Series A	***Series B***	***Series C***
200 200 200 200 200	200 205 202 203 190	1 989 2 3 5
Mean = 200	200	200

Measure of Dispersion

- *Scatteredness* (*homogeneity or heterogeneity*)

The degree to which numerical data tend to spread about an average value is called variation or dispersion of data.

***Measures of Dispersion**:* Range

Quartile Deviation

Mean Deviation

Standard Deviation

Absolute Measure of Dispersion:

Relative Measure of Dispersion:

Range

Definition: Difference between the value of the smallest item and the value of the largest item in the distribution.

$$\textbf{Range} = \boldsymbol{L - S}$$

L – Largest Value, S- Smallest Value

The *relative measure* corresponding to range is called the ***coefficient of range***,

$$\textbf{Coefficient of Range} = \boldsymbol{\frac{L-S}{L+S}}$$

Example 1: The following are the prices of shares of a company from Monday to Saturday:

Days	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Price(Rs.)	200	210	208	160	220	250

Calculate the range and its coefficient.

Solution: Range = $L - S$ = 250 – 160 = 90

Range = Rs. 90

Coefficient of Range = $\frac{L-S}{L+S} = \frac{250-160}{250+160} = \frac{90}{410} = 0.22$

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

Example 2:

Marks	10-20	20-30	30-40	40-50	50-60	60-70
No. of Students	12	18	27	20	17	6

Range = $L - S$ = 70 – 10 = 60

Coefficient of Range = $\frac{L-S}{L+S} = \frac{70-10}{70+10} = \frac{60}{80} = 0.75$

Quartile Deviation

Definition: Average amount by which the two quartiles differ from the median.

$$\textbf{Quartile Deviation (Q.D.)} = \frac{Q_3 - Q_1}{2}$$

- The Median $\pm$ Q.D. covers exactly 50 per cent of the observations.

- When Q.D. is very small, it describes high uniformity or small variation of the central 50% items, and a high Q.D. means that the variation among the central items is large.

Relative measure of Q.D.

Coefficient of Q.D. $= \frac{Q_3 - Q_1}{Q_3 + Q_1}$

It can be used to compare the degree of variation in different distributions.

Example 3: Calculate the value of Q.D. and its coefficient of Q.D. from the following data.

Roll No.	1	2	3	4	5	6	7
Marks	20	28	40	12	30	15	50

Solution: Marks in ascending order 12 15 20 28 30 40 50

Q_1 = Size of $\frac{N+1}{4}\,th$ item = Size of $\frac{7+1}{4}$ = 2^{nd} item.

Size of 2^{nd} item is 15. Hence $\boldsymbol{Q_1 = 15}$

Q_3 = Size of $3\left(\frac{N+1}{4}\right)\,th$ item = Size of $3\left(\frac{7+1}{4}\right)$= 6^{th} item.

Size of 6^{th} item is 40. Hence $\boldsymbol{Q_3 = 40.}$

.

$$\therefore \quad \boldsymbol{Q.D.} = \frac{Q_3 - Q_1}{2} = \frac{40-15}{2} = \boldsymbol{12.5}$$

$$\textbf{\textit{Coefficient of Q.D.}} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{40-15}{40+15} = \mathbf{0.455}$$

Example 4: Compute the value of Q.D. and its coefficient from the following data.

Marks	10	20	30	40	50	60
No. of Students	4	7	15	8	7	2

Solution:

Marks	*Frequency*	*cumulative frequency*
10	4	4
20	7	11
30	15	26
40	8	34
50	7	41
60	2	43

Q_1 = Size of $\frac{N+1}{4}\ th$ item = Size of $\frac{43+1}{4}$ = 11^{th} item.

Size of 11^{th} item is 20. Hence ***Q_1 = 20***

Q_3 = Size of $3\left(\frac{N+1}{4}\right)\ th$ item = Size of $3\left(\frac{43+1}{4}\right)$= 33^{rd} item.

Size of 33^{rd} item is 40. Hence ***Q_3 = 40.***

$$\boldsymbol{Q.D.} = \frac{Q_3 - Q_1}{2} = \frac{40-20}{2} = \boldsymbol{10}$$

$$\textbf{\textit{Coefficient of Q.D.}} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{40-20}{40+20} = \boldsymbol{0.333}$$

Example 4: Compute the value of Q.D. and coefficient of Q.D. from the following data

C.I.	10-20	20-30	30-40	40-50	50-60	60-70	70-80
f	12	19	5	10	9	6	6

Solution:

Marks	*Frequency*	*Cumulative Frequency*
10-20	12	12
20-30	19	31
30-40	5	36
40-50	10	46
50-60	9	55
60-70	6	61
70-80	6	67
	$\boldsymbol{N} = 67$	

Q_1 = Size of $\frac{N}{4}\, th$ item = Size of $\frac{67}{4}$ = 16.75^{th} item.

Q_1 lies in the interval **20-30**

$$Q_1 = l + \frac{\frac{N}{4} - c.f.}{f} \times i \qquad l = 20,\ N/4 = 16.75,\ c.f. = 12 \quad f = 19,\ i = 10$$

$$Q_1 = 20 + \frac{\frac{67}{4} - 12}{19} \times 10 = 20 + 2.5 = 22.5$$

Hence $\mathbf{Q_1 = 22.5}$

Q_3 = Size of $\frac{3N}{4}\,th$ item = Size of $\frac{3\times 67}{4}$ = 50.25th item.

Q_3 lies in the class **50-60**.

$$Q_3 = l + \frac{\frac{3N}{4} - c.f.}{f} \times i$$

l = 50, 3N/4 = 50.25, c.f. = 46 f = 9, i = 10

$$Q_3 = 50 + \frac{50.25-46}{9} \times 10 = 50+4.72 = 54.72$$

Hence **Q_3 = 54.72**

$$\boldsymbol{Q.D.} = \frac{Q_3-Q_1}{2} = \frac{54.72-22.5}{2} = \boldsymbol{16.11}$$

$$\textbf{\textit{Coefficient of Q.D.}} = \frac{Q_3-Q_1}{Q_3+Q_1} = \frac{54.72-22.5}{54.72+22.5} = \boldsymbol{0.4172}$$

Mean Deviation

Definition: M.D. is the average difference between the observations in a distribution and the *median* or *mean* of that series.

$$\text{Mean Deviation or M.D.} = \frac{\sum|D|}{N}$$

Where $|D|$ is the deviations from ***median*** ignoring signs.

For individual observations:

(i) Compute median of the series.

(ii) Take deviations of items from median ignoring $\pm$ signs and denote these deviations by $|D|$.

(iii) Obtain the total of these observations, $\sum|D|$.

(iv) Divide the total obtained in step (iii) by the number of observations to get the value of ***mean deviation***.

Relative Measure of M.D. :

$$\text{Co-efficient of M.D.} = \frac{M.D.}{Median}$$

Example 5: Calculate the mean deviation of the two income groups.

I (Rs.)	**II** (Rs.)
4000	3000
4200	4000
4400	4200
4600	4400
4800	4600
	4800
	5800

Solution:

Group I		Group II	
Rs.	$\|D\|$	**Rs.**	$\|D\|$
4000	400	3000	1400
4200	200	4000	400
4400	0	4200	200
4600	200	4400	0
4800	400	4600	200
		4800	400
		5800	1400
$N = 5$	$\sum\|D\| = 1200$	$N = 7$	$\sum\|D\| = 4000$

Mean deviation : **I group** M.D. = $\frac{\sum|D|}{N}$

Median = $\frac{N+1}{2}$ th item = $\frac{5+1}{2}$ = 3rd item. Size of the 3rd item = 4400.

$$\text{M.D.} = \frac{1200}{5} = \mathbf{240}$$

i.e., the average deviation of the individual incomes from the median income is **Rs. 240**.

Mean deviation : **II group** M.D. = $\frac{\sum|D|}{N}$

Median = $\frac{N+1}{2}\ th$ item = $\frac{7+1}{2}$ = 4th item.

Size of the 4th item = 4400.

M.D. = $\frac{4000}{5}$ = 571.43

i.e., the average deviation of the individual incomes from the median income is **Rs. 571.43**.

Co-efficient of M.D. (I - Group) = $\frac{M.D.}{Median} = \frac{240}{4400} = 0.055$

(II- Group) = $\frac{571.43}{4400} = 0.13$

Mean deviation – Discrete Series

$$\text{M.D.} = \frac{\sum f|D|}{N}$$

(i) Compute median of the series.

(ii) Take deviations of items from median ignoring $\pm$ signs and denote these deviations by $|D|$.

(iii) Multiply these deviations by the respective frequencies and Obtain the total, $\sum f|D|$.

(iv) Divide the total obtained in step (iii) by the number of observations to get the value of ***mean deviation***.

Example 6: The number of telephone calls received at an exchange in 245 successive one-minute intervals are shown in the following frequency distribution. Compute the mean deviation about the median.

Number of Calls	0	1	2	3	4	5	6	7
Frequency	14	21	25	43	51	40	39	12

Solution:

No. of Calls	*f*	*c.f.*	$\lvert D \rvert$	$f\lvert D \rvert$
0	14	14	4	56
1	21	35	3	63
2	25	60	2	50
3	43	103	1	43
4	51	154	0	0
5	40	194	1	40
6	39	233	2	78
7	12	245	3	36
	N = 245			$\boldsymbol{\sum f\lvert D \rvert = 366}$

Median = Size of $\frac{N+1}{2}$ th item = $\frac{245+1}{2}$ = 123rd item .

Hence the median value is 4.

M.D. $= \frac{\sum f|D|}{N} = \frac{366}{245} = \mathbf{1.49}$

In Continuous Series:

We have to obtain the mid-points of the various classes and take the deviations of these **mid-points** from median.

Calculate the coefficient of mean deviation from the following data:

Marks	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No. of Students	2	6	12	18	25	20	10	7

Mean deviation = 12.94

Standard deviation

For the frequency distribution $x_i \mid f_i$; i =1, 2, …, n ,

$$Variance = \sigma^2 = \frac{\sum(x_i - \bar{x})^2}{N} = \frac{1}{N}\sum x_i^2 - \left(\frac{1}{N}\sum x_i\right)^2 = \frac{1}{N}\sum x_i^2 - \bar{x}^2$$

(ii) Discrete or Continuous frequency distribution

$$\sigma^2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2$$

$$= \frac{1}{N}\sum f_i x_i^2 - \left(\frac{1}{N}\sum f_i x_i\right)^2 \ \textbf{or} \ \frac{1}{N}\sum f_i d_i^2 - \left(\frac{1}{N}\sum f_i d_i\right)^2$$

$$Standard\ deviation = \sqrt{variance} \quad \sigma = \sqrt{\frac{1}{N}\sum (x_i - \bar{x})^2} \quad \textbf{or} \quad \sqrt{\frac{1}{N}\sum f_i (x_i - \bar{x})^2}$$

$\boldsymbol{\bar{x}}$ − Arithmetic mean of the distribution

Coefficient of Variation

Coefficient of Variation : C.V. = $\frac{\sigma}{\bar{x}}$ x 100 (Relative Measure)

3) The score of two players A and B in ten innings during a certain season are:

A	32	28	47	63	71	39	10	60	96	14
B	19	31	48	53	67	90	10	62	40	80

Find which of the two players A, B is more consistent in scoring.

Solution: Calculation of Coefficient of Variation

X	$(X-\bar{X})$	$(X-\bar{X})^2$
32	-14	196
28	-18	324
47	+1	1
63	+17	289
71	+25	625
39	-7	49
10	-36	1296
60	+14	196
96	+50	2500
14	-32	1024
$\sum X$ = 460	0	6500

Y	$(Y-\bar{Y})$	$(Y-\bar{Y})^2$
19	-31	961
31	-19	361
48	-2	4
53	+3	9
67	+17	289
90	+40	1600
10	-40	1600
62	+12	144
40	-10	100
80	+30	900
$\sum Y$ =500	0	5968

$$\bar{X} = \frac{460}{10} = 46$$

$$\sigma_A^2 = \frac{\sum(x_i-\bar{x})^2}{N} = \frac{6500}{10} = 650$$

$$\bar{Y} = \frac{500}{10} = 50$$

$$\sigma_B^2 = \frac{\sum(y_i-\bar{y})^2}{N} = \frac{5968}{10} = 596.8$$

$$\sigma_A = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N}} = 25.5 \qquad \sigma_B = \sqrt{\frac{\sum(y_i - \bar{y})^2}{N}} = 24.43$$

$$\text{C.V.}_{(A)} = \frac{\sigma_A}{\bar{x}} \text{ x } 100 = 55.43 \qquad \text{C.V.}_{(B)} = \frac{\sigma_B}{\bar{y}} \text{ x } 100 = 48.86$$

$$\sum X = 460 \ ; \ \sum (X_i - \bar{x}) = 0 \ ; \sum (X_i - \bar{x})^2 = 6500$$

$$\sum Y = 500 \, ; \ \sum (Y_i - \bar{y}) = 0 \ ; \sum \left(Y_j - \bar{y}\right)^2 = 5968$$

$\sigma_A = 25.5$ $\sigma_B = 24.43$

C.V.$_{(A)}$ = 55.43 C.V.$_{(B)}$ = 48.86

4) Suppose that samples of polythene bags from two manufacturers, A and B, are tested by a prospective buyer for bursting pressure, with the following results:

Bursting Pressure (lb.)	***Number of Bags***	
	A	***B***
5.0 – 9.9	2	9
10.0 – 14.9	9	11
15.0 – 19.9	29	18
20.0 – 24.9	54	32
25.0 – 29.9	11	27
30.0 – 34.9	5	13

Which set of bags has the highest average burning pressure?
Which has more uniform pressure? If prices are the same,
which manufacturer's bags would be preferred by the buyer? **Why**?

For Manufacturer A

Bursting Pressure (lb.)	$\boldsymbol{m}$	$\boldsymbol{f}$	$\left(\frac{m-17.45}{5}\right)$ $\boldsymbol{d}$	$\boldsymbol{fd}$	$\boldsymbol{fd^2}$
4.95-9.95	7.45	2	-2	-4	4
9.95-14.95	12.45	9	-1	-9	9
14.95-19.95	17.45	29	0	0	0
19.95-24.95	22.45	54	1	54	54
24.95-29.95	27.45	11	2	22	44
29.95-34.95	32.45	5	3	15	45
		N = 110		$\sum \boldsymbol{fd}$**=78**	$\sum \boldsymbol{fd^2}$**=160**

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

Here, $A = 17.45, \sum fd$=78, $N = 110$, $i = 5$

$$\bar{X}_A = 17.45 + \frac{78}{110} \times 5 = 21$$

$$\sigma_{\boldsymbol{A}} = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times i$$

$$= \sqrt{1.455 - 0.503} \times 5 = 4.88$$

$$\text{C.V.} = \frac{\sigma_A}{\bar{x}} \text{ x } 100 = 23.24\%$$

For Manufacturer B

Bursting Pressure (lb.)	$\boldsymbol{m}$	$\boldsymbol{f}$	$\left(\frac{m-17.45}{5}\right)$ $\boldsymbol{d}$	$\boldsymbol{fd}$	$\boldsymbol{fd^2}$
4.95-9.95	7.45	9	-2	-18	36
9.95-14.95	12.45	11	-1	-11	11
14.95-19.95	17.45	18	0	0	0
19.95-24.95	22.45	32	+1	+32	32
24.95-29.95	27.45	27	+2	+54	108
29.95-34.95	32.45	13	+3	+39	117
		N = 110		$\boldsymbol{\sum fd = 96}$	$\boldsymbol{\sum fd^2 = 304}$

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

Here, $A = 17.45, \sum fd = 96, N = 110, i = 5$

$$\bar{X}_B = 17.45 + \frac{96}{110} \times 5 = 21.81$$

$$\sigma_{\boldsymbol{B}} = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times i$$

$$= \sqrt{2.764 - 0.762} \times 5 = 7.075$$

$$\text{C.V.} = \frac{\sigma_B}{\bar{x}} \text{ x } 100 = 32.44\%$$

$\bar{X}_A = 21$ $\bar{X}_B = 21.81$

$\sigma_A = 4.88$ $\sigma_B = 7.07$

C.V.$_{(A)}$ = 23.24% C.V.$_{(B)}$ = 32.44%

Since the average bursting pressure is higher for manufacturer B, the bags of manufacturer B have higher bursting pressure.