

Classification

1) Individual = X

2) Discrete =

3) Continuous =

1) Mean

↑

Direct Shortcut Step-deviation

1) Ind.

i) DM

X

10

20

30

40

50

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

2) S-C

A = mid value

$$X | d = X - A$$

$$\bar{X} = A + \frac{\sum d}{N}, N = \sum f$$

3) S-D

C = common

$$X | d | d' = \frac{d}{C}$$

$$\bar{X} = A + \left[\frac{\sum d'}{N} \times C \right]$$

2) Discrete

1) DM

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

Mean
variation
Variance
Standard Deviation

2) S-C

$$\bar{X} = A + \frac{\sum fd}{N}$$

3) S-D

$$\bar{X} = A + \left[\frac{\sum fd'}{N} \times C \right]$$

~~in~~
Dis
 $\left(\frac{N+1}{2} \right)^{th}$

3) Continuous

$$MP = \frac{DL+LL}{2}$$

cm

CI f X

0-10

Median

1) IS

N

Odd

$$\text{Med} = \left(\frac{N+1}{2} \right)^{th}$$

4 2 3 9

Even

$$\text{Med} = \frac{\left(\frac{N}{2} \right)^{th} + \left(\frac{N}{2} + 1 \right)^{th}}{2}$$

2) DS

$$\text{Med} = \left(\frac{N+1}{2} \right)^{th}$$

3) CS

$$\text{Med} = \left(\frac{N}{2} \right)^{th}$$

$$\text{Med} = L + \frac{\left(\frac{N}{2} - cf \right)}{f} x_i$$

202

Examp 17

Correction factor = $\frac{\text{lower limit of the 2nd class} - \text{Upper limit of the 1st class}}{2}$

Weekly wages	No. of workers.
800 - 899	5
<u>900 - 999</u>	10
1000 - 1099	15
1100 - 1199	8
1200 - 1299	2

$$\frac{900 - 899}{2} = 0.5$$

$$L \rightarrow 800 + 0.5 = 899.5 \\ U \rightarrow 899 + 0.5 = 900.5$$

~~Median for individual data points~~



Mean Deviation using mean

1) Ind S

$$X \mid |x - \bar{x}| \quad \bar{x} = \frac{\sum x}{N}$$

$$MD = \frac{\sum |x - \bar{x}|}{N}$$

2) D f C

$$X \mid f \mid |x - \bar{x}| \mid f|x - \bar{x}| \mid \bar{x} = \frac{\sum fx}{N}$$

$$MD = \frac{\sum f|x - \bar{x}|}{N}$$

Individual Series

short-cut Method.

(a) $X : 125, 128, 132, 135, 140, 148, 155, 157, 159, 161$

Sol : $\bar{x} = 144$

step - Deviation Method.

$125, 128, 132, 135, 140, 148, 155, 157, 159, 161$

(b) $X : 850, 700, 840, 750, 500, 800, 420, 2500$

(Income in R.) $2300, 1500$

Sol : $\bar{x} = 1,116$.

Family: A B C D E F G H I J

Monthly Income of 10 families are given below:

(b) The coins are tossed 1024 times. The theoretical frequency of 10 Heads ~~upto~~ are given below:

No. of Heads	Frequency.	
0	1	
1	10	
2	45	short-cut
3	120	Method.
4	210	
5	252	
6	210	
7	120	
8	45	
9	10	
10	1	

Calculate the mean number of heads per tossing.

$$\bar{x} = 5 \text{ heads.}$$

Continuous Series.

(a) Production in Tones: 10-20 20-30 30-40 40-50 50-60
 No. of factories: 5 4 7 12 10

60-70 70-80
 8 4

short-cut
Method.

$$\bar{x} = 46.6$$

(b) Production in tones	No. of factories.
5 - 9	15
10 - 14	14
15 - 19	17
20 - 24	22
25 - 29	8
30 - 34	7
35 - 39	11

Step-
Deviation
Method.

Sol: $\bar{X} = 20.51 \text{ tonnes}$.

(c)	<u>Marks</u>	<u>No. of Persons</u>	
	0 - 10	5	
	10 - 20	10	Direct Method.
	20 - 30	20	
	30 - 40	40	$\bar{X} = 39.39$.
	40 - 50	30	
	50 - 60	20	
	60 - 70	10	
	70 - 80	4	

MEDIAN

INDIVIDUAL SERIES

1.. Determine the median from the following:

25 15 23 40 27 25 23 25 20.

Median is the size of $\left(\frac{N+1}{2}\right)^{\text{th}}$ item.

$$\text{Mean} = \frac{10}{2} = 5^{\text{th}} \text{ item} = 25.$$

2.. Roll No.	Marks Obtained.
1	43
2	48
3	65
4	57
5	31
6	60
7	37
8	48
9	78
10	59

Calculate the median.

Median is the size of $\left(\frac{N+1}{2}\right)^{\text{th}}$ item

$$\Rightarrow \left(\frac{10+1}{2}\right)^{\text{th}} \text{ item} = 5.5^{\text{th}} \text{ item}$$

5.5 is between 5th item and 6th item.

$$\Rightarrow \frac{48+57}{2} = \frac{105}{2} = 52.5$$

Q3.. Discrete Series

FORMULA \Rightarrow Median is the size of $\left(\frac{N+1}{2}\right)^{\text{th}}$ item.

Q3. Calculate the median for the following data

size of the item : 4 6 8 10 12 14 16
 Frequency : 2 4 5 3 2 1 4

size of the item	frequency	Cumulative Frequency.
4	2	2
6	4	6
8	5	11
10	3	14
12	2	16
14	1	17
16	4	21

Median is the size of $\left(\frac{N+1}{2}\right)^{\text{th}}$ item = $\frac{21+1}{2}$

$$= \frac{22}{2} = 11^{\text{th}} \text{ item}$$

$$\text{Median} = 8$$

Q4. Calculate Median from the following data:

Marks : 50 40 30 20 10

Frequency : 10 40 20 12 16

$$\text{Median} = 40.$$

Q5: Mid value: 35 45 55 65 75 85
 frequency: 6 25 18 32 15 4

Class (x)	f	cf
30 - 40	6	6
40 - 50	25	31
50 - 60	18	49
60 - 70	32	81
70 - 80	15	96
80 - 90	4	100

$$M = 60.325$$

Continuous SERIES

FORMULA:

$$M = l + \frac{N/2 - cf}{f} \times c$$

M = Median

l = Lower limit of the median class

cf = Cumulative frequency of the class preceding the median class.

c = Size of class interval.

f = Frequency of the median class.

Q6. Calculate the Median of the following distribution.

Length (in inches)	No. of units
0 - 20	1
20 - 40	14
40 - 60	35
60 - 80	85
80 - 100	90
100 - 120	15

Length (in inches)	No. of units (f)	(cf.)
0 - 20	1	1
20 - 40	14	15
40 - 60	35	50
60 - 80	85	135
80 - 100	90	225
100 - 120	15	240

$$M = \frac{N}{2} = \frac{240}{2} = 120^{\text{th}} \text{ item, that is } 60-80 \text{ group}$$

$$M = l + \frac{N/2 - cf}{f} \times c$$

$$= 60 + \frac{120 - 50}{85} \times 20$$

$$= 76.47 \text{ inches.}$$

Q7. Calculate the median mark from the following frequency distribution.

Marks | No. of students

0 - 10	5
10 - 20	8
20 - 30	7
30 - 40	12
40 - 50	28
50 - 60	20
60 - 70	10

Sol: Median = 44.64 marks.

Q8. Calculate median for the following data:

Wages (x)	1 - 5	6 - 10	11 - 15	16 - 20	21 - 25	26 - 30	31 - 35
f	5	7	9	15	12	10	9

36 - 40	41 - 45	46 - 50
8	4	2

(19.67)

Sol: Median: 22.375

Q9. Calculate Median.

Mid value: 115 125 135 145 155 165

frequency: 6 25 48 72 60 19

Sol: M = 145

Q9. Obtain the value of median from the following data of the monthly income of 10 employees of a company in Rs:

4391 5384 5591 5407 6672 6522
 6777 6753 7850 7490

Sol: Median = 6597.

Q10. From the following data, find the value of median:

Income (Rs.): 4000 4500 5800 5060 6600 5380
 No. of Persons: 24 26 16 20 6 30

Sol: Rs 5060

Q11. Calculate the median for the following frequency distribution:

Marks	No. of students	Marks	No. of students.
45 - 50	10	10 - 15	15
40 - 45	15	5 - 10	7
35 - 40	26		
30 - 35	30		
25 - 30	42		
20 - 25	31		
15 - 20	24		

Continuous Series

Q1. Daily wages in Rs (X): 20-25 25-30 30-35 35-40 40-45

No. of workers (f): 1 3 8 12 7
45-50 5

$$Z = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times c \quad Z = 37.22$$

l = lower limit

Z = Mode

f_1 = frequency of the modal class

f_0 = frequency of the class preceding the modal class.

f_2 = frequency of the class succeeding the modal class.

c = size of class-interval.

Q2. Calculate Mode

X : 10-19	20-29	30-39	40-49	50-59	60-69
f : 5	12	22	25	14	10

X : 70-79	80-89
f : 8	4

Sol: $Z = 41.64$

Sol: Median = 27.74

Q12. Calculate the median from the following data

Height	No. of Apples	Height	No. of Apples
410 - 419	14	450 - 459	45
420 - 429	20	460 - 469	18
430 - 439	42	470 - 479	7
440 - 449	54		

Sol: Median \rightarrow 443.94

MODE:

Q1. Local Mode in the data 7, 12, 8, 5, 9, 6, 10, 9, 4, 9, 9
A: On inspection, $x_2 = 9$

Discrete series.

Mode:

Variable	Frequency
28	10
30	12
32	16
34	14
36	10
38	8
40	17
42	5
44	4

$Z = 40$ (Mode)

Q2. Calculate Mean and Median from the following data:

x :	10	15	20	25	30	35	40	45
f :	10	25	40	50	25	8	17	38

$$\text{Sol: } \bar{x} = 27.91$$

$$M = 25$$

Q3. Calculate the mean, mode, and Median from the data given below:

Weight (in Kgs):	58	60	61	62	63	64	65	66
No. of Persons:	4	12	24	32	32	16	8	2

$$\text{Sol: } \bar{x} = 62.25$$

$$M = 62$$

$$\text{Mode: } Z_2 = 61.5 \quad Z_2 = 63$$

Q4. Find Median and Mode.

Marks:	5-15	15-25	25-35	35-45
--------	------	-------	-------	-------

Students:	7	12	17	29
-----------	---	----	----	----

45-55	55-65	65-75
31	16	3

$$\text{Sol: Median} = 42.4$$

$$\bar{x} = 46.17$$

Q3. Midpoint Frequency.

120	5
130	12
140	32
150	40
160	20
170	10
180	9
190	11

Sol: Class interval of the distribution is 10.

115 - 125	5
125 - 135	12
135 - 145	32
145 - 155	40
155 - 165	20
165 - 175	10
175 - 185	9
185 - 195	11

$$\text{Sol: } Z = 147.85$$

Q3. Calculate Mean, Median and Mode from the following data.

Roll No: 1 2 3 4 5 6 7 8

Age : 21 18 19 21 20 17 22 23

$$\text{Sol: } \bar{X} = 20.12$$

$$M = 20.5$$

$$Z = 21.1$$

17 18 19 20 21 21 22 23

$$\frac{20+21}{2} = \frac{41}{2} = 20.5$$

Measures of Dispersion.

Dispersion is the measure of the variation of the items. In other words dispersion measures the extent to which the items vary from some central value.

An average is more meaningful when it is examined in the light of dispersion.

SERIES A	SERIES B	SERIES C.
100	100	1
100	105	489
100	102	2
100	103	3
100	90	5
Total	500	500
\bar{x}	100	100

SIGNIFICANCE OF MEASURING VARIATION.

1. To determine the reliability of an average.
2. To serve as a basis for the control of the variability.
3. To compare two or more series with regard to their variability.
4. To facilitate the use of other statistical measures.

Q5. Calculate mean, Median and Mode

Profit : (in lakhs)	4-7	8-11	12-15	16-19	20-23	24-27
No. of companies :	6	10	18	30	15	12
	28-31	32-35	36-39			
	10	6	2			
	$\bar{x} = 19.22$					
	Median = 18.23					
	Mode = 17.27					

✓ Q6. Compute the Mean and Median from the following data:

Mid-value:	115	125	135	145	155	165	175	185	195
Frequency:	6	25	48	72	116	60	38	22	3

$$\text{Sol: } \bar{x} = 153.64$$

$$\text{Median} = 153.79$$

Methods of studying Variation

1. Range.
2. Mean Deviation. \bar{D} - Variance.
3. Standard Deviation

$$\text{Range} = L - S$$

L = Largest item

S = Smallest item.

Range \downarrow Dispersion \downarrow .

MEAN DEVIATION

FOR INDIVIDUAL OBSERVATIONS.

$$M.D = \frac{1}{N} \sum |x_i - \bar{x}|$$

$$D = |x_i - \bar{x}|$$

$$MD = \frac{\sum |x_i - \bar{x}|}{N}$$

FOR DISCRETE AND CONTINUOUS SERIES.

$$M.D = \frac{1}{N} \sum f_i |x_i - \bar{x}|$$

$$D = |x_i - \bar{x}|$$

Q1. INDIVIDUAL OBSERVATIONS

Mean Deviation from Mean.

x : 68 49 32 21 54 38 59 66 41

$$MD = 12.94$$

Q2. Months : Jan Feb Mar Apr May June
Earnings : 1198 1200 1200 1202 1202 1204

Months : July Aug Sept Oct Nov Dec
Earnings : 1204 1206 1206 1208 1208 1210

Sol:

$$MD = 3.$$

MD in Discrete series

X	f
35	3
40	8
45	12
50	9
55	4
60	7
65	15
70	5
75	10
80	7
85	5
90	3
95	2

$$MD = 13.514$$

MD in Continuous series

Q1	CI.	f	MD
	10 - 20	5	
	20 - 30	4	
	30 - 40	7	
	40 - 50	12	
	50 - 60	10	
	60 - 70	8	
	70 - 80	4	

Q2	Marks obtained	No. of students
cheating	10 - 12	5
	12 - 14	8
	14 - 16	12
	16 - 18	20
	18 - 20	7
	20 - 22	10
	22 - 24	11

$$\begin{array}{ll} 24-26 & 15 \\ 26-28 & 10 \end{array}$$

$$MD = 4.22$$

Mean Deviations from Median.
for Individual series

$$MD =$$

$$\frac{\sum |D|}{N}$$

$$|D| = |x - \text{Median}|$$

$$\textcircled{Q1. } (x): 20 \ 25 \ 30 \ 38 \ 42 \ 55 \ 66 \ 67 \ 70 \\ \text{Median} = 42 \\ MD = 16.11$$

$$\textcircled{Q2. } x: 59 \ 32 \ 67 \ 43 \ 22 \ 17 \ 64 \ 55 \ 47 \ 80 \ 25 \\ \text{Median} = 47 \\ MD = 16.91$$

Discrete series

$$\textcircled{Q1. } x: 125 \ 126 \ 130 \ 132 \ 134 \ 138 \ 140 \ 141 \\ f: 8 \ 4 \ 3 \ 10 \ 7 \ 9 \ 4 \ 2 \\ \text{Median} = 132$$

Mean Deviation from Median

FORMULA: $MD = \frac{\sum f |D|}{N}$

$$D = |x - \text{Median}|$$

$$\rightarrow MD = 4.34$$

$$Q2. \quad X: 10 \ 20 \ 30 \ 40 \ 50 \ 60 \\ f: 4 \ 7 \ 15 \ 8 \ 7 \ 2 \quad \text{Median} = 22$$

$$MD = 10$$

Continuous Series

$$MD = \frac{\sum f |D|}{N}$$

$$D = |x - \text{Median}|$$

$$Q3: \quad X: 0-10 \ 10-20 \ 20-30 \ 30-40 \ 40-50 \\ f: 8 \ 30 \ 40 \ 12 \ 10 \quad \text{Median} = 23$$

$$MD = 8.28$$

Marks	No. of students
10-14	12
15-19	29
20-24	59
25-29	75
30-34	46
35-39	19

$$\text{Median} = 25.83 \\ = 26.$$

$$MD = 5.10$$

For INDIVIDUAL SERIES

MEDIAN FORMULA.

$$\text{MEDIAN}(M) = \frac{\text{size of} \left(\frac{N}{2} \right)^{\text{th}} \text{item} + \text{size of} \left[\frac{N}{2} + 1 \right]^{\text{th}} \text{item}}{2}$$

4, 1, 2, 6, 10, 8, 11, 3, 9, 7.

$$\text{Median} = \frac{6+7}{2} = \frac{13}{2} = 6.5.$$

• Variance

1) Individual Series

$$V = \frac{\sum (x - \bar{x})^2}{N-1}$$

2) Discrete / Continuous Series

$$\sigma = \sqrt{\text{Variance}}$$

$$V = \frac{\sum f(x - \bar{x})^2}{N-1}$$

Standard Deviation

$$\sigma = \sqrt{\text{Variance}}$$

Variance estimates how far a set of numbers (are) spread out from their mean value.

Q1. Find the variance for the following set of data representing trees heights in feet: 3, 21, 98, 203, 17, 9

$$\text{Variance} = 6219.9$$

$$\sigma = 78.86$$

Q2. Calculate the variance for the following data:

Class Interval	Frequency
200 - 201	13
201 - 202	27
202 - 203	18
203 - 204	10
204 - 205	1
205 - 206	1

$$\text{Variance} = 1.179$$

Q3. Find the variance and standard deviation of all the possibilities of rolling a die.

Sol: All possible outcomes of rolling a die are $\{1, 2, 3, 4, 5, 6\}$
 This data set has six values ($n = 6$)

$$\text{Variance} = 2.917 \times 3.5 \quad \text{needs checking}$$

$$\sigma = 1.708 \quad 1.87$$

Q4. Find the variance and standard deviation of all the even numbers less than 10.

Even nos. less than 10 are

$$\{0, 2, 4, 6, 8\} \quad \text{needs checking.}$$

Data set has 5 values

$$\text{Variance} = 8 \times 10$$

$$\sigma = 2.828$$

Q5 Data: 5, 8, 12, 7, 10

Find variance and standard deviation

$$\bar{x} = \frac{5+8+12+7+10}{5} = \frac{42}{5} = 8.4$$

$$X \quad (x - \bar{x})^2$$

$$5 \quad 11.56$$

$$8 \quad 0.16$$

$$12 \quad 13.44$$

$$7 \quad 1.96$$

$$10 \quad \frac{2.56}{29.68}$$

$$\text{Variance} = \frac{29.68}{4} = 7.42$$

$$\sigma = 2.73$$

Q6. Data: 2, 6, 4, 9, 10

Find variance and standard deviation

$$\bar{x} = 6.2$$

$$\text{Variance} = 11.2$$

$$s = 3.35$$

Q7. Data: 3, 5, 7, 8, 9

$$\bar{x} = 6.4$$

$$\text{Variance} = 5.8$$

$$\sigma = 2.41$$

Q8. Calculate Variance and standard deviation.

Marks (x)

frequency (f)

5

3

10

5

15

7

20

6

25

4

$$\text{Variance} = 38.6$$

$$\sigma = 6.21$$

Q9. The frequency distribution of the no. of hours worked per day is as follows:

Hours (x) : Frequency

2	4
4	6
6	8
8	5
10	2

$$\text{Mean} = 5.6$$

$$\text{Variance} = 5.4$$

$$\sigma = 2.32$$

Q10. A factory records the number of defects in 10 products. The no. of defects is as follows:

Defects (x) Frequency (f)

1	2
2	3
3	4
4	

$$\text{Mean} = 2.4$$

$$\text{Variance} = 0.84$$

$$\sigma = 0.92$$

Q11. The following frequency distribution shows the age of individuals in a group

Age (x)	frequency (f)
10 - 20	5
20 - 30	8
30 - 40	12
40 - 50	7
50 - 60	3
60 - 70	

$$\bar{x} = 33.57$$

$$\text{Variance} = 130.22$$

$$\sigma = 11.41$$

Q12. The following table shows the distribution of marks of students

Marks (x)	frequency (f)
0 - 10	4
10 - 20	6
20 - 30	10
30 - 40	5

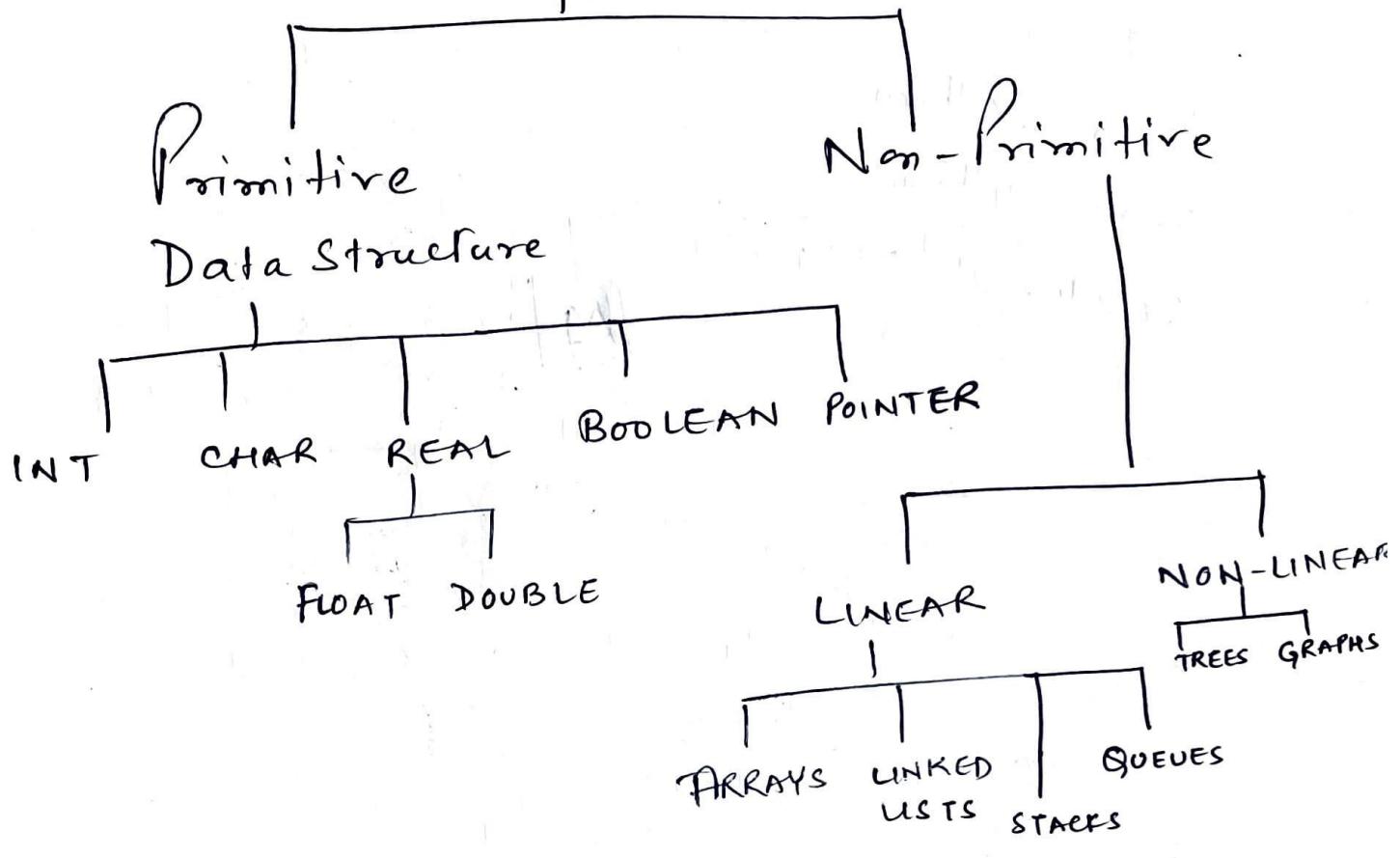
Find the variance and standard deviation

$$\text{Mean} = 21.4$$

$$\text{Variance} = 95.57$$

$$\sigma = 9.78$$

Data Structures



GEOMETRIC MEAN.

FOR INDIVIDUAL SERIES

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

FOR DISCRETE AND CONTINUOUS SERIES

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

Q Daily Income of ten families of a particular place is given below.

85, 70, 15, 500, 8, 45, 250, 95, 40, 36

X	$\log X$
85	1. 9294
70	1. 8451
15	1. 1761
75	1. 8751
500	2. 6990
8	0. 9031
45	1. 6532
250	2. 3979
40	1. 6021
36	1. 5563
	<hr/>
	17. 6373

$$\begin{aligned}
 G.M &= AL \left(\frac{17.6373}{N} \right) \\
 &= AL \left(\frac{17.6373}{10} \right) \\
 &= AL(1.7637) \\
 &= 5794 \\
 &\quad + 9 \\
 &= \underline{\underline{58.03}} \quad \checkmark
 \end{aligned}$$

Q2. calculate geometric mean from the following data :

125 1462 38 7 0.22 0.08 12.75 0.5

$$G.M = 6.952$$

Find the geometric mean for the data.

Marks : 4-8 8-12 12-16 16-20 20-24 24-28 28-32

Frequency : 6 10 18 30 115 12 10

32-36 36-40
6 2

Marks	(X)	f	$\log X$	$f \times \log X$
4-8	6	6	0.7782	4.6692
8-12	10	10	1.0000	10.0000
12-16	14	18	1.1461	20.6298
16-20	18	30	1.2553	37.6590
20-24	22	15	1.3424	20.1360
24-28	26	12	1.4150	16.9800
28-32	30	10	1.4771	14.7710
32-36	34	6	1.5315	9.1890
36-40	38	2	1.5798	<u>3.1596</u>
N	<u>109</u>			137.1936

$$G.M = \text{Antilog} \left(\frac{137.1936}{109} \right)$$

$$= \text{Antilog}(1.2587)$$

$$= 18.14$$

17.219

HARMONIC MEAN

FOR INDIVIDUAL SERIES

$$HM = \frac{N}{\sum \frac{1}{X}}$$

Q3. Compute Harmonic Mean for the following data

X. Marks	f. No. of students	f/x
10	20	2.0000
20	30	1.5000
25	50	2.0000
40	15	0.3750
50	5	0.1000
$N = \frac{120}{\sum f/x}$		5.9750

$$HM = \frac{N}{\sum f/x} = \frac{120}{5.9750} = 20.08$$

Q4. Calculate Harmonic Mean for the following data.

CI : 10 - 20	20 - 30	30 - 40	40 - 50	50 - 60
frequency : 4	6	10	7	3
HM = 29.88				

σ_m^2

Q From the following data.

X	R ₁	Y
36	7	50
56	3	35
20	9	70
65	1	25
42	6	58
33	8	75
44	5	60
50	4	45
15	10	80
60	2	38

$$\gamma = -0.93$$

X	Y
80	123
91	135
99	154
71	110
61	105
81	134
70	121
59	106

$$\gamma = 0.9524$$

Q	X	Y
60	73	
34	32	
40	34	
50	40	
45	45	
41	33	
22	12	$\{ = 0.8 \}$
43	30	
42	36	

Questions on Probability & Module-3.

Q1. A ball is drawn at random from a box containing 6 red balls, 4 white balls and 5 blue balls. Find the probability that the ball drawn is not red.

$$\text{Ans} \Rightarrow \text{Total number of balls} = 6 + 4 + 5 \\ = 15$$

$$\text{Probability of drawing a red ball} = \frac{6}{15}$$

$$\text{Probability of not red} = 1 - \frac{6}{15} \\ = \frac{9}{15}.$$