15.7.24

Measures of Central Tendency:

Different types of central tendency are:

1. (i) Arithmetic Mean ***
2. (ii) Median ***
3. (iii) Mode ***
4. (iv) Geometric Mean
5. (v) Harmonic Mean

Arithmetic Mean: 2 types of data.

i) Ungrouped data: If $x_1, x_2, \ldots, x_n$ represents the values of N items or observations, the arithmetic mean denoted by $\bar{x}$, is defined by $\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{N}$

$$= \frac{\sum_{i=1}^{N} x_i}{N}$$

ii) Grouped data: If $x_1, x_2, \ldots x_n$ represents the values of N items or observations with corresponding frequencies $f_1, f_2, \ldots f_n$, the arithmetic mean denoted by $\bar{x}$, is defined by $\bar{x} = \frac{f_1x_1 + f_2x_2 + \ldots + f_nx_n}{N} = \frac{\sum_{i=1}^{n} f_ix_i}{N}$

where, $N = f_1 + f_2 + \ldots + f_n = \sum_{i=1}^{n} f_i$

- **Example 1:** The monthly income of 10 employees working in a firm is as follows:

4487 4493 4502 4446 4475 4492
4572 4516 4468 4489

Find the average monthly income.

Solution: The total income

$$\sum x_i = 4487 + 4493 + 4502 + 4446 + 4475 + 4492 + 4572 + 4516 + 4468 + 4489$$

$$= 44,940$$

We know,

$$\text{average}, \bar{x} = \frac{\sum_{i=1}^{10} x_i}{N} = \frac{\sum_{i=1}^{10} x_i}{10} = \frac{44940}{10} = 4494$$

Hence the average monthly income is tk 4494.

- **Example 2:** Find the mean of the following data:

Class	8	10	15	20
frequency	5	8	8	4

Solution:

class (x_i)	frequency (f_i)	$f_i x_i$
8	5	40
10	8	80
15	8	120
20	4	80
	$\sum f_i = 25$	$\sum f_i x_i = 320$

We know,

average, $\bar{x} = \dfrac{\sum_{i=1}^{4} f_i x_i}{\sum_{i=1}^{4} f_i} = \dfrac{\sum f_i x_i}{N} = \dfrac{320}{25} = 12.8$

(ii) Using Shortcut Method:

$$\bar{x} = A + \frac{\sum_{i=1}^{n} f_i d_i}{N} \times h$$

Where,

h = The size of class interval

A = The assumed mean (middle number of the mid values) $\frac{N}{2}$

$d_i = \dfrac{x_i - A}{h}$ = Step deviation from A

x_i = Mid values of each class

N = The total frequency

- Example 1: Calculate mean for the following grouped data using short-cut method.

class	0-10	10-20	20-30	30-40	40-50
frequency	7	8	20	10	5

Solution:

Let, $A = 25$

Here, $h = 10$

We construct the following table:

class	Mid value x_i	frequency f_i	$d_i = \frac{x_i - A}{h}$	$f_i d_i$
0-10	5	7	-2	-14
10-20	15	8	-1	-8
20-30	25 → A	20	0	0
30-40	35	10	1	10
40-50	45	5	2	10
		$N = 50$		$\sum_{i=1}^{5} f_i d_i = -2$

We know,

$$\text{mean}, \bar{x} = A + \frac{\sum_{i=1}^{n} f_i d_i}{N} \times h$$

$$= 25 + \frac{-2}{50} \times 10$$

$$= 24.6$$

Example 2: Calculate mean for the following data representing the marks of statistics for 80 students in a class.

Marks	0-20	20-40	40-60	60-80	80-100	100-120	120-140
No. of Student	4	26	22	10	9	6	3

Solution: Let, $A = 70$

Here, $h = 20$

We construct the following table:

Marks	Mid value x_i	frequency f_i	$d_i = \frac{x_i - A}{h}$	$f_i d_i$
0-20	10	4	-3	-12
20-40	30	26	-2	-52
40-60	50	22	-1	-22
60-80	70 → A	10	0	0
80-100	90	9	1	9
100-120	110	6	2	12
120-140	130	3	3	9
		$N = 80$		$\sum f_i d_i = -56$

We know,

$$\text{Mean}, \bar{x} = A + \frac{\sum_{i=1}^{n} f_i d_i}{N} \times h$$

$$= 70 + \frac{-56}{80} \times 20$$

$$= 56$$

For practice

1. Calculate the mean of the following data

Height (cm)	65	66	67	68	69	70	71	72	73
No of plants	1	4	5	7	11	10	6	4	2

Solution: We construct the following table:

Height (x_i)	No of plants (f_i)	$f_i x_i$
65	1	65
66	4	264
67	5	335
68	7	476
69	11	759
70	10	700
71	6	426
72	4	288
73	2	146
	$\sum f_i = 50$	$\sum f_i x_i = \cancel{3469}$ 3459

We know,

$$\text{average}, \bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{\sum f_i x_i}{N} = \frac{3459}{50} = \cancel{68.98}\ 69.18$$

2. Find the mean of the following data

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
No. of Students	3	5	7	10	12	15	12	6	2	8

Solution: Let, $A = \cancel{12}\ 45$

Here, $h = 10$

We construct the following table:

Marks	Mid value x_i	No of students f_i	$d_i = \frac{x_i - A}{h}$	$f_i d_i$
0-10	5	3	-4	-12
10-20	15	5	-3	-15
20-30	25	7	-2	-14
30-40	35	10	-1	-10
40-50	45→A	12	0	0
50-60	55	15	1	15
60-70	65	12	2	24
70-80	75	6	3	18
80-90	85	2	4	8
90-100	95	8	5	40
		$N = 80$		$\sum f_i d_i = 44$

We know,

$$\text{mean}, \bar{x} = A + \frac{\sum_{i=1}^{n} f_i d_i}{N} \times h$$

$$= 45 + \frac{44}{80} \times 10$$

$$= 50.5$$

Median:

Ungrouped data:

i) If N is odd, then $\frac{N+1}{2}$ th observation is median.

ii) If N is even, then the median is the average of $\frac{N}{2}$ th and $\frac{N+1}{2}$ th observation.

- Example 1:

The weights of 11 mothers in kg were recorded as follows:

47 44 42 41 58 52 55 39 40 43 61

Find the median

Solution:

We arrange the weights in ascending order

39 40 41 42 43 44 47 52 55 58 61

Since, N=11 is odd, hence the median is

$\frac{11+1}{2} = 6$th observation.

6th observation is 44. So, median is 44.

- even হলে ——

39 40 41 42 43 44 47 52 55 58

Since, N=10 is even, hence

the median is the average of $\frac{10}{2} = 5$th and $\frac{10}{2}+1 = 6$th observation

5th observation $= 43$
6th observation $= 44$

$$\therefore \text{Median} = \frac{43+44}{2} = 43.5$$

- Find the median of the following

20 18 22 27 25 12 15

Solution:
We arrange the numbers in ascending order.

12 15 18 20 22 25 27

Since, $N = 7$ is odd, hence the median is $\frac{7+1}{2} = 4$th observation.

4th observation is 20. So, median is 20.

Calculation of Median (Grouped data):

For grouped data, Median = $L + \frac{\frac{N}{2} - P.c.f}{f} \times h$

where,

h = The size of class interval

L = Lower limit of median class.

P.c.f = Preceding cumulative frequency of median class. (Cumulative frequency above median class)

f = frequency of the median class

Example 1:

Calculate the median for the distribution of the weights of 150 students from the given below:

Weight	30-40	40-50	50-60	60-70	70-80	80-90
Frequency	18	37	45	27	15	8.

Solution:

We construct the following data:

Weight	Frequency	cumulative frequency
30-40	18	18
40-50	37	55
50-60	45	100
60-70	27	127
70-80	15	142
80-90	8	150
	N=150	

Median is $\frac{N}{2} = \frac{150}{2} = 75$ th observation.

75th observation lies in class 50-60.

So, median class 50-60.

Now, $L = 50$, $p.c.f = 55$, $f = 45$, $h = 10$

We know,

$$\text{Median} = L + \frac{\frac{N}{2} - p.c.f}{f} \times h$$

$$= 50 + \frac{75 - 55}{45} \times 10$$

$$= 54.44$$

- Example 2:

Following distribution gives the pattern of overtime done by 100 employees. Calculate the median.

Overtime	10-15	15-20	20-25	25-30	30-35	35-40
No of employee	11	20	35	20	8	6

Solution: We construct the following data:

Overtime	No of employee	Cumilative frequency
10-15	11	11
15-20	20	31
20-25	35	66
25-30	20	86
30-35	8	94
35-40	6	100
	N = 100	

Median is $\frac{N}{2} = \frac{100}{2} = 50$th observation.

50th observation lies in class 20-25.

median class is 20-25.

Now, L = 20, p.c.f = 31, f = 35, h = 5

We know,

$$\text{Median} = L + \frac{\frac{N}{2} - p.c.f}{f} \times h$$

$$= 20 + \frac{50-31}{35} \times 5$$

$$= 22.714$$

Hence 50% of the workers doing overtime up to 22.714 hrs and the remaining 50% of the workers doing overtime more than 22.714 hrs.

4.9.24

Example 3: Calculate the median from the following distribution gives the profit of 125 companies:

Profit (crore)	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No of companies	4	12	24	36	20	16	8	3

Comment on your result.

Solution:

We construct the following table:

Profit (crore)	No. of Companies	Cumulative frequency
0-10	4	4
10-20	12	16
20-30	24	40
30-40	36	76
40-50	20	96
50-60	16	112
60-70	8	120
70-80	3	125
	N=125	

Median is $\frac{N}{2} = \frac{125}{2} = 61.5 \approx 62$th observation.

62th observation lies in class 30-40.

Here, $L = 30$, $p.c.f = 40$, $f = 36$, $h = 10$

We know,

$$\text{Median} = L + \frac{\frac{N}{2} - p.c.f}{f} \times h$$

$$= 30 + \frac{61.5 - 40}{36} \times 10 = 35.97$$

Hence, 50% of the companies have profits up to ~~36.5~~ 35.97 crores and the remaining 50% of the companies have profits more than 35.97 crores.

Example 4:

Calculate the median from the following distribution.

No of days absent	5	10	15	20	25	30	35	40	45
No of students	29	195	241	117	52	10	6	3	2

Solution:

We construct the following data:

No of days absent	No of students	cumulative frequency
5	29	29
10	195	224
15	241	465
20	117	582
25	52	634
30	10	644
35	6	650
40	3	653
45	2	655
	N = 655	

Median is $\frac{N}{2} = \frac{655}{2} = 327.5$ ~~th~~ $\approx$ 328th observation

328th observation lies in class 10–15.

Here, $L = 10$, $p.c.f = 224$, $f = 241$, $h = 5$

We know,

$$Median = L + \frac{\frac{N}{2} - p.c.f}{f} \times h$$

$$= 10 + \frac{327.5 - 224}{241} \times 5$$

$$= 12.15$$

For Practice

1. Calculate the median of the following

Marks	0-20	20-40	40-60	60-80	80-100	100-120	120-140
No of students	4	26	22	10	9	6	3

Solution:

We construct the following table:

Marks	No. of students	cumulative frequency
0-20	4	4
20-40	26	30
40-60	22	52
60-80	10	62
80-100	9	71
100-120	6	77
120-140	3	80
	N= 80	

Median is $\frac{N}{2} = \frac{80}{2} = 40$th observation.

40th observation lies in class 40–60.

Here, $L = 40$, $p.c.f = 30$, $f = 22$, $h = 20$

We know,

$$\text{Median} = L + \frac{\frac{N}{2} - p.c.f}{f} \times h$$

$$= 40 + \frac{40 - 30}{22} \times 20$$

$$= 49.09$$

2. Find the median of the following data

Marks	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
No of students	7	32	56	106	180	164	86	44

Solution:

We construct the following table:

Marks	No of students	Cumulative frequency
0–10	7	7
10–20	32	39
20–30	56	95
30–40	106	201
40–50	180	381
50–60	164	545
60–70	86	631
70–80	44	675
	N = 675	

CamScanner

Median is $\frac{N}{2} = \frac{675}{2} = 337.5 \simeq 338$th observation.

338th observation lies in class 40-50

Here, $L = 40$, $p.c.f = 201$, $f = 180$, $h = 10$

We know,

$$\text{Median} = L + \frac{\frac{N}{2} - p.c.f}{f} \times h$$

$$= 40 + \frac{337.5 - 201}{180} \times 10$$

$$= 47.58$$

Quartiles:

Quartiles are those value which divide the total frequency into four parts. We need three values to divide the whole frequency into four parts. That is why, there are three quartiles.

Q_1 = First quartile

Q_2 = Second quartile

Q_3 = Third quartile

Q_1 Q_2 Q_3

Formula:

Quartiles, $Q_i = L + \dfrac{\frac{N}{4} \times i - p.c.f}{f} \times h$ $\quad i = 1,2,3$

Example 1:

The profits earned by 100 companies are given below:

Profits (lacs)	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
No of companies	4	8	18	30	15	10	8	7

Calculate Q_1, Median or Q_2, Q_3.

Solution:

We construct the following table:

Profits (lacs)	No of companies	cumulative frequency
20-30	4	4
30-40	8	12
40-50	18	30
50-60	30	60
60-70	15	75
70-80	10	85
80-90	8	93
90-100	7	100
	N = 100	

The first quartile, Q_1, is $\frac{N}{4} = 25$th observation

25th observation lies in class 40-50.

Here, $L = 40$, $p.c.f = 12$, $f = 18$, $h = 10$

We know;

First quartile, $Q_1 = L + \frac{\frac{N}{4} - p.c.f}{f} \times h$

$= 40 + \frac{25-12}{18} \times 10$

$= 47.22$

Comment: 25% of the companies earn an annual profit of 47.22 lacs.

The second quartile, Q_2, is $\frac{N}{2} = 50$th observation. 50th observation lies in class 50-60.

Here, $L = 50$, $p.c.f = 30$, $f = 30$, $h = 10$

We know,

Second quartile, $Q_2 = L + \frac{\frac{N}{2} - p.c.f}{f} \times h$

$= 50 + \frac{50-30}{30} \times 10$

$= 56.67$

Comment: 50% of the companies earn an annual profit of 56.67 lacs.

The third quartile, Q_3, is $\frac{3N}{4} = 75$th observation. 75th observation lies in class 60-70.

Here, $L = 60$, $p.c.f = 60$, $f = 15$, $h = 10$

We know,

Third quartile, $Q_3 = L + \frac{\frac{3N}{4} - p.c.f}{f} \times h$

$$= 60 + \frac{\frac{3 \times 100}{4} - 60}{15} \times 10$$

$$= 70$$

Comment: 75% of the companies earn an annual profit of 70 lacks.

Example 2: Following distribution gives the pattern of overtime done by 100 employee. Calculate first quartile Q_1.

Overtime	10-15	15-20	20-25	25-30	30-35	35-40
No of employee	11	20	35	20	8	6

Solution: We construct the following table:

Overtime	No of employee	cumulative frequency
10-15	11	11
15-20	20	31
20-25	35	66
25-30	20	86
30-35	8	94
35-40	6	100
	N = 100	

The first quartile, Q_1, is $\frac{N}{4} = 25$th observation.

25th observation lies in class 15-20

Here, L = 15, p.c.f = 11, f = 20, h = 5

We know,

$$\text{first quartile}, Q_1 = L + \frac{\frac{N}{4} - p.c.f}{f} \times h$$

$$= 15 + \frac{25 - 11}{20} \times 5$$

$$= 18.5$$

Comment: 25% of the employees done overtime of 18.5 hours.

Mode:

Mode is defined as the value which occurs the maximum number of times ie. having the maximum frequency.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$$

where,

h = The size of class interval

L = Lower limit of modal class (The class having maximum frequency)

Δ_1 = Difference between the frequency of the modal class and the pre-modal class.

Δ_2 = Difference between the frequency of the modal class and the post-modal class

Example 1:

Calculate the mode for the distribution of the weights of 150 students from the given below:

weight	30-40	40-50	50-60	60-70	70-80	80-90
frequency	18	37	45	27	15	8

Solution:

We have the following table:

Weight	Frequency
30-40	18
40-50	37
50-60	45
60-70	27
70-80	15
80-90	8

Here the highest frequency is 45 which lies in the class 50-60. So, modal class is 50-60.

Now, $\Delta_1 = 45-37 = 8$

$\Delta_2 = 45-27 = 18$

$L = 50$

$h = 10$

We know,

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1+\Delta_2} \times h$$

$$= 50 + \frac{8}{8+18} \times 10$$

$$= 53.08$$

9.9.24

Example 2:

Find the mode of the following data:

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No of students	7	32	56	106	180	164	86	44

Solution:

We have the following table:

Marks	No. of students
0-10	7
10-20	32
20-30	56
30-40	106
40-50	180
50-60	164
60-70	86
70-80	44

Here the highest frequency is 180 which lies in the class 40-50. So, modal class is 40-50.

Now, $\Delta_1 = 180 - 106 = 74$

$\Delta_2 = 180 - 164 = 16$

$L = 40$

$h = 10$

We know,

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$$

$$= 40 + \frac{74}{74+16} \times 10$$

$$= 48.22$$

For practice

Example 1: Find mode of the following data relates to the sales of 100 companies.

Sales	0-60	60-62	62-64	64-66	66-68	68-70	70-72
No of companies	12	18	25	30	10	3	2

Solution:
We have the following table:

Sales	No of companies
0-60	12
60-62	18
62-64	25
64-66	30
66-68	10
68-70	3
70-72	2

Here the highest frequency is 30, which lies in the class 64-66. So, modal class is 64-

Now, $\Delta_1 = 30 - 25 = 5$

$\Delta_2 = 30 - 10 = 20$

$L = 64$

$h = 2$

We know,

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$$

$$= 64 + \frac{5}{5+20} \times 2$$

$$= 64.4$$

Empirical Relation between Mean, Median, Mode

$$\boxed{\text{Mode} = 3\ \text{Median} - 2\ \text{Mean}}$$

Example 1:

Calculate the median and Mode of the frequency distribution given below. Hence calculate the mean using empirical relation between them.

Weight	30-40	40-50	50-60	60-70	70-80	80-90
Frequency	18	37	45	27	15	8

Solution:
We construct the following table:

Weight	Frequency	cumulative frequency
30-40	18	18
40-50	37	55
50-60	45	100
60-70	27	127
70-80	15	142
80-90	8	150
	N = 150	

Median is $\frac{N}{2} = 75$th observation. 75th observation lies in class 50-60.

Here, $L = 50$, $p.c.f = 55$, $f = 45$, $h = 10$

We know,

$$\text{Median} = L + \frac{\frac{N}{2} - p.c.f}{f} \times h$$

$$= 50 + \frac{75-55}{45} \times 10$$

$$= 54.44$$

The highest frequency is 45, which lies in the class 50-60. So, modal class is 50-60.

Now,

$$\Delta_1 = 45 - 37 = 8$$

$$\Delta_2 = 45 - 27 = 18$$

$$L = 50$$

$$h = 10$$

We know,

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$$

$$= 50 + \frac{8}{8+18} \times 10$$

$$= 53.08$$

Now,

We know,

$$\text{Mode} = 3\,\text{median} - 2\,\text{mean}$$

$$53.08 = 3 \times 54.44 - 2\,\text{mean}$$

$$2\,\text{mean} = 163.32 - 53.08$$

$$\text{mean} = \frac{110.24}{2} = 55.12$$

Example 2:

Calculate the arithmetic Mean and Median of the frequency distribution given below. Hence calculate the mode using empirical relation between them.

Height	130-134	135-139	140-144	145-149	150-154	155-159	160-164
No of students	5	15	28	24	17	10	1

Solution:

We construct the following table:

Height	Mid value x_i	No of students f_i	$d_i = \frac{x_i - A}{h}$	$f_i d_i$	Cumulative frequency
129.5–134.5	132	5	−3	−15	5
134.5–139.5	137	15	−2	−30	20
139.5–144.5	142	28	−1	−28	48
144.5–149.5	147 → A	24	0	0	72
149.5–154.5	152	17	1	17	89
154.5–159.5	157	10	2	20	99
159.5–164.5	162	1	3	3	100
		N = 100		$\sum f_i d_i = -33$	

We know,

$$\text{Mean, } \bar{x} = A + \frac{\sum_{i=1}^{n} f_i d_i}{N} \times h$$

$$= 147 + \frac{-33}{100} \times 5$$

$$= 145.35$$

Median is $\frac{N}{2} = 50th$ observation. 50th observation lies in class 144.5 - 149.5

So, median class is 144.5 - 149.5

Now,

$L = 144.5$, $p.c.f = 48$, $f = 24$, $h = 5$

We know,

$$\text{Median} = L + \frac{\frac{N}{2} - p.c.f}{f} \times h$$

$$= 144.5 + \frac{50 - 48}{24} \times 5$$

$$= 144.916$$

Now,

We know,

$$\text{Mode} = 3\,\text{Median} - 2\,\text{Mean}$$

$$= 3 \times 144.916 - 2 \times 145.35$$

$$= 144.051$$

For practice

Example 1:

Calculate the arithmetic mean and median of the frequency distribution given below. Hence calculate the mode using empirical relation between them:

Marks	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
No of students	2	12	15	20	18	10	9	4

Solution:
We construct the following table:

Marks	Mid value x_i	No of students f_i	$d_i = \frac{x_i - A}{h}$	$f_i d_i$	Cumulative frequency
19.5 - 29.5	~~24.75~~ 24.5	2	-3	-6	2
29.5 - 39.5	34.5	12	-2	-24	14
39.5 - 49.5	44.5	15	-1	-15	29
49.5 - 59.5	54.5 → A	20	0	0	49
59.5 - 69.5	64.5	18	1	18	67
69.5 - 79.5	74.5	10	2	20	77
79.5 - 89.5	84.5	9	3	27	86
89.5 - 99.5	94.5	4	4	16	90
		N = 90		$\sum f_i d_i = 36$	

We know,

$$\text{Mean}, \bar{x} = A + \frac{\sum_{i=1}^{n} f_i d_i}{N} \times h$$

$$= 54.5 + \frac{36}{90} \times 10$$

$$= 58.5$$

Median is $\frac{N}{2} = 45$th observation. 45th observation lies in class 49.5 - 59.5

So, median class is 49.5 - 59.5

Now, $L = 49.5$, $p.c.f = 29$; $f = 20$, $h = 10$

We know,

$$Median = L + \frac{\frac{N}{2} - p.c.f}{f} \times h$$

$$= 49.5 + \frac{45 - 29}{20} \times 10$$

$$= 57.5$$

Now,

We know,

$$Mode = 3\ Median - 2\ Mean$$

$$= 3 \times 57.5 - 2 \times 58.5$$

$$= 172.5 - 117$$

$$= 55.5$$

Measures of Dispersion

Different types of measures of dispersion.

i) The range

ii) The interquartile range or quartile deviation

iii) The mean deviation ***

iv) The variance ***

v) The standard deviation ***

Range:

The range is the absolute difference between the largest value and the smallest value in the set of data. Symbolically

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

Where, L = Largest value
S = Smallest value

Example 1:

The following are the prices of shares of a company from Saturday to Thursday:

Day	Sat	Sun	Mon	Tue	Wed	Thu
Price (tk)	200	210	208	160	220	250

Solution:

$$\text{Range } R = L - S = 250 - 160 = 90 \text{ tk}$$

In the frequency distribution, range is calculated by taking the difference between the lower limit of the lowest class and upper limit of the highest class.

Example:

Calculate the range of the following data:

Profit (lakh)	10-20	20-30	30-40	40-50	50-60
No of companies	200	210	208	160	220

Solution:

$$\text{Range } R = L - S = 60 - 10 = 50 \text{ lakh}$$

Limitation:

Range cannot tell us anything about the character about the distribution within two extreme observations.

Data 1	6	46	46	46	46	46	46	R=40
Data 2	6	6	6	6	46	46	46	R= 40
Data 3	6	10	15	25	36	39	46	R= 40

In all the three series of data range is same (i.e, 40) but it does not mean that the distribution of data are same.

Quartile Deviation:

$$Q.D = \frac{Q_3 - Q_1}{2}$$

Quartiles, $Q_i = L + \frac{\frac{i \times N}{4} - p.c.f}{f} \times h.$ $\quad i = 1, 2, 3$

Example 1:

The profits earned by 100 companies are given below:

Profits (lakh)	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
No of companies	4	8	18	30	15	10	8	7

i) Calculate the range within which middle 50% companies fall.

ii) Calculate quartile deviation.

11.9.24

Solution:

We construct the following table:

Profits (lakhs)	No of companies	Cumulative frequency
20-30	4	4
30-40	8	12
40-50	18	30
50-60	30	60
60-70	15	75
70-80	10	85
80-90	8	93
90-100	7	100
	N=100	

The first quartile is $\frac{N}{4} = 25$th observation which lies in the class 40-50.

Here, $L=40$, $p.c.f = 12$, $f=18$, $h=10$.

First quartile, $Q_1 = L + \frac{\frac{N}{4} - p.c.f}{f} \times h$

$= 40 + \frac{25-12}{18} \times 10$

$= 47.22$ lakh

The third quartile is $\frac{3N}{4} = 75$th observation which lies in the class 60-70.

Here, $L=60$, $p.c.f = 60$, $f=15$, $h=10$

Third quartile, $Q_3 = L + \frac{\frac{3N}{4} - p.c.f}{f} \times h$

$$= 60 + \frac{75-60}{15} \times 10$$

$$= 70 \text{ lakh}$$

Range within which middle 50% companies fall $= Q_3 - Q_1 = 70 - 47.22$

$$= 22.78$$

Quartile deviation $= \frac{Q_3 - Q_1}{2}$

$$= \frac{70 - 47.22}{2}$$

$$= 11.39 \text{ lakh}$$

Example 2:

Based on the frequency distribution given below calculate quartile deviation:

Tax Paid (Lakh)	5-10	10-15	15-20	20-25	25-30	30-35	35-40
No of Managers	18	30	46	28	20	12	6

Solution: We construct the following table:

Tax Paid (Lakh)	No. of Managers	cumulative frequency
5-10	18	18
10-15	30	48
15-20	46	94
20-25	28	122
25-30	20	142
30-35	12	154
35-40	6	160
	N = 160	

The first quartile is $\frac{N}{4} = 40$th observation, which lies in the class 10-15.

Here, $L = 10$, $p.c.f = 18$, $f = 30$, $h = 5$

First quartile, $Q_1 = L + \frac{\frac{N}{4} - p.c.f}{f} \times h$

$$= 10 + \frac{40 - 18}{30} \times 5$$

$$= 13.67 \text{ Lakh}$$

The third quartile is $\frac{3N}{4} = 120$ th observation which lies in the class 20-25.

Here, $L = 20$, $p.c.f = 94$, $f = 28$, $h = 5$

Third quartile, $Q_3 = L + \frac{\frac{3N}{4} - p.c.f}{f} \times h$

$$= 20 + \frac{120 - 94}{28} \times 5$$

$$= 24.64 \text{ Lakh}$$

$\therefore$ Quartile deviation $= \frac{Q_3 - Q_1}{2}$

$$= \frac{24.64 - 13.67}{2}$$

$$= 5.485 \text{ Lakh}$$

Example 3: For the following data:

Age (yr)	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No of Members	6	5	8	15	7	6	3

i) Calculate the range within which middle 50% members fall

ii) Calculate quartile deviation.

Solution:
We construct the following table;

Age (yr)	No of members	cumulative frequency
0-10	6	6
10-20	5	11
20-30	8	19
30-40	15	34
40-50	7	41
50-60	6	47
60-70	3	50
	N=50	

The first quartile is $\frac{N}{4} = 12.5 \approx 13th$ observation, which lies in the class 20-30.
Here, $L = 20$, $p.c.f = 11$, $f = 8$, $h = 10$
First quartile, $Q_1 = L + \frac{\frac{N}{4} - p.c.f}{f} \times h$

$$= 20 + \frac{12.5 - 11}{8} \times 10$$

$$= 21.875 \text{ year}$$

The Third quartile is $\frac{3N}{4} = 37.5 \approx 38th$ observation, which lies in the class 40-50.
Here, $L = 40$, $p.c.f = 34$, $f = 7$, $h = 10$
Third quartile, $Q_3 = L + \frac{\frac{3N}{4} - p.c.f}{f} \times h$

$$= 40 + \frac{37.5 - 34}{7} \times 10$$

$$= 45 \text{ year}$$

Range within which middle 50% members fall $= Q_3 - Q_1 = 45 - 21.875 = 23.125$

Quartile deviation $= \frac{Q_3 - Q_1}{2}$

$$= \frac{45 - 21.875}{2}$$

$$= 11.5625 \text{ year}$$

Mean Deviation:

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

Example 1:

Based on the frequency distribution given below calculate mean deviation.

Batsman 1	49	50	55	54

Solution:

Batsman 1

x_i	$\bar{x}$	$\lvert x_i - \bar{x} \rvert$
49		3
50	$\frac{208}{4} = 52$	2
55		3
54		2
		$\sum \lvert x_i - \bar{x} \rvert = 10$

$$M.D = \frac{\sum |x_i - \bar{x}|}{N} = \frac{10}{4} = 2.5$$

Example 2:

Based on the frequency distribution given below calculate mean deviation.

Batsman 2	10	68	90	40

Solution:

Batsman 2

x_i	$\bar{x}$	$\lvert x_i - \bar{x} \rvert$
16	$\frac{208}{4} = 52$	42
68		16
90		38
40		12
		$\sum \lvert x_i - \bar{x} \rvert = 108$

$$M.D = \frac{\sum \lvert x_i - \bar{x} \rvert}{N} = \frac{108}{4} = 27$$

Mean deviation (Grouped data):

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

Where,

$\bar{x}$ = Arithmetic mean $\left(\bar{x} = A + \frac{\sum f_i d_i}{N} \times h\right)$

x_i = Mid values of each class

N = The total frequency

Example 1:

Calculate mean deviation for the following data

Sales (Lakhs)	10-20	20-30	30-40	40-50	50-60
No of Days	3	6	11	3	2

Solution:

We construct the following data.

Sales (lac)	Mid value x_i	No of days f_i	d_i	$f_i d_i$	$\bar{x} = A + \frac{\sum f_i d_i}{N} \times h$	$\lvert x_i - \bar{x} \rvert$	$f_i \lvert x_i - \bar{x} \rvert$
10-20	15	3	-2	-6	Here,	18	54
20-30	25	6	-1	-6	$A = 35$	8	48
30-40	35	11	0	0	$\bar{x} = 35 + \frac{-5}{25} \times 10$	2	22
40-50	45	3	1	3	$= 33$	12	36
50-60	55	2	2	4		22	44
		$N = 25$		$\sum f_i d_i = -5$			$\sum f_i \lvert x_i - \bar{x} \rvert = 204$

So, the mean deviation is $M.D = \frac{\sum f_i \lvert x_i - \bar{x} \rvert}{N}$

$$= \frac{204}{25}$$

$$= 8.16 \text{ lakhs}$$

Example 2:

Calculate mean deviation for the following data

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No of Students	6	5	8	15	7	6	3

Solution:
We construct the following data:

marks	Mid value x_i	f_i	d_i	f_id_i	$\bar{x}$	$\lvert x_i-\bar{x}\rvert$	$f_i\lvert x_i-\bar{x}\rvert$
0-10	5	6	-3	-18	Here, $A=35$	28.4	170.4
10-20	15	5	-2	-10	$\bar{x}=A+\frac{\sum f_id_i}{N}\times h$	18.4	92
20-30	25	8	-1	-8	$=35+\frac{-8}{50}\times 10$	8.4	67.2
30-40	35	15	0	0	$=33.4$	1.6	24
40-50	45	7	1	7		11.6	81.2
50-60	55	6	2	12		21.6	129.6
60-70	65	3	3	9		31.6	94.6
		$N=50$		$\sum f_id_i=-8$			$\sum f_i\lvert x_i-\bar{x}\rvert=658.$

So, the mean deviation is $M.D=\frac{\sum f_i\lvert x_i-\bar{x}\rvert}{N}$

$$=\frac{658.4}{50}$$

$$=13.168$$

Example 3:

Calculate mean deviation for the following data.

class	0-6	6-12	12-18	18-24	24-30
No of days	8	10	12	9	5

Solution:
We construct the following data:

Class	Mid values x_i	f_i	d_i	$f_i d_i$	$\bar{x}$	$\lvert x_i-\bar{x}\rvert$	$f_i\lvert x_i-\bar{x}\rvert$
0-6	3	8	-2	-16	Here, $A=15$	11.045	88.36
6-12	9	10	-1	-10	$\bar{x}=A+\frac{\sum f_i d_i}{N}\times h$	5.045	50.45
12-18	15	12	0	0	$=15+\frac{-7}{44}\times 6$	0.955	11.46
18-24	21	9	1	9		6.955	62.595
24-30	27	5	2	10	$=14.045$	12.955	64.775
		$N=44$		$\sum f_i d_i=-7$			$\sum f_i\lvert x_i-\bar{x}\rvert=227.64$

So, the mean deviation is $M.D=\frac{\sum f_i\lvert x_i-\bar{x}\rvert}{N}$

$$=\frac{227.64}{44}$$

$$=6.31$$

Variance:

$$\sigma^2=\frac{\sum(x_i-\bar{x})^2}{N} \quad \text{(Ungrouped Data)}$$

4) **Example 1:** Find variance from the weekly wages of 10 workers working in a factory.

1320 1310 1315 1322 1326 1340 1325 1321 1320 1331

Solution:

x_i	$\bar{x}$	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$
1320		-3	9
1310		-13	169
1315	$\frac{13230}{10}$	-8	64
1322		-1	1
1326		3	9
1340	$=1323$	17	289
1325		2	4
1321		-2	4
1320		-3	9
1331		8	64
			$\sum(x_i-\bar{x})^2=622$

$$\sigma^2=\frac{\sum(x_i-\bar{x})^2}{N}=\frac{622}{10}=62.2 \text{ tk}$$

Variance (Grouped Data):

$$\sigma^2 = \frac{\sum f_i(x_i-\bar{x})^2}{N}$$

or, $$\sigma^2 = \left[\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2\right] \times h^2$$

Example 1: Calculate variance for the following data

Profit (lakhs)	10-20	20-30	30-40	40-50	50-60
No. of Companies	8	12	20	6	4

Solution:
We construct the following data:

Profit	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
10-20	15	8	-2	-16	32
20-30	25	12	-1	-12	12
30-40	35	20	0	0	0
40-50	45	6	1	6	6
50-60	55	4	2	8	16
		$N=50$		$\sum f_i d_i = -14$	$\sum f_i d_i^2 = 66$

Here, $h = 10$

Therefore,

Variance,

$$\sigma^2 = h^2 \times \left[\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2\right]$$

$$= 10^2 \times \left[\frac{66}{50} - \left(\frac{-14}{50}\right)^2\right]$$

$= 124.16$ lakh

Example 2:

Calculation of variance for the following data

Profit (Lakhs)	0-10	10-20	20-30	30-40	40-50
No of Companies	6	25	36	20	13

Solution:

We construct the following data:

Profit (Lakhs)	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
0-10	5	6	-2	-12	24
10-20	15	25	-1	-25	25
20-30	25	36	0	0	0
30-40	35	20	1	20	20
40-50	45	13	2	25	52
		N=100		$\sum f_i d_i = 9$	$\sum f_i d_i^2 = 121$

Here, $h = 10$

Therefore,

variance,

$$\sigma^2 = h^2 \times \left[\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N} \right)^2 \right]$$

$$= 10^2 \times \left[\frac{121}{100} - \left(\frac{9}{100} \right)^2 \right]$$

$$= 120.19 \text{ lakh}$$

Example 3:

Calculate variance for the following data:

Profit (Lakhs)	0-10	10-20	20-30	30-40	40-50	50-60
No of Companies	8	12	20	30	20	10

Solution:

We construct the following data:

Profit	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
0-10	5	8	-2	-16	32
10-20	15	12	-1	-12	12
20-30	25	20	0	0	0
30-40	35	30	1	30	30
40-50	45	20	2	40	80
50-60	55	10	3	30	90
		$N = 100$		$\Sigma f_i d_i = 72$	$\Sigma f_i d_i^2 = \cancel{3800}$ 244

Here,

$h = 10$

Therefore,

variance,

$$\sigma^2 = h^2 \times \left[\frac{\Sigma f_i d_i^2}{N} - \left(\frac{\Sigma f_i d_i}{N} \right)^2 \right]$$

$$= 10^2 \times \left[\frac{\overset{244}{\cancel{3800}}}{100} - \left(\frac{72}{100} \right)^2 \right]$$

$= 192.16$ Lakh

Example 4:

Calculate variance for the following data

Tax (Thousand)	5-10	10-15	15-20	20-25	25-30	30-35	35-40
No of Managers	18	30	46	28	20	12	6

Solution:

We construct the following data:

Tax (Thousand)	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
5-10	7.5	18	-3	-54	162
10-15	12.5	30	-2	-60	120
15-20	17.5	46	-1	-46	46
20-25	22.5	28	0	0	0
25-30	27.5	20	1	20	20
30-35	32.5	12	2	24	48
35-40	37.5	6	3	18	54
		$N=160$		$\sum f_i d_i = -98$	$\sum f_i d_i^2 = 450$

Here, $h=5$

Therefore;

variance,

$$\sigma^2 = h^2 \times \left[\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2\right]$$

$$= 5^2 \times \left[\frac{450}{160} - \left(\frac{-98}{160}\right)^2\right]$$

$= 60.93$ Thousand

Standard Deviation (grouped data)

For grouped data,
Standard deviation, $\sigma = \sqrt{\frac{\sum f_i(x_i-\bar{x})^2}{N}}$

$$\Rightarrow \sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$$

Example 1:
Calculate standard deviation for the following data:

Profit (Lakhs)	10-20	20-30	30-40	40-50	50-60
No of Companies	8	12	20	6	4

Solution: We construct the following table:

Profit	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
10-20	15	8	-2	-16	32
20-30	25	12	-1	-12	12
30-40	35	20	0	0	0
40-50	45	6	1	6	6
50-60	55	4	2	8	16
		N=50		$\sum f_i d_i = -14$	$\sum f_i d_i^2 = 66$

$$\therefore \sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$$

$$= \sqrt{\frac{66}{50} - \left(\frac{-14}{50}\right)^2} \times 10$$

$= 11.14$ lakh

Example 2:
Calculate standard deviation for the following data

Profit (lakhs)	0-10	10-20	20-30	30-40	40-50
No of companies	6	25	36	20	13

Solution:
We construct the following table:

Profit	Midvalue x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
0-10	5	6	-2	-12	24
10-20	15	25	-1	-25	25
20-30	25	36	0	0	0
30-40	35	20	1	20	20
40-50	45	13	2	25	52
		$N=100$		$\sum f_i d_i = 9$	$\sum f_i d_i^2 = 121$

$$\therefore \sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$$

$$= \sqrt{\frac{121}{100} - \left(\frac{9}{100}\right)^2} \times 10$$

$$= 10.96 \text{ lakh}$$

Example 3:

Calculate mean and standard deviation:

Reject amount	21–25	26–30	31–35	36–40	41–45	46–50	51–55
No of Operators	5	15	28	42	15	12	3

Solution:

We construct the following table:

Reject amount	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
20.5–25.5	23	5	−3	−15	45
25.5–30.5	28	15	−2	−30	60
30.5–35.5	33	28	−1	−28	28
35.5–40.5	38→A	42	0	0	0
40.5–45.5	43	15	1	15	15
45.5–50.5	48	12	2	24	48
50.5–55.5	53	3	3	9	27
		$N=120$		$\sum f_i d_i = -25$	$\sum f_i d_i^2 = 223$

We know,

$$\text{Mean}, \bar{x} = A + \frac{\sum f_i d_i}{N} \times h$$

$$= 38 + \frac{-25}{120} \times 10$$

$$= 36.96$$

And we know,

$$\text{Standard deviation}, \sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$$

$$= \sqrt{\frac{223}{120} - \left(\frac{-25}{120}\right)^2} \times 5$$

$$= 6.375$$

For Practice

Example 1:
Calculate standard deviation for the following data:

Profit (lakhs)	0-10	10-20	20-30	30-40	40-50	50-60
No of Companies	8	12	20	30	20	10

Solution:
We construct the following table:

Profit	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
0-10	5	8	-3	-24	72
10-20	15	12	-2	-24	48
20-30	25	20	-1	-20	20
30-40	35	30	0	0	0
40-50	45	20	1	20	20
50-60	55	10	2	20	40
		$N=100$		$\sum f_id_i=-28$	$\sum f_id_i^2=200$

We know,

Standard deviation, $\sigma = \sqrt{\frac{\sum f_id_i^2}{N} - \left(\frac{\sum f_id_i}{N}\right)^2} \times h$

$= \sqrt{\frac{200}{100} - \left(\frac{-28}{100}\right)^2} \times 10$

$= 13.862$

Example 2:

Calculate standard deviation for the following data:

Tax (Thousand)	5-10	10-15	15-20	20-25	25-30	30-35	35-40
No of Managers	18	30	46	28	20	12	6

Solution:

We construct the following table:

Tax	Mid value x_i	f_i	~~f_ix_i~~ d_i	f_id_i	$f_id_i^2$
5-10	7.5	18	-3	-54	162
10-15	12.5	30	-2	-60	120
15-20	17.5	46	-1	-46	46
20-25	22.5	28	0	0	0
25-30	27.5	20	1	20	20
30-35	32.5	12	2	24	48
35-40	37.5	6	3	18	54
		$N=160$		$\Sigma f_id_i=-98$	$\Sigma f_id_i^2=450$

We know,

Standard deviation, $\sigma = \sqrt{\frac{\Sigma f_id_i^2}{N} - \left(\frac{\Sigma f_id_i}{N}\right)^2} \times h$

$= \sqrt{\frac{450}{160} - \left(\frac{-98}{160}\right)^2} \times 5$

$= \sqrt{2.8125 - 0.375} \times 5$

$= 7.806$

Coefficient of Variance

$$C.V = \frac{\sigma}{\bar{x}} \times 100$$

Where,

σ = Standard deviation = $\sqrt{\frac{\sum(x_i-\bar{x})^2}{N}}$

$\bar{x}$ = Arithmetic mean = $\frac{\sum x_i}{N}$

Example 1:
Find standard deviation from the price of a company share during the last 10 months in Dhaka stock exchange.
105 120 115 118 130 127 109 110 104 112

Solution:
We construct the following table:

x_i	$\bar{x}$	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$
105		-10	100
120		5	25
115		0	0
118	$\frac{1150}{10}=115$	3	9
130		15	225
127		12	144
109		6	36
110		-5	25
104		-11	121
112		-3	9
			$\sum(x_i-\bar{x})^2=694$

Standard deviation, $\sigma = \sqrt{\frac{\sum(x_i-\bar{x})^2}{N}} = \sqrt{\frac{694}{10}} = 8.33$ tk

$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{8.33}{115} \times 100 = 7.24\%$

Example 2:

Find standard deviation from the price of a company share during the last 10 months in Chittagong stock exchange.

108 117 120 130 100 125 125 120 110 135

Solution:

We construct the following table:

x_i	$\bar{x}$	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$
108		-11	121
117		-2	4
120		1	1
130	$\frac{1190}{10}=119$	11	121
100		-19	361
125		6	36
125		6	36
120		1	1
110		-9	81
135		16	256
			$\Sigma(x_i-\bar{x})^2=1018$

Standard deviation, $\sigma = \sqrt{\frac{\Sigma(x_i-\bar{x})^2}{N}} = \sqrt{\frac{1018}{10}} = 10.09$ tk

$CV = \frac{\sigma}{\bar{x}} \times 100 = \frac{10.09}{119} \times 100 = 10.09\ \%$

Coefficient of variance (Grouped data)

For grouped data,
coefficient of variance, $C.V = \frac{\sigma}{\bar{x}} \times 100$

where, $\sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$

and $\bar{x} = A + \frac{\sum f_i d_i}{N} \times h$

Example 1;
Calculate standard deviation and coefficient of variance for the following data.

Profit (lakhs)	10-20	20-30	30-40	40-50	50-60
No of companies	8	12	20	6	4

Solution:
We construct the following table:

Profit	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
10-20	15	8	-2	-16	32
20-30	25	12	-1	-12	12
30-40	35	20	0	0	0
40-50	45	6	1	6	6
50-60	55	4	2	8	16
		$N=50$		$\sum f_i d_i = -14$	$\sum f_i d_i^2 = 66$

Mean, $\bar{x} = A + \frac{\sum f_i d_i}{N} \times h$

$= 35 + \frac{-14}{50} \times 10$

$=$ ~~11.14 lakh~~ 32.2 lakh

Standard deviation, $\sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$

$= \sqrt{\frac{66}{50} - \left(\frac{-14}{50}\right)^2} \times 10 = 11.14$ lakh

$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{11.14}{32.2} \times 100 = 34.6\ \%$

Example 2:
Calculate standard deviation and coefficient of variance for the following data:

Profit (lakhs)	0-10	10-20	20-30	30-40	40-50
No of Companies	6	25	36	20	13

Solution:
We construct the following table:

Profit	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
0-10	5	6	-2	-12	24
10-20	15	25	-1	-25	25
20-30	25	36	0	0	0
30-40	35	20	1	20	20
40-50	45	13	2	25	52
		$N=100$		$\Sigma f_i d_i = 9$	$\Sigma f_i d_i^2 = 121$

Mean, $\bar{x} = A + \frac{\Sigma f_i d_i}{N} \times h = 25 + \frac{9}{100} \times 10 = 25.9$ lakh

Standard deviation:

$$\sigma = \sqrt{\frac{\Sigma f_i d_i^2}{N} - \left(\frac{\Sigma f_i d_i}{N}\right)^2} \times h$$

$$= \sqrt{\frac{121}{100} - \left(\frac{9}{100}\right)^2} \times 10$$

$= 10.96$ lakh

$$C.V = \frac{\sigma}{\bar{x}} \times 100$$

$$= \frac{10.96}{25.9} \times 100$$

$= 42.32\%$

Example 3:

Calculate standard deviation and coefficient of variance.

Reject amount	21-25	26-30	31-35	36-40	41-45	46-50	51-55
No of operators	5	15	28	42	15	12	3

Solution:

We construct the following table:

Reject amount	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
20.5-25.5	23	5	-3	-15	45
25.5-30.5	28	15	-2	-30	60
30.5-35.5	33	28	-1	-28	28
35.5-40.5	38	42	0	0	0
40.5-45.5	43	15	1	15	15
45.5-50.5	48	12	2	24	48
50.5-55.5	53	3	3	9	27
		$\Sigma N = 120$		$\Sigma f_id_i = -25$	$\Sigma f_id_i^2 = 223$

Mean $= \bar{x} = A + \frac{\Sigma f_id_i}{N} \times h = 38 + \frac{-25}{120} \times 10 = 36.96$

Standard deviation:

$$\sigma = \sqrt{\frac{\Sigma f_id_i^2}{N} - \left(\frac{\Sigma f_id_i}{N}\right)^2} \times h$$

$$= \sqrt{\frac{223}{120} - \left(\frac{-25}{120}\right)^2} \times 5$$

$$= 6.375$$

$$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{6.375}{36.96} \times 100 = 17.25\%$$

For practice

[1] Example 1:

Calculate standard deviation and coefficient of variance for the following data:

Profit	0-10	10-20	20-30	30-40	40-50	50-60
No of companies	8	12	20	30	20	10

Solution:

We construct the following table:

Profit	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
0-10	5	8	-2	-16	32
10-20	15	12	-1	-12	12
20-30	25←A	20	0	0	0
30-40	35	30	1	30	30
40-50	45	20	2	40	80
50-60	55	10	3	30	90
		N = 100		$\Sigma f_id_i = 72$	$\Sigma f_id_i^2 = 244$

Mean, $\bar{x} = A + \frac{\Sigma f_id_i}{N} \times h = 25 + \frac{72}{100} \times 10 = 32.2$

Standard deviation:

$$\sigma = \sqrt{\frac{\Sigma f_id_i^2}{N} - \left(\frac{\Sigma f_id_i}{N}\right)^2} \times h$$

$$= \sqrt{\frac{244}{100} - \left(\frac{72}{100}\right)^2} \times 10$$

$$= 13.86$$

$$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{13.86}{32.2} \times 100 = 43.05\ \%$$

Example 2:

Calculate standard deviation and coefficient of variance for the following data:

Tax (Thousand)	5–10	10–15	15–20	20–25	25–30	30–35	35–40
No of Managers	18	30	46	28	20	12	6

Solution:

We construct the following table:

Tax	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
5–10	7.5	18	−3	−54	162
10–15	12.5	30	−2	−60	120
15–20	17.5	46	−1	−46	46
20–25	22.5	28	0	0	0
25–30	27.5	20	1	20	20
30–35	32.5	12	2	24	48
35–40	37.5	6	3	18	54
		$N=160$		$\sum f_i d_i = -98$	$\sum f_i d_i^2 = 450$

Mean, $\bar{x} = A + \frac{\sum f_i d_i}{N} \times h = 22.5 + \frac{-98}{160} \times 5 = 19.43$

Standard deviation:

$$\sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$$

$$= \sqrt{\frac{450}{160} - \left(\frac{-98}{160}\right)^2} \times 5$$

$$= 7.806$$

$$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{7.806}{19.437} \times 100 = 40.16\%$$

Example 3:

Calculate standard deviation and coefficient of variance for the following data:

Wages	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No of workers	1	2	4	13	21	9	6	4

Solution:

We construct the following table:

Wages	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
10-20	15	1	-4	-4	16
20-30	25	2	-3	-6	18
30-40	35	4	-2	-8	16
40-50	45	13	-1	-13	13
50-60	55	21	0	0	0
60-70	65	9	1	9	9
70-80	75	6	2	12	24
80-90	85	4	3	12	36
		N=60		$\sum f_id_i = 2$	$\sum f_id_i^2 = 132$

Mean, $\bar{x} = A + \frac{\sum f_id_i}{N} \times h = 55 + \frac{2}{60} \times 10 = 55.33$

Standard deviation,

$$\sigma = \sqrt{\frac{\sum f_id_i^2}{N} - \left(\frac{\sum f_id_i}{N}\right)^2} \times h$$

$$= \sqrt{\frac{132}{60} - \left(\frac{2}{60}\right)^2} \times 10$$

$= 14.83$

$$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{14.83}{55.33} \times 100 = 26.8\%$$

Example 4:

Calculate standard deviation and coefficient of variance for the following data.

Profits	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
No of Companies	4	8	18	30	15	10	8	7

Solution:

We construct the following table:

Profits	Mid value x_i	f_i	~~fi~~ d_i	$f_i d_i$	$f_i d_i^2$
20-30	25	4	-4	-16	64
30-40	35	8	-3	-24	72
40-50	45	18	-2	-36	72
50-60	55	30	-1	-30	30
60-70	65	15	0	0	0
70-80	75	10	1	10	10
80-90	85	8	2	16	32
90-100	95	7	3	21	63
		$N = 100$		$\Sigma f_i d_i = -59$	$\Sigma f_i d_i^2 = 343$

Mean, $\bar{x} = A + \frac{\Sigma f_i d_i}{N} \times h = 65 + \frac{-59}{100} \times 10$

$= 59.1$

Standard deviation, $\sigma = \sqrt{\frac{\Sigma f_i d_i^2}{N} - \left(\frac{\Sigma f_i d_i}{N}\right)^2} \times h$

$= \sqrt{\frac{343}{100} - \left(\frac{-59}{100}\right)^2} \times 10$

$= 17.56$

$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{17.56}{59.1} \times 100 = 29.70\ \%$

Example 5:
Calculate standard deviation and coefficient of variance for the following data.

Weights	210-215	215-220	220-225	225-230	230-235	235-240	240-245	245-250
No of Boys	8	13	16	29	14	10	7	3

Solution:
We construct the following data:

Weights	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
210-215	212.5	8	-4	-32	128
215-220	217.5	13	-3	-49	147
220-225	222.5	16	-2	-32	64
225-230	227.5	29	-1	-29	29
230-235	232.5	14	0	0	0
235-240	237.5	10	1	10	10
240-245	242.5	7	2	14	28
245-250	247.5	3	3	9	27
		$N=100$		$\sum f_id_i=-109$	$\sum f_id_i^2=433$

Mean, $\bar{x} = A + \frac{\sum f_id_i}{N} \times h = 232.5 + \frac{-109}{100} \times 5 = 227.05$

Standard deviation, $\sigma = \sqrt{\frac{\sum f_id_i^2}{N} - \left(\frac{\sum f_id_i}{N}\right)^2} \times h$

$= \sqrt{\frac{433}{100} - \left(\frac{-109}{100}\right)^2} \times 5$

$= 8.86$

$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{8.86}{227.05} \times 100 = 3.90\%$

Example 6:
Calculate standard deviation and coefficient of variance for the following data.

Turnover (Lakh)	5-10	10-15	15-20	20-25	25-30	30-35	35-40
No of companies	8	18	42	62	30	10	4

Solution:
We construct the following table.

Turnover	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
5-10	7.5	8	-3	-24	72
10-15	12.5	18	-2	-36	72
15-20	17.5	42	-1	-42	42
20-25	22.5	62	0	0	0
25-30	27.5	30	1	30	30
30-35	32.5	10	2	20	40
35-40	37.5	4	3	12	36
		$N=174$		$\sum f_id_i=-40$	$\sum f_id_i^2=292$

Mean, $\bar{x} = A + \frac{\sum f_id_i}{N} \times h = 22.5 + \frac{-40}{174} \times 5 = 21.35$

Standard deviation, $\sigma = \sqrt{\frac{\sum f_id_i^2}{N} - \left(\frac{\sum f_id_i}{N}\right)^2} \times h$

$$= \sqrt{\frac{292}{174} - \left(\frac{-40}{174}\right)^2} \times 5$$

$$= 6.374$$

$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{6.374}{21.35} \times 100 = 29.86\%$

Empirical Relation between Measures of variation

Quartile Deviation (Q.D) = $\frac{2}{3}$ Standard Deviation (σ)

Mean Deviation (M.D) = $\frac{4}{5}$ Standard Deviation (σ)

Quartile Deviation (Q.D) = $\frac{5}{6}$ Mean Deviation (M.D)

Example 1:

Calculate standard deviation and then calculate mean deviation using empirical relation for the following data:

Profit (Lakh)	10–20	20–30	30–40	40–50	50–60
No of Companies	8	12	20	6	4

Solution:

We construct the following table.

Profit	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
10–20	15	8	−2	−16	32
20–30	25	12	−1	−12	12
30–40	35	20	0	0	0
40–50	45	6	1	6	6
50–60	55	4	2	8	16
		N = 50		$\Sigma f_id_i=$	$\Sigma f_id_i^2 = 66$

Standard deviation: $\sigma = \sqrt{\frac{\Sigma f_id_i^2}{N} - \left(\frac{\Sigma f_id_i}{N}\right)^2} \times h$

$= \sqrt{\frac{66}{50} - \left(\frac{-14}{50}\right)^2} \times 10 = 11.14 \text{ lakh}$

Mean deviation:

$M.D = \frac{4}{5} \times \sigma = \frac{4}{5} \times 11.14 = 8.912 \text{ lakh}$

Example 2:
Calculate mean deviation and then calculate quartile deviation using empirical relation for the following data.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No of students	6	5	8	15	7	6	3

Solution:
We construct the following data:

Marks	Mid value x_i	f_i	d_i	$f_i d_i$	$\bar{x}$	$\lvert x_i - \bar{x}\rvert$	$f_i \lvert x_i - \bar{x}\rvert$
0-10	5	6	-3	-18		28.4	170.4
10-20	15	5	-2	-10		18.4	92
20-30	25	8	-1	-8	$35+\frac{-8}{50}\times 10$	8.4	67.2
30-40	35	15	0	0	$=33.4$	1.6	24
40-50	45	7	1	7		11.6	81.2
50-60	55	6	2	12		21.6	129.6
60-70	65	3	3	9		31.6	94.6
		$N=50$		$\Sigma f_i d_i = -8$			$\Sigma f_i \lvert x_i - \bar{x}\rvert = 658.4$

$$M.D = \frac{\Sigma f_i \lvert x_i - \bar{x}\rvert}{N} = \frac{658.4}{50} = 13.168$$

Quartile deviation, (QD) $= \frac{5}{6} \times M.D$

$$= \frac{5}{6} \times 13.168$$

$$= 10.97$$

Skewness

Coefficient of skewness, $Sk_p = \frac{\text{Mean} - \text{Mode}}{\sigma}$

If $Sk_p > 0$, the distribution positively skewed (Mean > Mode)

If $Sk_p < 0$, the distribution negatively skewed (Mean < Mode)

Example 1:
Calculate coefficient of skewness and comment on its value:

Profit (lakhs)	100-120	120-140	140-160	160-180	180-200	200-220	220-240
No of companies	17	53	199	194	327	208	2

Solution:
We construct the following table:

Profit	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
100-120	110	17	-3	-51	153
120-140	130	53	-2	-106	212
140-160	150	199	-1	-199	199
160-180	170	194	0	0	0
180-200	190	327	1	327	327
200-220	210	208	2	416	832
220-240	230	2	3	6	18
		N=1000		$\sum f_id_i = 393$	$\sum f_id_i^2 = 1741$

Mean, $\bar{x} = A + \frac{\sum f_id_i}{N} \times h = 170 + \frac{393}{1000} \times 20 = 177.86$

Since highest frequency is 327 which lies in the class 180-200, So, Modal class is 180-200.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 180 + \frac{133}{133+119} \times 20 = 190.56$$

Standard deviation, $\sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$

$$= \sqrt{\frac{1741}{1000} - \left(\frac{393}{1000}\right)^2} \times 20$$

$$= 25.2$$

Coefficient of skewness, $Sk_p = \frac{\text{Mean} - \text{Mode}}{\sigma}$

$$= \frac{177.86 - 190.56}{25.2}$$

$$= -0.504$$

The distribution is negatively skewed, that is mode of the distribution is greater than mean.

An analysis of workers resulted in the following distribution:

Age	20-25	25-30	30-35	35-40	40-45	45-50	50-55
No of employees	8	12	20	25	15	12	8

Calculate coefficient of skewness and comment on the result.

Solution:

We construct the following table:

Age	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
20-25	22.5	8	-3	-24	72
25-30	27.5	12	-2	-24	48
30-35	32.5	20	-1	-20	20
35-40	37.5	25	0	0	0
40-45	42.5	15	1	15	15
45-50	47.5	12	2	24	48
50-55	52.5	8	3	24	72
		N=100		$\sum f_id_i = -5$	$\sum f_id_i^2 = 275$

Mean, $\bar{x} = A + \frac{\sum f_id_i}{N} \times h = 37.5 + \frac{-5}{100} \times 5 = 37.25$

Since, the highest frequency is 25 which lies in the class 35-40. So, Modal class is 35-40.

Mode $= L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 35 + \frac{5}{5+10} \times 5 = 36.67$

Standard deviation, $\sigma = \sqrt{\frac{\sum f_id_i^2}{N} - \left(\frac{\sum f_id_i}{N}\right)^2} \times h$

$= \sqrt{\frac{275}{100} - \left(\frac{-5}{100}\right)^2} \times 5$

$= 8.29$

Coefficient of skewness, $SK_p = \frac{\text{Mean} - \text{Mode}}{\sigma}$

$= \frac{37.25 - 36.67}{8.29}$

$= 0.07$

Coefficient of skewness $= 0.07$ indicates that the distribution is positively skewed, that is mode of the distribution is less than mean.

Example 3:

Calculate coefficient of skewness and comment on the result.

No of rejects	21–25	26–30	31–35	36–40	41–45	46–50	51–55
No of companies	5	15	28	42	15	12	3

Solution:

We construct the following table:

No of rejects	Mid values x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
20.5–25.5	23	5	−3	−15	45
25.5–30.5	28	15	−2	−30	60
30.5–35.5	33	28	−1	−28	28
35.5–40.5	38	42	0	0	0
40.5–45.5	43	15	1	15	15
45.5–50.5	48	12	2	24	48
50.5–55.5	53	3	3	9	27
		$N=120$		$\Sigma f_i d_i = -25$	$\Sigma f_i d_i^2 = 223$

Mean, $\bar{x} = A + \frac{\Sigma f_i d_i}{N} \times h = 38 + \frac{-25}{120} \times 5 = 36.96$

Since, highest frequency is 42, which lies in the class 35.5–40.5. So, Modal class is 35.5–40.5.

Mode $= L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 35.5 + \frac{14}{14+17} \times 5 = 37.21$

Standard deviation, $\sigma = \sqrt{\frac{\Sigma f_i d_i^2}{N} - \left(\frac{\Sigma f_i d_i}{N}\right)^2} \times h = \sqrt{\frac{223}{120} - \left(\frac{-25}{120}\right)^2} \times 5$

$= 6.736$

Coefficient of skewness, $Sk_p = \frac{\text{Mean} - \text{Mode}}{\sigma} = \frac{36.96 - 37.21}{6.736}$ $\cancel{= 0.03}$

$= -0.037$.

The distribution is negatively skewed, that is mode of the distribution is greater than mean.

For practice

Example 1:
Calculate coefficient of skewness for the following distribution.

Marks	0-20	20-40	40-60	60-80	80-100
No of students	18	22	30	20	10

Solution:
We construct the following table:

Marks	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
0-20	10	18	-2	-36	72
20-40	30	22	-1	-22	22
40-60	50	30	0	0	0
60-80	70	20	1	20	20
80-100	90	10	2	20	40
		$N=100$		$\sum f_id_i=-18$	$\sum f_id_i^2=154$

Mean, $\bar{x} = A + \frac{\sum f_id_i}{N} \times h = 50 + \frac{-18}{100} \times 20 = 46.4$

Since the highest frequency is 30, which lies in the class 40-60. So, ~~Median~~ Modal class is 40-60.

Mode $= L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 40 + \frac{8}{8+10} \times 20 = 48.89$

Standard deviation, $\sigma = \sqrt{\frac{\sum f_id_i^2}{N} - \left(\frac{\sum f_id_i}{N}\right)^2} \times h$

$= \sqrt{\frac{154}{100} - \left(\frac{-18}{100}\right)^2} \times 20 = 24.56$

Coefficient of skewness, $Sk_p = \frac{\text{Mean} - \text{Mode}}{\sigma}$

$= \frac{46.4 - 48.89}{24.56} = -0.101$

The distribution is negatively skewed, that is mode of the distribution is greater than Mean.

Example 2:
An analysis of electricity consumption resulted in the following distribution

Consumption (kw/h)	0-10	10-20	20-30	30-40	40-50
No of users	6	25	36	20	13

Calculate coefficient of skewness and comment on the result.

Solution:
We construct the following table:

Consumption	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$
0-10	5	6	-2	-12	24
10-20	15	25	-1	-25	25
20-30	25	36	0	0	0
30-40	35	20	1	20	20
40-50	45	13	2	26	52
		$N=100$		$\sum f_id_i = 9$	$\sum f_id_i^2 = 121$

Mean, $\bar{x} = A + \frac{\sum f_id_i}{N} \times h = 25 + \frac{9}{100} \times 10 = 25.9$

Since the highest frequency is 36, which lies in the class 20-30. So, Modal class is 20-30.

Mode $= L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 20 + \frac{11}{11+16} \times 10 = 24.07$

Standard deviation, $\sigma = \sqrt{\frac{\sum f_id_i^2}{N} - \left(\frac{\sum f_id_i}{N}\right)^2} \times h$

$= \sqrt{\frac{121}{100} - \left(\frac{9}{100}\right)^2} \times 10$

$= 10.96$

Coefficient of skewness, $Sk_p = \frac{\text{Mean} - \text{Mode}}{\sigma} = \frac{25.9 - 24.07}{10.96}$

$= 0.167$

The distribution is positively skewed, that is mode of the distribution is less than Mean.

CamScanner

Example 3:

Calculate coefficient of skewness for the following distribution.

Class	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
frequency	11	22	30	35	21	11	6	5

Solution:

We construct the following table:

Class	Mid Value x_i	f_i	d_i	f_id_i	$f_id_i^2$
0-10	5	11	-3	-33	99
10-20	15	22	-2	-44	88
20-30	25	30	-1	-30	30
30-40	35	35	0	0	0
40-50	45	21	1	21	21
50-60	55	11	2	22	44
60-70	65	6	3	18	54
70-80	75	5	4	20	80
		$N=141$		$\Sigma f_id_i=-26$	$\Sigma f_id_i^2=416$

Mean, $\bar{x} = A + \frac{\Sigma f_id_i}{N} \times h = 35 + \frac{-26}{141} \times 10 = 33.156$

The highest frequency is 35, which lies in the class 30-40, So, modal class is 30-40.

Mode $= L + \frac{\Delta_1}{\Delta_1+\Delta_2} \times h = 30 + \frac{5}{5+14} \times 10 = 32.63$

Standard deviation, $\sigma = \sqrt{\frac{\Sigma f_id_i^2}{N} - \left(\frac{\Sigma f_id_i}{N}\right)^2} \times h$

$= \sqrt{\frac{416}{141} - \left(\frac{-26}{141}\right)^2} \times 10 = 17.08$

Coefficient of skewness, $Sk_p = \frac{\text{Mean}-\text{Mode}}{\sigma}$

$= \frac{33.156 - 32.63}{17.08} = 0.031$

The distribution is positively skewed, That is mode of the distribution is less than Mean.

Example 4:
Calculate coefficient of skewness for for the following distribution.

Class	4000-4200	4200-4400	4400-4600	4600-4800	4800-5000	5000-5200	5200-5400
Frequency	22	38	65	75	80	70	50

Solution:
We construct the following table:

Class	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
4000-4200	4100	22	-3	-66	198
4200-4400	4300	38	-2	-76	152
4400-4600	4500	65	-1	-65	65
4600-4800	4700	75	0	0	0
4800-5000	4900	80	1	80	80
5000-5200	5100	70	2	140	280
5200-5400	5300	50	3	150	450
		$N=400$		$\Sigma f_i d_i = 163$	$\Sigma f_i d_i^2 = 1225$

Mean, $\bar{x} = A + \frac{\Sigma f_i d_i}{N} \times h = 4700 + \frac{163}{400} \times 200 = 4781.5$

The highest frequency is 80, which lies in the Class 4800-5000. So, Modal class is 4800-5000.

Mode $= L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 4800 + \frac{5}{5+10} \times 200 = 4866.67$

Standard deviation, $\sigma = \sqrt{\frac{\Sigma f_i d_i^2}{N} - \left(\frac{\Sigma f_i d_i}{N}\right)^2} \times h$

$= \sqrt{\frac{1225}{400} - \left(\frac{163}{400}\right)^2} \times 200$

$= 340.4$

Coefficient of skewness, $Sk_p = \frac{\text{Mean} - \text{Mode}}{\sigma}$

$= \frac{4781.5 - 4866.67}{340.4}$

$= -0.25$

The distribution is negatively skewed, That is Mode of the distribution is greater than Mean.

Moments

Raw Moments (About A)	Central Moments (About $\bar{x}$)
$\mu_1' = \frac{\sum f_i d_i}{N} \times h$	$\mu_1 = 0$
$\mu_2' = \frac{\sum f_i d_i^2}{N} \times h^2$	$\mu_2 = \mu_2' - \mu_1'^2$
$\mu_3' = \frac{\sum f_i d_i^3}{N} \times h^3$	$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$
$\mu_4' = \frac{\sum f_i d_i^4}{N} \times h^4$	$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$

Example 1:
An analysis of workers resulted in the following distribution:

Earnings (tk)	50-70	70-90	90-110	110-130	130-150	150-170	170-190
No of Employees	4	8	12	20	6	7	3

Calculate the first four moments about assumed mean. Convert the result into moments about the mean.

Solution:
We construct the following table:

Earnings (tk)	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$	$f_i d_i^3$	$f_i d_i^4$
50-70	60	4	-3	-12	36	-108	324
70-90	80	8	-2	-16	32	-64	128
90-110	100	12	-1	-12	12	-12	12
110-130	120	20	0	0	0	0	0
130-150	140	6	1	6	6	6	6
150-170	160	7	2	14	28	56	112
170-190	180	3	3	9	27	81	243
		$N=60$		$\sum f_i d_i = -5$	$\sum f_i d_i^2 = 141$	$\sum f_i d_i^3 = -41$	$\sum f_i d_i^4 = 825$

Moments about assumed mean:

$$\mu_1' = \frac{\sum f_i d_i}{N} \times h = \frac{-5}{60} \times 20 = -3.67$$

$$\mu_2' = \frac{\sum f_i d_i^2}{N} \times h^2 = \frac{141}{60} \times 20^2 = 940$$

$\therefore \mu_3' = \frac{\sum f_i d_i^3}{N} \times h^3 = \frac{-41}{60} \times 20^3 = -5466.67$

$\mu_4' = \frac{\sum f_i d_i^4}{N} \times h^4 = \frac{825}{60} \times 20^4 = 2200000$

Moments about mean:

$\mu_1 = 0$

$\mu_2 = \mu_2' - \mu_1'^2 = 940 - (-3.67)^2 = 926.56$

$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$

$= -5466.67 - 3(940)(-3.67) + 2(-3.67)^3 = 4774.832$

$\mu_4 = \mu_4' - 4\mu_3'\mu_1 + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$

$= 2200000 - 4(-5466.67)(-3.67) + 6(940)(-3.67)^2$

$- 3(-3.67)^4 = 2195107.3$

Example 2:

An analysis of companies resulted in the following distribution.

Profit (lakhs)	10-20	20-30	30-40	40-50	50-60
No of companies	18	20	30	22	10

Calculate the first four moments about assumed mean. Convert the result into moments about the mean.

Solution:

We construct the following table:

Profit	Mid value x_i	f_i	d_i	f_id_i	$f_id_i^2$	$f_id_i^3$	$f_id_i^4$
10-20	15	18	-2	-36	72	-144	288
20-30	25	20	-1	-20	20	-20	20
30-40	35	30	0	0	0	0	0
40-50	45	22	1	22	22	22	22
50-60	55	10	2	40	40	80	160
		N = 100		$\Sigma f_id_i = -14$	$\Sigma f_id_i^2 = 154$	$\Sigma f_id_i^3 = -62$	$\Sigma f_id_i^4 = 490$

Moments about assumed mean:

$$\mu_1' = \frac{\Sigma f_id_i}{N} \times h = \frac{-14}{100} \times 10 = -1.4$$

$$\mu_2' = \frac{\Sigma f_id_i^2}{N} \times h^2 = \frac{154}{100} \times 10^2 = 154$$

$$\mu_3' = \frac{\Sigma f_id_i^3}{N} \times h^3 = \frac{-62}{100} \times 10^3 = -620$$

$$\mu_4' = \frac{\Sigma f_id_i^4}{N} \times h^4 = \frac{490}{100} \times 10^4 = 49000$$

Moments about Mean:

$\mu_1 = 0$

$\mu_2 = \mu_2' - \mu_1'^2 = 154 - (-1.4)^2 = 152.04$

$\mu_3 = \mu_3' - 3\mu_2'\mu_1 + 2\mu_1'^3 = -620 - 3(154)(-1.4)$

$+ 2(-1.4)^3 = 21.312$

$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$

$= 49000 - 4(-620)(-1.4) + 6(154)(-1.4)^2 - 3(-1.4)^4$

$= 47327.51$