Chapter: Measures of Dispersion

Measures of dispersion or variation measures the extent of variation or deviation of individual values from central value. It helps us to measure variation among the data.

Batsman 1	49	50	55	54	$\bar{x} = 52$
Batsman 2	10	68	90	40	$\bar{x} = 52$

Which batsman is more consistent?

Objectives:

- To find variability among the data.
- To find the nature of distribution about the entire data.
- Find uniformness among the data.

Types:

Different types of measures of dispersion or variation are:

1. The range
2. The interquartile range or quartile deviation
3. The mean deviation
4. The variance
5. The standard deviation

Absolute Measures of Variation: Absolute measures of variation are expressed in the same statistical unit in which the original data are given. **Ex:** Salary range between a manager and average salary of workers.

Relative Measures of Variation: A measure of relative variation is the ratio of a measure of absolute variation to an average. It also called coefficient of variation. **Ex:** The percentage of salary a manager get more from the average salary of workers.

Absolute Measures of Variation

Range:

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

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

Where L= Largest Value

S= Smallest Value

Example: 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:

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

- 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(Lakhs)	10-20	20-30	30-40	40-50	50-60
No of Companies	200	210	208	160	220

Solution:

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

Limitation:

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

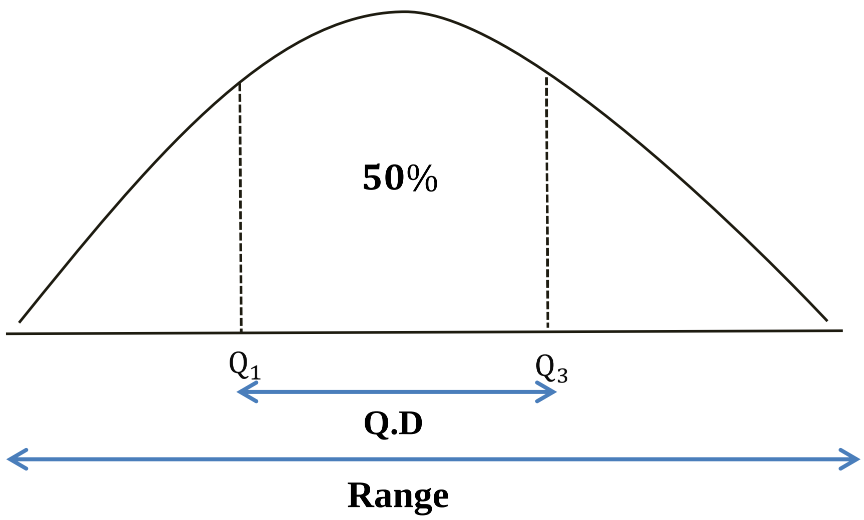
Data 1	6	46	46	46	46	46	46	**R=40**
Data 2	6	6	6	6	46	46	46	**R=40**
Data 3	6	10	15	25	36	39	46	**R=40**

- 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.

Interquartile range or quartile deviation:

Interquartile range or quartile deviation represents the difference between the third quartile Q_3 and the first quartile Q_1. Symbolically

$$Q.D = \frac{Q_3 - Q_1}{2} \qquad \text{Q.D = Quartile deviation}$$



- When quartile deviation is very small it describes small variation of the central 50% observations, and a high quartile deviation means that the variation among central observations is large.

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

Quartile class identified by $Q_i = \frac{i \times N}{4}$ **th observation.**

Example: The profits earned by 100 companies are given below:

Profits (lakhs)	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.

Solution:

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 $Q_1 = \frac{1 \times 100}{4} = 25$ th observation. 25 th observation lies in 40 - 50. Quartile class is 40 - 50.

We know $Q_i = L + \frac{\frac{i \times N}{4} - \text{p.c.f}}{f} \times h$

$$Q_1 = 40 + \frac{\frac{1 \times 100}{4} - 12}{18} \times 10 = 47.22 \text{ lakhs}$$

The third quartile $Q_3 = \frac{3 \times 100}{4} = 75$ th observation. 75 th observation lies in 60 - 70. Quartile class is 60 - 70.

$$Q_3 = 60 + \frac{\frac{3 \times 100}{4} - 60}{15} \times 10 = 70 \text{ lakhs}$$

Range within which middle 50% companies fall = $Q_3 - Q_1 = 70 - 47.22 = 22.78$

Quartile deviation $\quad Q.D = \frac{Q_3 - Q_1}{2} = \frac{70 - 47.22}{2} = 11.39$ lakhs

Example: 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

ANS: $\mathbf{Q_1 = 13.67, Q_3 = 24.64,\ Q.D = 5.485}$

Example: 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.

Limitation:

Quartile deviation ignores 50% items, i.e., the first 25% and the last 25%. Since it does not depend upon every observation it is not regarded as good method of measuring variation.

Mean Deviation:

Mean deviation is an average of absolute deviation of each observations from the mean. It is obtained by calculating the absolute deviation of each observations from the mean, and then averaging the deviations by taking their mean.

Calculation of mean deviation (Ungrouped Data)

For ungrouped data

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

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

| 49-52 |

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

Example: Based on the frequency distribution given below calculate mean deviation.

Batsman 2	10	68	90	40

Solution:

Batsman 2

x_i	$\bar{x}$	$\lvert x_i - \bar{x} \rvert$
10	$\frac{208}{4} = 52$	42
68		16
90		38
40		12
		$\sum \lvert x_i - \bar{x} \rvert = 108$

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

Calculation of mean deviation (Grouped Data)

For grouped data

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

Where

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

x_i = Mid values of each class.

N = The total frequency.

Example: 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:

Sales	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		18	54
20-30	25	6	-1	-6	$35 + \frac{-5}{25} \times 10 = 33$	8	48
30-40	35	11	0	0		2	22
40-50	45	3	+1	+3		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

$$M.D = \frac{\sum f_i |x_i - \bar{x}|}{N} = \frac{204}{25} = 8.16 \text{ lakhs}$$

Example: 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:

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	$35 + \frac{-8}{50} \times 10 = 33.4$	28.4	170.4
10-20	15	5	-2	-10		18.4	92
20-30	25	8	-1	-8		8.4	67.2
30-40	35	15	0	0		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	94.8
		N = 50		$\sum f_i d_i$ = -8			$\sum f_i\lvert x_i - \bar{x}\rvert$= 659.2

$$\mathrm{M.D} = \frac{\sum f_i |x_i - \bar{x}|}{N} = \frac{659.2}{50} = 13.184$$

Example: 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

ANS: 6.3

Limitation:

In calculation algebraic signs are ignored while taking the deviations of the items. If the signs of the deviations are not ignored, the net sum of the deviations will be zero.

Variance:

Variance is the average of the squares of the deviations of the given values from their arithmetic mean. It is denoted by σ^2

Calculation of variance (Ungrouped Data)

For ungrouped data, variance

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

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

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

Calculation of variance (Grouped Data)

For grouped data, variance

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

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: 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:

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

f_1=8, d_1=-2
$8\times(-2)^2$

$$\sigma^2 = \left[\frac{\sum f_i\, {d_i}^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2\right] \times h^2 = \left[\frac{66}{50} - \left(\frac{-14}{50}\right)^2\right] \times 10^2 =$$

$$= 1.2416 \times 100 = 124.16 \text{ lakhs}$$

Example: Calculate 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:

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	+26	52
		N = 100		$\sum f_i d_i = 9$	$\sum f_i {d_i}^2 = 121$

$$\sigma^2 = \left[\frac{\sum f_i\, {d_i}^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2\right] \times h^2 = \left[\frac{121}{100} - \left(\frac{9}{100}\right)^2\right] \times 10^2 =$$

$$= 1.2019 \times 100 = 120.19 \text{ lakhs}$$

Example: 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

ANS: 192.155

Example: 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

ANS: 60.93

Standard Deviation:

Standard deviation is the positive square root of the average of the squares of the deviations of the given values from their arithmetic mean. It is denoted by σ. That is

$$\text{Standard deviation} \qquad \text{S.D or } \sigma = \sqrt{\sigma^2} = \sqrt{\text{variance}}$$

Calculation of standard deviation (Ungrouped Data)

For ungrouped data, standard deviation

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

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

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N}} = \sqrt{\frac{622}{10}} = \sqrt{62.2} = 7.89$$

Calculation of standard deviation (Grouped Data)

For grouped data, standard deviation

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

or

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

Example: 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:

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

f_1=8, d_1=-2
$8\times(-2)^2$

$$\sigma = \sqrt{\frac{\sum f_i\, d_i{}^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h = \sqrt{\frac{66}{50} - \left(\frac{-14}{50}\right)^2} \times 10$$

$$= 11.14 \text{ lakhs}$$

Example: 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:

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		$\sum f_i d_i = 9$	$\sum f_i d_i^2 = 121$

$$\sigma = \sqrt{\frac{\sum f_i\, d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h = \sqrt{\frac{121}{100} - \left(\frac{9}{100}\right)^2} \times 10$$

$$= 10.96 \text{ lakhs}$$

Example: An analysis of production rejects resulted in the following data

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

Calculate mean and standard deviation.

Solution: Using class boundaries we get

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	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

Mean: $$\bar{x} = A + \frac{\sum f_i d_i}{N} \times h = 38 + \frac{-25}{120} \times 10 = 36.96$$

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

$$= \sqrt{1.858 - .043} \times 5 = 6.375$$

Example: Calculate standard deviation for the following data

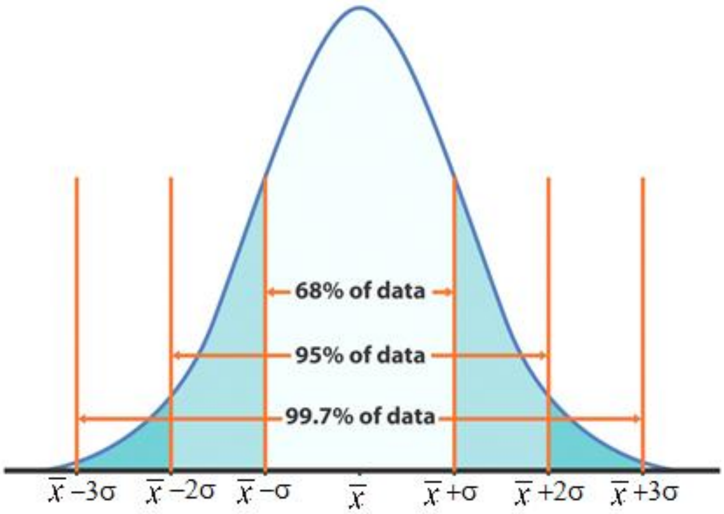
Profit(Lakhs)	0-10	10-20	20-30	30-40	40-50	50-60
No of Companies	8	12	20	30	20	10

ANS: 13.862

Example: 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

ANS: 7.806



- $\bar{x} \pm \sigma$ covers 68.27% observations.
- $\bar{x} \pm 2\sigma$ covers 95.45% observations.
- $\bar{x} \pm 3\sigma$ covers 99.73% observations.

Example: The breaking strength of 80 test pieces of a certain alloy is give in the following table, the unit being given to the nearest pounds per square inch.

Breaking strength	44-46	46-48	48-50	50-52	52-54
No of pieces	3	24	27	21	5

Calculate the average breaking strength of the alloy and the standard deviation. Calculate the percentage of observations lying between mean $\pm 2\sigma$.

Solution:

Breaking strength	Mid value x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$
44-46	45	3	-2	-6	12
46-48	47	24	-1	-24	24
48-50	49	27	0	0	0
50-52	51	21	+1	+21	21
52-54	53	5	+2	+10	20
		$N = 80$		$\sum f_i d_i = 1$	$\sum f_i d_i^2 = 77$

$$\bar{x} = A + \frac{\sum f_i d_i}{N} \times h = 49 + \frac{1}{80} \times 2 = 49.025$$

$$\sigma = \sqrt{\frac{\sum f_i\, d_i^{\,2}}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h = \sqrt{\frac{77}{80} - \left(\frac{1}{80}\right)^2} \times 2$$

$$= 1.96$$

Now $\bar{x} \pm 2\sigma = 49.025 \pm 2(1.96) = 45.105 \text{ to } 52.945 = 45 \text{ to } 53 \text{ (approx.)}$. We have to calculate the percentage of items lying between 45 to 53. Since the number are equally distributed, then the frequency at 45 be 1.5 and the frequency at 53 be 2.5.

Thus the total frequency between 45 to 53 is (1.5 + 24 + 27 + 21 + 2.5) = 76. The percentage is $\frac{76}{80} \times 100 = 95$. Thus there are 95 percent observations lying between mean $\pm\, 2\sigma$.

Relative Measures of Variation

Coefficient of Variance:

Measures of variance discussed above are absolute measures. One of the relative measure is known as coefficient of variance. A coefficient of variance is computed as a percentage of the standard deviation of the distribution to the mean of the same distribution. This measure represents the spread of the distribution relative to the mean of the same distribution. Symbolically:

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

Where $\sigma =$ Standard deviation

$\bar{x} =$ Arithmetic mean

Consider a statistic of height and weight of the 150 children:

	Height	**Weight**
Mean	40 inch	10 kg
S.D	5 inch	2 kg
C.V	**12.5%**	**20%**

Which one has more variability among the children, height or weight?

- The coefficient of variance helpful in comparing the relative variation in several data set having different measure of unit.

Consider the blood pressure of a group of patient were measured at two levels:

	Systolic (mm)	**Diastolic (mm)**
Average	130	60
S.D	15	8
C.V	**11.5%**	**13.3%**

Which pressure shows greater variability among the patients?

- The coefficient of variance helpful in comparing the relative variation in several data set having different means and different standard deviations.

Calculation of coefficient of variance (Ungrouped Data)

For ungrouped data, coefficient of variance

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

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

Example: Find coefficient of variance 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:

x_i	$\bar{x}$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
105		-10	100
120		+5	25
115		0	0
118		+3	9
130	$\frac{1150}{10} = 115$	+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 $\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N}} = \sqrt{\frac{694}{10}} = \sqrt{69.4} = 8.33$ tk

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

Example: Find coefficient of variance 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:

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

Standard deviation $\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N}} = \sqrt{\frac{1018}{10}} = \sqrt{101.8} = 10.09$ tk

$$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{10.09}{119} \times 100 = 10.09\ \%$$

- Since the coefficient of variance of Dhaka is less, therefore it is safe to invest in Dhaka stock exchange.

Calculation of 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: 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:

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 $\bar{x} = A + \frac{\sum f_i d_i}{N} \times h = 35 + \frac{-14}{50} \times 10 = 32.2$ 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: 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:

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		$\sum f_i d_i = 9$	$\sum f_i {d_i}^2 = 121$

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

Standard deviation: $$\sigma = \sqrt{\frac{\sum f_i {d_i}^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h = \sqrt{\frac{121}{100} - \left(\frac{9}{100}\right)^2} \times 10 = 10.96 \text{ lakh}$$

$$\text{C.V} = \frac{\sigma}{\bar{x}} \times 100 = \frac{10.96}{25.9} \times 100 = 42.32\ \%$$

Example: An analysis of production rejects resulted in the following data

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

Calculate standard deviation and coefficient of variance.

Solution: Using class boundaries we get

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
		N = 120		$\sum f_id_i$ = -25	$\sum f_id_i^2$= 223

Mean: $\bar{x} = A + \frac{\sum f_i d_i}{N} \times h = 38 + \frac{-25}{120} \times 10 = 36.96$

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

$$= \sqrt{1.858 - .043} \times 5 = 6.375$$

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

For Practice

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

ANS: $\bar{\mathbf{x}} = \mathbf{32.2}, \boldsymbol{\sigma} = \mathbf{13.862}, \mathbf{C.V} = \mathbf{43.05}\%$

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

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

ANS: $\bar{\mathbf{x}} = \mathbf{17.6}, \boldsymbol{\sigma} = \mathbf{7.806}, \mathbf{C.V} = \mathbf{44.35}\%$

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

ANS: $\bar{\mathbf{x}} = \mathbf{55.33}, \boldsymbol{\sigma} = \mathbf{14.83}, \mathbf{C.V} = \mathbf{26.8}\%$

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

ANS: $\bar{\mathbf{x}} = \mathbf{59.1}, \boldsymbol{\sigma} = \mathbf{17.56}, \mathbf{C.V} = \mathbf{29.71}\%$

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

ANS: $\bar{\mathbf{x}} = \mathbf{227.55}, \boldsymbol{\sigma} = \mathbf{8.732}, \mathbf{C.V} = \mathbf{3.84}\%$

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

ANS: $\bar{\mathbf{x}} = \mathbf{21.35}, \boldsymbol{\sigma} = \mathbf{6.375}, \mathbf{C.V} = \mathbf{29.86}\%$

Empirical Relation between Measures of variation

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

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

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

Example: Calculate standard deviation and then calculate mean deviation using empirical relation 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:

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

Standard deviation: $\sigma = \sqrt{\frac{\sum f_i\, d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h = \sqrt{\frac{66}{50} - \left(\frac{-14}{50}\right)^2} \times 10 = 11.14 \text{ lakh}$

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

Example: 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:

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	$35 + \frac{-8}{50} \times 10 = 33.4$	8.4	67.2
30-40	35	15	0	0		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		$\sum f_i d_i$ = -8			$\sum f_i\lvert x_i - \bar{x}\rvert$= 658.4

$$\mathrm{M.D} = \frac{\sum f_i |x_i - \bar{x}|}{N} = \frac{658.4}{50} = 13.168$$

Quartile deviation $\quad \mathrm{Q.D} = \frac{5}{6} \times \mathrm{M.D} = \frac{5}{6} \times 13.168 = 10.97$