	8	40	6.5 – 9.5	10
10 – 12	$4f_1$	11	44	9.5 – 12.5	14
13 – 15	$6f_m$	14	84	12.5 – 15.5	20
16 – 18	$2f_2$	17	34	15.5 – 18.5	22
19 – 21	1	20	20	18.5 – 21.5	23
	$\sum f=23$		$\sum fX = 241$		

(i) Mean: **(A.B)**

Formula:

$$\overline{X} = \frac{\sum fX}{\sum f}$$

Putting the values

$$= \frac{241}{23}$$

$$\Rightarrow \overline{X} = 10.478kg$$

(ii) Median: **(A.B)**

Formula:

$$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - c\right)$$

$$= 9.5 + \frac{3}{4}(11.5 - 10)$$

$$= 9.5 + 1.125$$

$$= 10.625$$

$$\Rightarrow \text{Median} = 10.625kg$$

(iii) Mode: **(A.B)**

Formula:

$$\text{Mode} = l + \frac{f_m - f_1}{2f_m - f_1 - f_2} \times h$$

Putting the values

$$= 12.5 + \frac{6-4}{2(6)-4-2} \times 3$$

$$= 12.5 + \frac{2 \times 3}{6}$$

$$= 12.5 + 1$$

$$\Rightarrow \text{Mode} = 13.5kg$$

Q.10 A student obtained the following marks at a certain examination: English 73, Urdu 82, Mathematics 80, History 67 and Science 62. **(A.B + U.B)**

(i) If the weights accorded these marks are 4, 3, 3, 2 and 2, respectively, what is an appropriate average mark?

(ii) What is the average mark if equal weights are used?

Given Data:

Subject	Marks	Weights
English	73	4
Urdu	82	3
Mathematics	80	3
History	67	2
Science	62	2

Required:

(i) What is an appropriate average mark? (i-e Weighted means)

(ii) What is the average mark if equal weights are used (simple means)

Solution:

Subject	Marks (x)	Weights (w)	Wx
English	73	4	292
Urdu	82	3	246
Mathematics	80	3	240
History	67	2	134
Science	62	2	124
Total	364	14	1036

(i) Weighted Mean:

$$\bar{X}_\omega = \frac{\sum wx}{\sum w}$$

$$= \frac{1036}{14}$$

$$\Rightarrow \bar{X}_\omega = 74 \text{marks}$$

(ii) When equal weights are used

$$\bar{X} = \frac{\sum X}{n}$$

$$= \frac{364}{5}$$

$$= 72.8 \ \textit{marks}$$

Q.11 On a vacation trip a family bought 21.3 liters of petrol at 39.90 rupees per liter, 18.7 liters at 42.90 rupees per liter, and 23.5 liters at 40.90 rupees per liter. Find the mean price paid per liter. (A.B)

Given Data:

No. of liters	Prices (Rs)
21.3	39.90
18.7	42.90
23.5	40.90

Required:

Mean price paid per liters = $\bar{X}_\omega$ =?

No. of liters (w)	Price (x)	wx
21.3	39.90	849.87
18.7	42.90	802.23
23.5	40.90	961.15
$\sum w = 63.5$		$\sum wx = 2613.25$

Formula

$$\bar{X}_\omega = \frac{\text{Total payment}}{\text{Total liters}} = \frac{\sum wX}{\sum w}$$

$$= \frac{2613.25}{63.5}$$

$$= 41.15$$

Result:

Mean price paid per liter $= Rs\ 41.15$

Q.12 Calculate simple moving average of 3 years from the following data: (A.B)

Given Data:

Years	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Values	102	108	130	140	158	180	196	210	220	230

Required:

3 years moving average:

Solution:

Years	Values	3 years moving total	3 years moving average
2001	102	-	-
2002	108	340	113.33
2003	130	378	126
2004	140	428	142.67
2005	158	478	159.33
2006	180	534	178
2007	196	586	195.33
2008	210	626	208.67
2009	220	660	220
2010	230	-	-

Q.13 Determine graphically for the following data and check your answer by using formulae.

(i) Median and Quartiles using cumulative frequency polygon.

(ii) Mode using Histogram. (A.B)

Given Data:

Class boundaries	Frequency
10 – 20	2
20 – 30	5
30 – 40	9
40 – 50	6
50 – 60	4
60 – 70	1

Required:

(i) Median and quartiles using cumulative frequency polygon and formula.

(ii) Mode using histogram and using formula.

Solution:

C – B	f	$C.f$	
0-10	0	0	
10 – 20	2	2	
20 – 30	5 f_1	7	Q_1 class
30 – 40	9 f_m	16	Median class/Modal class
40 – 50	6 f_2	22	Q_3 class
50 – 60	4	26	
60 – 70	1	27	
	27		

Median using Formula (K.B)

$$\text{Median} = \left(\frac{n}{2}\right)^{th} \text{ observation}$$

$$= \left(\frac{27}{2}\right)^{th} \text{ observation}$$

$$= 13.5^{th} \text{ observation}$$

$$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - C\right)$$

Putting the values

$$\text{Median} = 30 + \frac{10}{9}(13.5 - 7)$$

$$= 30 + \frac{10}{9}(6.5)$$

$$= 30 + 7.22$$

$$\Rightarrow \text{Median} = 37.22$$

Quartiles using Formula (K.B)

First Quartile

$$\mathbf{Q_1} = \frac{n}{4} \text{th observation}$$

$$= \frac{27}{4} \text{th observation}$$

$$= 6.75 \text{th observation}$$

$$Q_1 = l + \frac{h}{f}\left(\frac{n}{4} - C\right)$$

Putting the values

$$= 20 + \frac{10}{5}(6.75 - 2)$$

$$= 20 + 2(4.75)$$

$$= 20 + 9.50$$

$$\Rightarrow Q_1 = 29.50$$

Third Quartile (Q_3)

$$\mathbf{Q_3} = \frac{3n}{4} \text{th observation}$$

$$= 3 \times \frac{27}{4} \text{th observation}$$

$$= \frac{81}{4} \text{th observation}$$

$$= 20.25 \text{ th observation}$$

$$Q_3 = l + \frac{h}{f}\left(\frac{3n}{4} - C\right)$$

Putting the values

$$= 40 + \frac{10}{6}(20.25 - 16)$$

$$= 40 + \frac{10}{6}(4.25)$$

$$= 40 + 7.80$$

$$\Rightarrow Q_3 = 47.08$$

Mode (K.B)

$$\text{Mode} = l + \frac{f_m - f_1}{2f_m - f_1 - f_2} \times h$$

Putting the values

$$= 30 + \frac{9 - 5}{2(9) - 5 - 6} \times 10$$

$$= 30 + \frac{4}{18 - 11} \times 10$$

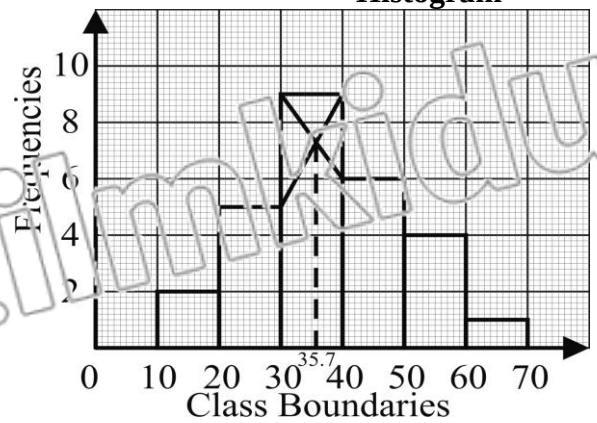
$$= 30 + \frac{40}{7}$$

$$= 30 + 5.71$$

$$\Rightarrow \text{Mode} = 35.71$$

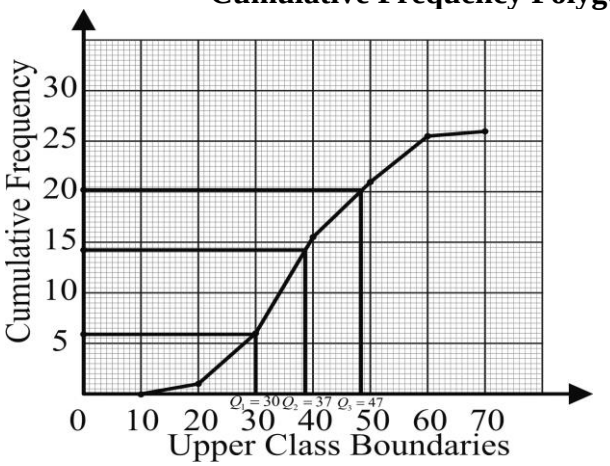
Mode Using Histogram

Histogram



Scale:
On X-axis
1 big box = 20
On Y-axis
1 big box =5

Cumulative Frequency Polygon



Scale:
On X-axis
1 big box = 20
On Y-axis
1 big box =5

DISPERSION (K.B)

Spread or scatter ness of the data is known as dispersion.

Types of Dispersion (K.B)

(i) Range
(ii) Variance
(iii) Standard deviation

Measure of Dispersion (K.B)

The techniques that are used to determine the degree or extent of variation in a data set is called measure of dispersion.

Range (K.B)

The difference between the largest and smallest value in a given data is called range.

Mathematically it is defined as:

Range (R) = $X_m - X_1$

X_m = The largest value

X_1 = The smallest value

Example:

Given

16, 5, 6, 10, 11

$X_m = 16$, $X_1 = 5$

Range = $X_m - X_1$

= 16 – 5

= 11

Variance (K.B)

A value obtained by dividing the sum of squares of deviations taken form arithmetic mean by the number of observations in the given data is called variance. It is denoted by S^2.

Formulae

(i) $S^2 = \dfrac{\sum\left(X - \bar{X}\right)^2}{n}$

(Proper mean formula)

(ii) $S^2 = \dfrac{\sum X^2}{n} - \left(\dfrac{\sum X}{n}\right)^2$

(Direct Formula)

Standard Deviation **(K.B)**

The positive square root of variance is called standard deviation. It is abbreviated as S.D and is denoted by S.

Formulae

(i) $S = \dfrac{\sqrt{\sum\left(X - \bar{X}\right)^2}}{n}$

(Proper mean formula)

(ii) $S = \sqrt{\dfrac{\sum X^2}{n} - \left(\dfrac{\sum X}{n}\right)^2}$

(Direct formula)

Example 5: (Page # 140) **(A.B)**

Compare the variation about mean for the two groups of students who obtained the following marks in statistics:

X = Marks (Section A)	Y = Marks (section B)
60	62
70	62
30	65
90	68
80	67
40	48

Solution:

In order to compare variation about mean we compute standard deviation for the two groups as follows:

X	Y	$X-\bar{X}$	$(X-\bar{X})^2$	$Y-\bar{Y}$	$(Y-\bar{Y})^2$
60	62	−2	4	0	0
70	62	8	64	0	0
30	65	−32	1024	3	9
90	68	28	784	6	36
80	67	18	324	5	25
40	48	−20	400	−14	196
$\sum x = 372$	$\sum y = 372$	0	$\sum(x-\bar{x})^2 = 2600$	0	$\sum(y-\bar{y})^2 = 266$

Mean for group $A = \bar{X} = \dfrac{\sum X}{n} = \dfrac{372}{6} = 62$

Mean for group $B = \bar{Y} = \dfrac{\sum Y}{n} = \dfrac{372}{6} = 62$

$$S.D(X) = \sqrt{\frac{\sum\left(X - \bar{X}\right)^2}{n}}$$

$$= \sqrt{\frac{2600}{6}} = \sqrt{433.333}\,.$$

$$= 20.82\text{gm}$$

$$S.D(Y) = \sqrt{\frac{\sum\left(Y - \bar{Y}\right)^2}{n}}$$

$$= \sqrt{\frac{266}{6}} = \sqrt{4.333}\,.$$

$$= 6.66\text{gm}$$

Comment: we note that the variation in Group B is smaller than that of Group A. This implies the marks of students in Group B are closer to their Mean than that of group A.

Note **(K.B + U.B)**

Smaller variation is more consistent in performance.

Exercise 6.3

Q.1 What do you understand by dispersion? **(K.B)**

Ans: See Definition page # 206

Q.2 How do you define measures of dispersion? **(K.B)**

Ans: See Definition page # 206

Q.3 Define Range, Standard deviation and Variance. **(K.B + A.B)**

Ans: See Definition page # 206, 207

Q.4 The salaries of five teachers in Rupees are as follows. 11500, 12400, 15000, 14500, 14800). Find the range and standard deviation. **(A.B)**

Given Data:

11500, 12400, 15000, 14500, 14800

Required

(i) Range **(ii)** Standard Derivation

Solution:

X	X^2
11500	132250000
12400	153760000
15000	225000000
14500	210250000
14800	219040000
$\sum X = 68200$	$\sum X^2 = 940300000$

Range **(A.B)**

Max. value = $X_m = 15{,}000$

Min. value = $X_o = 11{,}500$

Range = $X_m - X_o = 15000 - 11500$

= 3,500 Rs.

Standard Deviation **(A.B)**

$$S=\sqrt{\frac{\sum X^2}{n}-\left(\frac{\sum X}{n}\right)^2}$$

$$=\sqrt{\frac{940300000}{5}-\left(\frac{68200}{5}\right)^2}$$

$$=\sqrt{188060000-186049600}$$

$$=\sqrt{2010400}$$

$$\Rightarrow S = 1417.886$$

Q.5 a- Find the standard deviation "S" of each set of numbers: **(A.B)**

(i) 12, 6, 7, 3, 15, 10, 18, 5

(ii) 9, 3, 8, 8, 9, 8, 9, 18.

b- Calculate variance for the data: 10, 8, 9, 7, 5, 12, 8, 6, 8, 2.

Given Data:

(i) 12,6,7,3,15,10,18,5 **(FSD 2017)**

Required:

Standard Deviation = S =?

(a) Solution:

X	X^2
12	144
6	36
7	49
3	9
15	225
10	100
18	324
5	25
$\sum X = 76$	$\sum X^2 = 912$

Standard Deviation **(A.B)**

$$S = \sqrt{\frac{\sum X^2}{n}-\left(\frac{\sum X}{n}\right)^2}$$

$$=\sqrt{\frac{912}{8}-\left(\frac{76}{8}\right)^2}$$

$$= \sqrt{114 - (9.5)^2}$$
$$= \sqrt{114 - 90.25}$$
$$= \sqrt{23.75}$$
$$\Rightarrow S = 4.87$$

(ii) 9,3,8,8,9,8,9,18 **(FSD 2015)** **(A.B)**

X	$X - \overline{X} = X - 9$	$(X - \overline{X})^2$
9	0	0
3	-6	36
8	-1	1
8	-1	1
9	0	0
8	-1	1
9	0	0
18	9	81
$\sum X = 72$		**120**

Mean

$$\overline{X} = \frac{\sum X}{n}$$
$$= \frac{72}{8}$$
$$= 9$$

Standard Deviation **(A.B)**

$$S = \sqrt{\frac{\sum (X - \overline{X})^2}{n}}$$
$$= \sqrt{\frac{120}{8}} = \sqrt{15}$$
$$\Rightarrow S = 3.87$$

(b) **Given Data:** **(D.G.K 2015)** **(A.B)**

10, 8, 9, 7, 5, 12, 8, 6, 8, 2

Required

Variance = S^2

Solution

X	X^2
10	100
8	64
9	81
7	49
5	25
12	144
8	64
6	36
8	64
2	4
$\sum X = 75$	$\sum X^2 = 631$

Variance

$$S^2 = \frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2$$
$$= \frac{631}{10} - \left(\frac{75}{10}\right)^2$$
$$= 63.1 - 56.25$$
$$S^2 = 6.85$$

Result

Variance $= S^2 = 6.85$

The length of 32 items are given below. Find the mean length and standard deviation of the distribution.

Length	20-22	23-25	26-28	29-31	32-34
Frequency	3	6	12	9	2

Given Data:

Length	Frequency
20 – 22	3
23 – 25	6
26 – 28	12
29 – 31	9
32 – 34	2

Required

(i) Mean length **(ii)** Standard Deviation

Solution:

C – I	f	X	fX	fX^2
20-22	3	21	63	1,323
23-25	6	24	144	3,456
26-28	12	27	324	8,748
29-31	9	30	270	8,100
32-34	2	33	66	2,178
	$\sum f = 32$		$\sum fX = 867$	$\sum fX^2 = 23{,}805$

Mean

$$\bar{X} = \frac{\sum fX}{\sum f}$$

$$= \frac{867}{32}$$

$$\Rightarrow \bar{x} = 27.09$$

Standard Deviation

$$S = \sqrt{\frac{\sum fX^2}{n} - \left(\frac{\sum fX}{n}\right)^2}$$

$$= \sqrt{\frac{23805}{32} - \left(\frac{867}{32}\right)^2}$$

$$= \sqrt{743.906 - 733.87}$$

$$= \sqrt{9.8347} = 3.136$$

Q.6 Find the range of the given data: **(A.B)**

Class	33-40	41-50	51-60	61-70	71-75
Frequency	28	31	12	9	5

Solution:

Class	Frequency	C-B
33-40	28	32.5-40.5
41-50	31	40.5-50.5
51-60	12	50.5-60.5
61-70	9	60.5-70.5
71-75	5	70.5-75.5

Here

Range = upper class boundary of last class – lower class boundary of 1^{st} class

Range = 75.5 – 32.5 = 43

FORMULAE

(Exercise # 6.2,6.3)

(A.B + K.B + U.B)

For Ungrouped Data	For Grouped Data
Arithmetic means	
Ungroup Data (Direct Method)	**Grouped Data**
$\bar{X} = \frac{\sum x}{n}$	$\bar{X} = \frac{\sum fx}{\sum f}$
Indirect Method (Short Cut/ Deviation Method)	
$\bar{X} = A + \frac{\sum D}{n} \quad ;(D = x - A)$	$\bar{X} = A + \frac{\sum fD}{\sum f}$
(Coding Method)	
$A.M = A + \frac{\sum u}{n} \times h \quad ;\left(u = \frac{x-A}{h}\right)$	$A.M = A + \frac{\sum fu}{\sum f} \times h$
(Median)	
Median = $\left(\frac{n+1}{2}\right)^{th}$ item (n is odd)	Median $= l + \frac{h}{f}\left(\frac{n}{2} - C\right)$
Median $= \frac{1}{2}\left\{\left(\frac{n}{2}\right)^{th} \text{item} + \left(\frac{n+2}{2}\right)^{th} \text{item}\right\}$ (n is even)	

(Mode)	
Mode = most repeated value of the data	Mode = $l + \frac{(f_m - f_1)}{2fm - f_1 - f_2} \times h$
(Geometric Mean)	
G.M $= (x_1, x_2, x_3, \ldots x_n)^{\frac{1}{n}}$ By definition method	G.M $= Anti\log\left(\frac{\sum f \log X}{\sum f}\right)$
G.M $= Anti\log\left(\frac{\sum \log X}{n}\right)$ By log method	
Harmonic Mean	
H.M $= \frac{n}{\sum \frac{1}{X}}$	H.M $= \frac{\sum f}{\sum \frac{f}{X}}$
Weighted Arithmetic Mean $= \bar{X}_w = \frac{\sum wx}{\sum w}$	
Range	
Range = $X_m - X_1$	Range = upper class boundary of last class – lower class boundary of first class

<table>
<tr><th colspan="2">Variance</th></tr>
<tr><td>(i) $S^2 = \frac{\sum\left(X-\bar{X}\right)^2}{n}$
Proper mean formula
(ii) $S^2 = \frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2$
Direct Formula</td><td>(i) $S^2 = \frac{\sum f\left(X-\bar{X}\right)^2}{\sum f}$
Proper mean formula
(ii) $S^2 = \frac{\sum fX^2}{\sum f} - \left(\frac{\sum fX}{\sum f}\right)^2$
Direct Formula</td></tr>
<tr><th colspan="2">Standard Deviation</th></tr>
<tr><td>(i) $S = \sqrt{\frac{\sum\left(X-\bar{X}\right)^2}{n}}$
Proper mean formula
(ii) $S = \sqrt{\frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2}$
Direct Formula</td><td>(i) $S = \sqrt{\frac{\sum f\left(X-\bar{X}\right)^2}{\sum f}}$
Proper mean formula
(ii) $S = \sqrt{\frac{\sum fX^2}{\sum f} - \left(\frac{\sum fX}{\sum f}\right)^2}$
Direct Formula</td></tr>
</table>

Miscellaneous Exercise 6

Q.1 Multiple choice questions

Three possible answers are given for the following question. Tick (✓) the correct answer.

(i) A grouped frequency table is also called (SWL 2014, MTN 2015) **(K.B)**
(a) Data **(b)** Frequency Distribution **(c)** Frequency Polygon

(ii) A histogram is a set of adjacent (LHR 2014, 15, MTN 2015) **(K.B)**
(a) Squares **(b)** Rectangles **(c)** Circles

(iii) A frequency polygon is a many sided (LHR 2014) **(K.B)**
(a) Closed figure **(b)** Rectangle **(c)** Square

(iv) A cumulative frequency table is also called **(K.B)**
(a) Frequency distribution **(b)** Data **(c)** Less than cumulative frequency distribution

(v) In a cumulative frequency polygon frequencies are plotted against **(K.B)**
(a) Midpoints **(b)** Upper class boundaries **(c)** Class limits

(vi) Arithmetic mean is a measure that determines a value of the variable under study by dividing the sum of all values of the variable by their **(K.B)**
(a) Number **(b)** Group **(c)** Denominator

(vii) A deviation is defined as a difference of any value of the variable from a **(K.B)**
(a) Constant **(b)** Histogram **(c)** Sum

(viii) A data in the form of frequency distribution is called **(K.B)**
(a) Grouped data **(b)** Ungrouped data **(c)** Histogram

(ix) Mean of a variable with similar observations say constant k is **(K.B)**
(a) Negative **(b)** k itself **(c)** Zero

(x) Mean is affected by change in (LHR 2015, FSD 2014) **(K.B)**
(a) Value **(b)** Ratio **(c)** Origin

(xi) Mean is affected by change in **(K.B)**
(a) Place **(b)** Scale **(c)** Rate

(xii) Sum of the deviations of the variable X from its mean is always **(K.B)**
(FSD 2016, 18, D.G.K 2015)
(a) Zero **(b)** One **(c)** Same

(xiii) The n^{th} positive root of the product of the $x_1, x_2, x_3, ...x_n$ observations is called **(K.B)**
(a) Mode **(b)** Mean **(c)** Geometric mean

(xiv) The value obtained by reciprocating the mean of the reciprocal of $x_1, x_2, x_3, ..., x_n$ observations is called **(K.B)**
(a) Geometric mean **(b)** Median **(c)** Harmonic mean

(xv) The most frequent occurring observation in a data set is called **(K.B)**
(GRW 2014, FSD 2014, 15, SWL 2014)
(a) Mode **(b)** Median **(c)** Harmonic mean

(xvi) The measure which determines the middlemost observation in a data set is called **(K.B)**
(a) Median **(b)** Mode **(c)** Mean

(xvii) The observations that divide a data set into four equal parts are called (SWL 2015, 16) **(K.B)**
(a) Deciles **(b)** Quartiles **(c)** Percentiles

(xviii) The spread or scatterness of observations in a data set is called **(K.B)**
(FSD 2014, SGD 2014, D.G.K 2014)
(a) Average **(b)** Dispersion **(c)** Central tendency

(xix) The measures that are used to determine the degree or extent of variation in a data set are called measures of **(K.B)**
(a) Dispersion **(b)** Central tendency **(c)** Average

(xx) The extent of variation between two extreme observations of a data set is measured by (RWP 2015) **(K.B)**
(a) Average **(b)** Range **(c)** Quartiles

(xxi) The mean of the squared deviations of $x_i\ (i=1,2,...n)$ observations from their arithmetic mean is called **(K.B)**
(a) Variance **(b)** Standard deviation **(c)** Range

(xxii) The positive square root of mean of the squared deviations of $X_i\ (i=1,2,....n)$ observations from their arithmetic mean is called **(K.B)**
(a) Harmonic mean **(b)** Range **(c)** Standard deviation

ANSWER KEY

i	b	v	b	ix	b	xiii	c	xvii	b	xxi	a
ii	b	vi	a	x	c	xiv	c	xviii	b	xxii	c
iii	a	vii	a	xi	b	xv	a	xix	a		
iv	c	viii	a	xii	a	xvi	a	xx	b		

Q.2 Writhe short answers of the following questions.

(i) Define class limits. (RWP 2015, BWP 2016, D.G.K 2014) **(K.B)**
Ans: See definition page # 184

(ii) Define class mark. **(K.B)**
Ans: See definition page # 184

(iii) What is cumulative frequency? (GRW 2014, BWP 2011, 14, RWP 2016, SGD 2014, 16) **(K.B)**
Ans: See definition page # 185

(iv) Define a frequency distribution. (BWP 2015, SWL 2016, MTN 2016) **(K.B)**
Ans: See definition page # 184

(v) What is Histogram? (GRW 2016, FSD 2014, SWL 2017, SGD 2015, RWP 2014) **(K.B)**
Ans: See definition page # 185

(vi) Name two measures of central tendency. **(K.B)** (LHR 2014, GRW 2016, SWL 2015, D.G.K 2015)
Ans: See definitiDEon page # 196

(vii) Define Arithmetic mean. (LHR 2016, GRW 2016, FSD 2015, SGD 2015, D.G.K 2014) **(K.B)**
Ans: See definition page # 188

(viii) Write three properties of Arithmetic mean. **(K.B)**
(LHR 2016, 17, GRW 2014, FSD 2014, 15, 17, SWL 2016, RWP 2014, BWP 2014, SGD 2015, MTN 2017)
Ans: See definition page # 189

(ix) Define Median. (LHR 2014, BWP 2015, SGD 2016, RWP 2017, MTN 2017) **(K.B)**
Ans: See definition page # 188

(x) Define Mode? **(K.B)**
(LHR 2014, 17, GRW 2017, SWL 2015, 16, BWP 2014, 17, MTN 2015, 16, 17, D.G.K 2017)
Ans: See definition page # 189

(xi) What do you mean by Harmonic mean? **(K.B)** (GRW 2015, SWL 2015, MTN 2015, SGD 2015)
Ans: See definition page # 189

(xii) Define Geometric mean (LHR 2015, FSD 2015, SWL 2016, D.G.K 2016) **(K.B)**
Ans: See definition page # 189

(xiii) What is Range? (SWL 2014) **(K.B)**
Ans: See definition page # 206

(xiv) Define Standard deviation. **(K.B)**
(LHR 2016, FSD 2015, SWL 2014, 15, BWP 2017, MTN 2015, 17, SGD 2014, 17, D.G.K 2014)
Ans: See definition page # 207

SELF TEST

Time: 40 min **Marks: 25**

Q.1 Four possible answers (A), (B), (C) & (D) to each question are given, mark the correct answer. (7×1=7)

1 The value obtained by reciprocating the mean of the reciprocal of $x_1, x_2, x_3, \ldots\ldots\ldots\ldots, x_n$ observations is called:

(A) Geometric mean **(B)** Median
(C) Harmonic mean **(D)** Mode

2 If $D = x - 10$, $\sum D = 50$ and $n = 5$ then arithmetic mean is:

(A) 10 **(B)** 20
(C) 50 **(D)** 15

3 In 4,3,2,1,5,6,7 the mode is:

(A) 0 **(B)** 4
(C) 7 **(D)** No mode

4 If $\sum\left(x - \bar{x}\right)^2 = 40$, $n = 5$ then standard deviation is:

(A) 8 **(B)** 2.83
(C) 200 **(D)** 14.1

5 In a cumulative frequency polygon frequencies are plotted against:

(A) Midpoints **(B)** Upper class boundaries
(C) Class limits **(D)** Lower class boundaries

6 A histogram is a set of adjacent:

(A) Squares **(B)** Rectangles
(C) Circles **(D)** Closed figure

7 The measure which determines the middle most observation in a data set is called:

(A) Mean **(B)** Mode
(C) Median **(D)** None

Q.2 Give Short Answers to following Questions. (5×2=10)

(i) Find the standard deviation for the data: 12,6,7,3,15,10,18,5.

(ii) Find arithmetic mean by indirect method for the set of data: 200,225,350,375,270,320,290.

(iii) Find the geometric mean of 2, 4, 8 using logarithmic formula.

(iv) Find median.

Class Mark (X)	5	10	15	20	25	30
Frequency (f)	2	12	25	32	14	5

(v) On a vacation trip a family bought 21.3 liters of petrol at 39.90 rupees per liter, 18.7 liters at 42.90 rupees per liter, and 23.5 liters at 40.90 rupees per liter. Find the mean price paid per liter.

Q.3 Answer the following Questions. (4+4=8)

(a) The length of 32 items are given below, find the mean length and standard deviation of the distribution.

Length	20 – 22	23 – 25	26 – 28	29 – 31	32 – 34
Frequency	3	6	12	9	2

(b) On a prize distribution day, 50 students brought pocket money as under. Find mode.

Rupees	5-10	10-15	15-20	20-25	25-30
Frequency (f)	12	9	18	7	4

NOTE: Parents or guardians can conduct this test in their supervision in order to check the skill of students.