Measures of dispersion

All measures of dispersion can be divided into two categories:

(*i*) Measures of spread between certain specified values. They include range, inter-quartile range and quartile deviation;

(*ii*) Measures computed from the deviation of every variable from a measure of central tendency. They include mean deviation from median, mean deviation from mean and the standard deviation.

The above measures can be further classified into two categories:

(*i*) Absolute Measures

(*ii*) Relative measures

Absolute Measures

Such measures of dispersion that are expressed in their original units are called absolute measures. They are not suitable for comparing dispersion of more than one data set. They are:

(*i*) Range

(*ii*) Inter-quartile Range

(*iii*) Quartile Deviation or *Q.D.*

(*iv*) Mean Deviation or *M.D.*

(*v*) Standard Deviation or S.D.

Relative Measures

A relative measure of dispersion is suitable for comparing the extent of dispersion in more than one data set. It is expressed in pure number. They are:

(*i*) Coefficient of Range

(*ii*) Coefficient of *Q.D.*

(*iii*) Coefficient of *M.D.*

(*iv*) Coefficient of Variation or C.V.

An important point about all these absolute and relative measures is that all of them give answer in positive numbers.

RANGE

Range is the difference between the highest and the lowest observed values. Greater the range, larger is the variability among data set. Range is often used when the number of observations is less (generally ten or fewer observations). It is calculated by the following formula:

Range = $L - S$

Where L = Largest value and S = Smallest value.

Uses of Range

Range is often used in quality control and in the construction of control charts (Range Chart), stock price quotations, measurement of risk in investment, information about weather like maximum and minimum temperatures, etc. Of late, range is used in other areas like range of marks of students. Just look at the information relating to the range of marks of candidates admitted to IITs in the year 2008 under various categories:

Category	*Total marks Max-Min*
General	433-180
OBC	347-173
SC	322-104
ST	292-104

Source: TOI, New Delhi, 8.8.2008.

Since the focus of range is on the two extreme data and it does not consider other data, therefore, it is treated as a crude measure of dispersion and it is not very popular in real business applications.

Mathematical properties of Range

If a constant is added (or subtracted) to all the values, range is not affected but if all the values are multiplied (or divided) by a constant range is multiplied (or divided) by the constant. Thus, range is independent of origin but not of scale.

Coefficient of Range

The relative measure of range is known as coefficient of range. Its formula is:

$$\text{Coefficient of range} = \frac{L-S}{L+S}$$

EXAMPLE 1. A sugar mill packs sugar in packets of one kilogram for retail marketing. Ten packets were checked at random and their weights were found as under:

Weights (in kg.) : 1.050, 1.280, 0.995, 0.987, 1.116, 1.085, 0.990, 0.955, 1.205, 0.980.

Calculate range and coefficient of range.

SOLUTION:

$$\text{Range} = L - S$$
$$= 1.280 - 0.955$$
$$= 0.325 \text{ kg.}$$

$$\text{Coefficient of range} = \frac{L-S}{L+S} = \frac{1.280 - 0955}{1.280 + 0.955} = \frac{0.325}{2.235} = 0.145$$

EXAMPLE 2. Popular Mutual Fund issued units of Rs.10 each and the Net Present Value (NPV) during the last 12 months was recorded as follows:

Month:	1	2	3	4	5	6	7	8	9
NPV (Rs.)	11	9	15	22	13	9	8	9	13

Month:	10	11	12
NPV (Rs.)	20	24	17

Calculate the range of the above NPVs.

SOLUTION: Range = $L - S = 24 - 8 =$ Rs.16.

EXAMPLE 3. 8 members of the award staff were promoted as Manager – 1 in a bank. After their promotion, they attended a training programme in the Training College of the bank. They were evaluated after the training and obtained the following marks:

57, 84, 63, 34, 87, 69, 48, 72.

Compute range and coefficient of range.

SOLUTION: Range = $L - S = 87 - 34 = 53$ marks.

$$\text{Coefficient of range} = \frac{L-S}{L+S} = \frac{87-34}{87+34} = 0.438.$$

EXAMPLE 4. Export earnings of companies located in a particular Special Export Promotion Zone are mentioned in the following table:

Earnings (Rs. Crores) :	5 –10	10-15	15–20	20–25	25-30	30-35	35-40
No. of companies :	25	28	32	21	17	14	8

Calculate range and coefficient of range.

SOLUTION:

Range = $L - S = 40 - 5 =$ Rs. 35crores.

$$\text{Coefficient of range} = \frac{L-S}{L+S} = \frac{40-5}{40+5} = 0.78.$$

Merits of Range

(*a*) It is easy to calculate.

(*b*) It is not affected by frequencies.

(*c*) It is used in the application of control charts.

Demerits of Range

(*a*) It is based on highest and lowest value; thus, ignores other values and the pattern of the distribution. As a result some times two data set may have the same range but their composition may be different.

(*b*) It is influenced by extreme values.

(*c*) It is not suitable for open-ended distributions.

(*d*) It shows too much variation in different samples.

(*e*) It cannot be used for further algebraic manipulations.

Because of the above limitations range is seldom used alone in business applications.

INTER-QUARTILE RANGE or IQR

It measures the range of the middle fifty percent values. It is suitable for interval or ordinal data. Businessmen use *IQR* when they have to emphasize on the middle fifty per cent values. Usually realtors prefer to quote *IQR* to their customers. It is an absolute measure and it has no relative measure. The most important advantage of inter-quartile range is that it is not affected by extremes of the lower and higher values, thus, an improvement over range. Its main demerit is that it is unfit for further mathematical treatment. It is calculated by the following formula:

$IQR : Q_3 - Q_1$

EXAMPLE 5. For 120 students of 1st Trimester of a Business School Q_3 = 186.6 marks and Q_1 = 122 marks. Calculate *IQR* of marks.

SOLUTION:

$IQR = Q_3 - Q_1 = 186.6 - 122 = 64.6$ **marks.**

QUARTILE DEVIATION (*Q.D.*)

This is the average of the inter-quartile range. It is an absolute measure of dispersion. Like *IQR*, *Q.D* is suitable for interval and ordinal data. It is calculated by the following formula:

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

Coefficient of *Q.D.*

The relative measure of *Q.D.* is known as Coefficient of *Q.D.* and is computed by the following formula:

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

EXAMPLE 6. A group of 15 friends visit a departmental store. The variety and the manner in which the items are displayed compelled everybody to make some purchases. The amount spent (in Rs.) by them is shown below:

1134, 1095, 3387, 965, 2050, 3141, 2805, 4500, 1700, 5190, 1685, 1632, 5630, 2200, 3100.

Calculate Inter quartile Range, Quartile Deviation and Coefficient of *Q.D.*

SOLUTION:

S.No.	1	2	3	4	5	6	7	8	9	10	11
x :	965	1095	1134	1632	1685	1700	2050	2200	2805	3100	3141

S.No.	12	13	14	15
x :	3387	4500	5190	5630

$$Q_1 = \left(\frac{n+1}{4}\right)\text{th value} = 4^{th} \text{ value} = \text{Rs.}1632$$

$$Q_3 = \left(\frac{3(n+1)}{4}\right)\text{th value} = 12^{th} \text{ value} = \text{Rs.}3387$$

$$\text{Interquartile Range} = Q_3 - Q_1 = 3387 - 1632 = \text{Rs.}1755$$

$$Q.D. = \frac{Q_3 - Q_1}{2} = \frac{3387 - 1632}{2} = \text{Rs.}877.5$$

$$\text{Coefficient of } Q.D. = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{3387 - 1632}{3387 + 1632} = \frac{1755}{5019} = 0.35$$

EXAMPLE 7. Municipal Corporation of Delhi (MCD) collected the following amount of tax from roaming traders in a busy market:

Amount of Tax (Rs.):	25	50	75	100	200	300
No. of traders:	135	120	90	60	55	40

From the above data of tax collection, calculate *Q.D.* and Coefficient of *Q.D.*

SOLUTION:

Amt. (Rs.) = x	*No. of traders = f*	*Cf*
25	135	135
50	120	255
75	90	345
100	60	405
200	54	459
300	40	499
	$n = 499$	

$$Q_1 = \left(\frac{n+1}{4}\right)\text{th item} = 125\text{th item} = \text{Rs. } 25$$

$$Q_3 = 3\left(\frac{n+1}{4}\right)\text{th item} = 375\text{th item} = \text{Rs. } 100$$

$$Q.D. = \frac{Q_3 - Q_1}{2} = \frac{100-25}{2} = \text{Rs. } 37.5$$

$$\text{Coeff.}Q.D. = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{100-25}{100+25} = 0.6$$

EXAMPLE 8. State Roadways recorded the following collection on different days during last year. Calculate quartile deviation and coefficient of quartile deviation from the following data:

Daily Collection (Rs. Crores)	4-9	10-14	15-19	20-24	25-29	30-34
No. of days	66	72	86	53	49	30

SOLUTION:

Collection = x	*No. of days = f*	*Cf*
4.5 - 9.5	66	66
9.5 – 14.5	72	138
14.5 – 19.5	86	224
19.5 – 24.5	53	277
24.5 – 29.5	49	326
29.5 – 34.5	30	356

$$Q_1 = L + \frac{n/4 - Cf}{f} \times i = 9.5 + \frac{89-66}{72} \times 5 = \text{Rs.}11.1 \text{ crores}$$

$$Q_3 = L + \frac{3n/4 - Cf}{f} \times i = 19.5 + \frac{267-224}{53} \times 5 = \text{Rs.}23.56 \text{ crores}$$

$$Q.D. = \frac{Q_3 - Q_1}{2} = \frac{23.56 - 11.1}{2} = \text{Rs.}6.23 \text{ crores}$$

$$\text{Coeff.of } Q.D. = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{23.56 - 11.1}{23.56 + 11.1} = \frac{12.46}{34.66} = 0.36$$

EXAMPLE 9. If the first quartile is 142 and the Semi-interquartile Range is 18, find the Median (assuming the distribution to be symmetrical).

SOLUTION: $Q_1 = 142 \quad Q.D. = 18 \quad \therefore 18 = \frac{Q_3 - 142}{2}; \ \therefore Q_3 = 178.$

In a symmetrical distribution:

$Q_3 - \text{Median} = \text{Median} - Q_1$

$2\text{Median} = Q_3 + Q_1$

$$\therefore \text{Median} = \frac{Q_3 + Q_1}{2} = \frac{178 + 142}{2} = 160.$$

Merits of *Q.D.*

1. It is most suitable technique of dispersion when data are skewed and open-ended.
2. It is not affected by extremely large values in the data set as it is based on middle 50% values, thus, it is a better measure than range.

Limitations of *Q.D.*

1. It is affected by sampling fluctuation.
2. It is not fit for further mathematical calculations.
3. It ignores the first and the last quarter of values and not based on all observations.

The above limitations of quartile deviation restrict its applications in real business environment.

MEAN DEVIATION (*M.D.*)

Another measure of dispersion is known as Mean deviation. *Average of sum of absolute deviations from Median (or Mean) is called Mean deviation*. Usually, Median is subtracted from each observation then sum of the deviations is obtained after ignoring + or – signs. Therefore, some managers prefer to call it Mean Absolute Deviation (MAD). Mean deviation is least when deviations are taken from Median. That is the reason that statisticians prefer to take deviations from Median. It is a good measure of dispersion if further algebraic manipulations are not required. When specified, deviations are taken from Mean also. It is an absolute measure. It is suitable for interval scale data.

It is calculated by the following formula:

$$M.D. = \frac{\Sigma|D|}{n} \quad \text{Where } |D| = |x - \text{Median}|$$

In case of discrete or continuous series *M.D.* is calculated as follows:

$$M.D. = \frac{\Sigma f|D|}{n}$$

Coefficient of *M.D.*

The relative measure of *M.D.* is known as Coefficient of *M.D.* and calculated as:

$$\text{Coefficient of } M.D. = \frac{M.D.}{\text{Median}}$$

Sometimes *M.D.* is calculated based on Arithmetic Mean. In that case, the following changes are made in the formula:

$$M.D. = \frac{\Sigma f|D|}{n}$$

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

$$\text{Coefficient of } M.D. = \frac{M.D.}{\bar{x}}$$

Mean deviation is used to calculate errors in business forecasting. Otherwise it is of limited practical use because it is based on absolute deviations and unfit for further mathematical treatment. *M.D.* from Mean is independent of origin.

EXAMPLE 10. Nine firms deposited the following amount of income tax for the quarter April – June:

Firms:	A	B	C	D	E	F	G	H	I
Amount of tax deposited (in Rs. Lakhs):	5.0	6.7	10.4	12.0	15.7	17.9	18.8	19.6	19.9

Calculate *M.D.* and Coefficient of *M.D.* from Median as well as Mean.

SOLUTION:

Firms	x	$\lvert D \rvert = \lvert x\text{-Med.} \rvert$	$\lvert D \rvert = \lvert x - \bar{x} \rvert$
A	5.0	10.7	9
B	6.7	9	7.3
C	10.4	5.3	3.6
D	12.0	3.7	2
E	15.7	0	1.7
F	17.9	2.2	3.9
G	18.8	3.1	4.8
H	19.6	3.9	5.6
I	19.9	4.2	5.9
$n = 9$	$\Sigma x = 126$	$\Sigma\lvert D \rvert = 42.1$	$\Sigma\lvert D \rvert = 43.8$

$$\text{Median} = \left(\frac{n+1}{2}\right)\text{th value or } 5^{th} \text{ value i.e. Rs.15.7 Lakhs}$$

$$M.D. = \frac{\Sigma|D|}{n} = \frac{42.1}{9} = \text{Rs.4.7Lakhs}$$

$$\text{Coefficient of } M.D. = \frac{\text{M.D.}}{\text{Median}} = \frac{4.7}{15.7} = 0.3$$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{126}{9} = \text{Rs.14Lakhs}$$

$$M.D. = \frac{\Sigma|D|}{n} = \frac{43.8}{9} = \text{Rs.4.88Lakhs}$$

$$\text{Coefficient of } M.D. = \frac{M.D.}{\bar{x}} = \frac{4.88}{14} = 0.35$$

EXAMPLE 11. A survey was conducted on the inmates of a prison and the following data were found:

Period of Jail (in months):	6	10	12	20	24	36	40	48
No. of prisoners:	7	11	14	36	16	9	5	1

Calculate *M.D.* and Coefficient of *M.D.*

SOLUTION:

x	f	Cf	$\|D\|$	$f\|D\|$
6	7	7	14	98
10	11	18	10	110
12	14	32	8	112
20	36	68	0	0
24	16	84	4	64
36	9	93	16	144
40	5	98	20	100
48	1	99	28	28
	$n = 99$			$\Sigma f\|D\| = 656$

$$\text{Median} = \left(\frac{n+1}{2}\right)\text{th value} = \left(\frac{99+1}{2}\right)\text{th value} = 50\text{th value} = 20 \text{ months}$$

$$M.D. = \frac{\Sigma f|D|}{n} = \frac{656}{99} = 6.63 \text{ months}$$

$$\text{Coefficient of } M.D. = \frac{M.D.}{\text{Median}} = \frac{6.63}{20} = 0.33$$

EXAMPLE 12. A petrol selling station recorded the following sales in a day:

Sales (in Rs.):	0 – 100	100 – 200	200 – 300	300 – 400	400 – 500	500 – 600	600 – 700
No. of customers:	38	44	56	120	42	33	17

Calculate Mean deviation and Coefficient of Mean deviation; Range and Coefficient of Range; Quartile deviation and Coefficient of Quartile deviation for the above sales.

SOLUTION:

Sales	*f*	*Cf*	*Mid-Point = x*	$\|D\| = \|x\text{-Med.}\|$	$f\|D\|$
0 – 100	38	38	50	280.83	10,671.54
100 – 200	44	82	150	180.83	7,956.52
200 – 300	56	138	250	80.83	4,526.48
300 – 400	120	258	350	19.17	2,300.40
400 – 500	42	300	450	119.17	5,005.14
500 – 600	33	333	550	219.17	7,232.61
600 – 700	17	350	650	319.17	5,425.89
	$n = 350$				$\Sigma f\|D\| =$ 43,118.58

$$\text{Median} = L + \frac{\frac{n}{2} - Cf}{f} \times i$$

$$= 300 + \frac{175 - 138}{120} \times 100; \text{ Mean deviation} = \frac{\Sigma f|D|}{n} = \frac{43{,}118.58}{350} = \text{Rs.}123.2$$

$= \text{Rs.}330.83$

Coefficient of Mean deviation = $M.D.$/Median = 123.2/330.83 = 0.37

Range = L – S = 700 – 0 = Rs. 700

Note: Range can also be calculated for the Mid-values. In that case Range = 650 – 50 = Rs.600.

$$\text{Coefficient of Range} = \frac{L-S}{L+S} = \frac{700-0}{700+0} = 1$$

$$Q_1 = L + \frac{n/4 - Cf}{f} \times i = 200 + \frac{87.5 - 82}{56} \times 100 = \text{Rs.}209.82$$

$$Q_3 = L + \frac{\frac{3n}{4} - Cf}{f} \times i = 400 + \frac{262.5 - 258}{42} \times 100 = \text{Rs.}410.71$$

$$\text{Quartile Deviation} = \frac{Q_3 - Q_1}{2} = \frac{410.71 - 209.82}{2} = \text{Rs.}100.445$$

$$\text{Coefficient of } Q.D. = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{410.71 - 209.82}{410.71 + 209.82} = \frac{200.89}{620.53} = 0.32$$

Merits of *M.D.*

1. It is based on all items; thus, it is a better measure of dispersion than Range and Quartile deviation.
2. It is easy to calculate and understand.
3. It is not affected by extremely large values.

Demerits of *M.D.*

1. It is not suitable for open-ended data set.
2. It cannot be used for further algebraic manipulations.
3. It is based on absolute deviations which makes it unsound in comparison to standard deviation. This demerit also restricts its applications in business.

VARIANCE

Variance is another absolute measure of dispersion. It is the square of standard deviation. In other words, standard deviation is the positive square root of variance. It is capable of further mathematical manipulations. In some cases it is more useful than standard deviation. It is applied in the analysis of variance, ANOVA. Since variance is expressed in squared units and not easy to understand hence it is less popular than standard deviation.

If the two data sets have the same $\bar{x}$ combined variance can be calculated by the following formula:

$$\sigma_{1,2}^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1 + n_2}$$

where n_1, n_2 refer to the number of observations in the first and the second data set.

σ_1^2= variance of the first data set and σ_2^2= variance of the second data set.

(Note: The same logic can be applied for finding combined variance for the more than two series.)

Procedure of calculating standard deviation and variance

In case of individual series, calculate mean; take deviation of all values from the mean; square these deviations; total these squared deviations; divide this total by the number of observations; find the square root. This square root value is the standard deviation. The appropriate formula is:

$$\sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

$$\text{Variance} = \sigma^2 = \frac{\Sigma(x-\bar{x})^2}{n}$$

Coefficient of Variation

This relative measure is known as Coefficient of Variation (C.V.). The credit for Coefficient of Variation goes to Karl Pearson. For comparing variation in two or more distributions we need a relative measure that will give us a feel for the magnitude of the deviation relative to the magnitude of the mean. It is not expressed in absolute values but expressed in percentage. For calculating C.V., none of the variables should be negative, otherwise it cannot be calculated. For a much skewed distribution C.V. may be more than 100%. Symbolically,

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

EXAMPLE 13. The per kilogram retail prices of apples in Delhi in different months were found as under:

Month :	September	October	November	December	January	February	March
Prices :	60	50	35	45	48	55	64

Calculate Variance, Standard Deviation and Coefficient of Variation.

SOLUTION:

Month	*Prices = x*	$(x - \bar{x})^2$
September	60	81
October	50	1
November	35	256
December	45	36
January	48	9
February	55	16
March	64	169
$n = 7$	$\Sigma x = 357$	$\Sigma(x - \bar{x})^2 = 568$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{357}{7} = \text{Rs.}51$$

$$\sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} = \sqrt{\frac{568}{7}} = \sqrt{81.14} = \text{Rs.9 approx.}$$

$$\text{Variance} = 81.14 \text{ rupees}^2$$

$$\text{Coefficient of Variation} : \frac{\sigma}{\bar{x}} \times 100 = \frac{9}{51} \times 100 = 17.65\%$$

The above procedure is not very suitable when mean is in fractional value. In such a case, first find the sum of all observations; square all observations and total the squared observation. Apply the following formula to obtain standard deviation:

$$\sigma = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2} \; or \sqrt{\frac{\Sigma x^2}{n} - (\bar{x})^2}$$

EXAMPLE 14. Following is the earning per share (in Rupees) earned by an investor from eight scripts:

18, 11, 16, 9, 15, 13, 17, 8.

From these earnings calculate standard deviation by direct method

SOLUTION:

x	x^2
18	324
11	121
16	256
9	81
15	225
13	169
17	289
8	64
$\Sigma x = 107$	$\Sigma x^2 = 1{,}529$

$$\sigma = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2} = \sqrt{\frac{1{,}529}{8} - \left(\frac{107}{8}\right)^2} = \sqrt{191.125 - 178.891} = \sqrt{12.234} = \text{Rupees } 3.5.$$

Shortcut method

Standard deviation is also calculated by a shortcut method. From an assumed Mean, deviation of all values is taken (similar to the procedure used in computing Mean); total these deviations; square all deviations; total the squared deviation and apply the following formula:

$$\sigma = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}$$ Where $d = (x - A)$; A being assumed mean.

EXAMPLE 15. 9 randomly selected customers of a bank visited their original bank branch and spent the following amount of time (in minutes) to complete their business. Calculate Standard Deviation of the time spent by shortcut method.

11, 14, 22, 25, 28, 35, 40, 42, 50.

SOLUTION: Let assumed Mean be taken as 28.

x	$d = (x - A)$	d^2
11	– 17	289
14	– 14	196
22	– 6	36
25	– 3	9
28	0	0
35	7	49
40	12	144
42	14	196
50	22	484
$n = 9$	$\Sigma d = 15$	$\Sigma d^2 = 1{,}403$

$$\sigma = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2} = \sqrt{\frac{1,403}{9} - \left(\frac{15}{9}\right)^2} = \sqrt{155.89 - 2.78} = \sqrt{153.11} = 12.37 \text{ minutes}$$

In case of discrete series, the basic procedure remains the same as explained above but every term is multiplied by the corresponding frequency. The modified formulas are shown below.

$$\sigma = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{n}}$$

$$\sigma = \sqrt{\frac{\Sigma fx^2}{n} - \left(\frac{\Sigma fx}{n}\right)^2} \text{ or } \sqrt{\frac{\Sigma fx^2}{n} - (\bar{x})^2}$$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2}$$

EXAMPLE 16. A travel and tourist consultancy firm offers a variety of tour packages for different destinations in India both to the foreign tourists as well as domestic tourists. In the tourist season of the last winter the following bookings were made by the firm:

Tariff (Rs. thousands):	4	7	10	15	25	50
No. of bookings:	80	75	40	32	15	8

Calculate the following:

(*i*) Total amount of booking,
(*ii*) Mean amount per booking,
(*iii*) Standard Deviation per booking,
(*iv*) Coefficient of Variation.

SOLUTION:

x (Amt. in 000)	*f*	*fx*	$f(x-\bar{x})^2$
4	80	320	2,880
7	75	525	675
10	40	400	0
15	32	480	800
25	15	375	3,375
50	8	400	12,800
	$n = 250$	$\Sigma fx = 2,500$	$\Sigma f(x-\bar{x})^2 = 20,530$

(*i*) Total amount of booking = Rs.2,500 thousands or Rs.25,00,000.00

(*ii*) Mean amount per booking $\bar{x} = \frac{\Sigma fx}{n} = \frac{2500}{250} = 10$ or Rs.10,000.00

(*iii*) S.D. $= \sqrt{\frac{\Sigma f(x-\bar{x})^2}{n}} = \sqrt{\frac{20,530}{250}} = 9.062$ or Rs. 9062.00

(*iv*) C.V. $= \frac{\sigma}{\bar{x}} \times 100 = \frac{9.062}{10} \times 100 = 90.62\%$

EXAMPLE 17. Number of patients increased tremendously after every rainy season. They have to spend lot of money to buy medicines. This increases the sales of pharmaceutical companies. The amount of bills of medicines dispatched to retail chemists in the preceding July by a pharmaceutical company is given below:

Amount (Rs.000):	4	8	10	13	18	24	30
No. of Bills:	18	22	45	60	52	36	17

Calculate Standard Deviation and C.V. for the above data by direct method.

Solution:

x	f	fx	fx^2
4	18	72	288
8	22	176	1,408
10	45	450	4,500
13	60	780	10,140
18	52	936	16,848
24	36	864	20,736
30	17	510	15,300
	$n = 250$	$\Sigma fx = 3,788$	$\Sigma fx^2 = 69,220$

$$\sigma = \sqrt{\frac{\Sigma fx^2}{n} - \left(\frac{\Sigma fx}{n}\right)^2} = \sqrt{\frac{69,220}{250} - \left(\frac{3,788}{250}\right)} = \sqrt{276.88 - 229.58} = \sqrt{47.3} = \text{Rs.}6.88 \text{ thousand}$$

$$\bar{x} = \frac{\Sigma fx}{n} = \frac{3,788}{250} = \text{Rs.}15.15 \text{ thousand.}$$

$$C.V. = \frac{\sigma}{\bar{x}} \times 100 = \frac{6.88}{15.15} \times 100 = 45.41\%$$

EXAMPLE 18. Sizes of land holdings of small farmers in a district are given below. From these data calculate (*i*) Mean; (*ii*) Standard Deviation; (*iii*) Coefficient of Variation. Apply shortcut method.

Farm Size (Acres):	5	8	10	12	15	25	50	75
No. of Farms:	24	35	42	58	63	16	9	3

SOLUTION:

x	f	$d = (x - 25)$	fd	fd^2
5	24	– 20	– 480	9,600
8	35	– 17	– 595	10,115
10	42	– 15	– 630	9,450
12	58	– 13	– 754	9,802
15	63	– 10	– 630	6,300
25	16	0	0	0
50	9	25	225	5,625
75	3	50	150	7,500
	$n = 250$		$\Sigma fd = -2,714$	$\Sigma fd^2 = 58,392$

$$\bar{x} = A + \frac{\Sigma fd}{n} = 25 - \frac{2,714}{250} = 25 - 10.856 = 14.144 \text{ acres}$$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2} = \sqrt{\frac{58,392}{250} - \left(\frac{-2,714}{250}\right)^2} = \sqrt{233.57 - 117.85} = \sqrt{115.72} = 10.76 \text{ acres.}$$

$$C.V. = \frac{\sigma}{\bar{x}} \times 100 = \frac{10.76}{14.144} \times 100 = 76.07\%$$

Standard Deviation is calculated by the following formula in case of continuous series:

$$\sigma = \sqrt{\frac{\Sigma fd'^2}{n} - \left(\frac{\Sigma fd'}{n}\right)^2} \times i$$

Where $d' = \frac{x - A}{i}$ A = assumed mean and i = class interval

Steps required to calculate *S.D.* in case of continuous series:

(*i*) Find mid-point for each class;

(*ii*) Total the frequency column and call the total n;

(*iii*) Take step deviation d' as explained above;

(*iv*) Multiply d' by each corresponding frequency and total the column; call it Σfd'.

(*v*) Multiply fd' by d' to obtain fd'^2 and then total the column; call it $\Sigma fd'^2$.

(*vi*) Put these values in the above stated formula and then multiply the square root values by the class interval i.

EXAMPLE 19. The HRD section of a company conducted an examination of those candidates who applied for the post of computer operator. In all 150 candidates appeared in the written examination. Their scores, out of 100, are given in the following table. Calculate Standard Deviation and Coefficient of Variation of the scores.

Scores	*No. of candidates*
0-10	8
10-20	15
20-30	17
30-40	28
40-50	25
50-60	24
60-70	18
70-80	9
80-90	6

SOLUTION:

Score	*f*	*Mid-Point=x*	*d'*	*fd'*	*fd'²*
0-10	8	5	-4	– 32	128
10-20	15	15	-3	– 45	135
20-30	17	25	-2	– 34	68
30-40	28	35	-1	– 28	28
40-50	25	45	0	0	0
50-60	24	55	1	24	24
60-70	18	65	2	36	72
70-80	9	75	3	27	81
80-90	6	85	4	24	96
	$n = 150$			$\Sigma fd' = -28$	$\Sigma fd'^2 = 632$

$$\bar{x} = A + \frac{\Sigma fd'}{n} \times i = 45 - \frac{28}{150} \times 10 = 43.13$$

$$\sigma = \sqrt{\frac{\Sigma fd'^2}{n} - \left(\frac{\Sigma fd'}{n}\right)^2} \times i = \sqrt{\frac{632}{150} - \left(\frac{-28}{150}\right)^2} \times 10 = \sqrt{4.213 - 0.035} \times 10 = 20.44$$

$$C.V. = \frac{\sigma}{\bar{x}} \times 100 = \frac{20.44}{43.13} \times 100 = 47.39\%$$

EXAMPLE 20. Calculate Variance and Coefficient of Variation for the following data:

Weights	*No. of students*
60-64	10
65-74	40
75-79	35
80-84	22
85-94	26
95-99	10
100-104	7

Solution: Before further calculation, the series should be converted into an exclusive series.

Weights	*No. of students* *f*	*Mid-point* *x*	*A = 82* *d=(x – A)*	*fd*	*fd²*
59.5- 64.5	10	62	-20	–200	4,000
64.5- 74.5	40	69.5	-12.5	–500	6,250
74.5- 79.5	35	77	-5	–175	875
79.5- 84.5	22	82	0	0	0
84.5- 94.5	26	89.5	+7.5	+195	1,462.5
94.5- 99.5	10	97	15	+150	2,250
99.5-104.5	7	102	20	+140	2,800
	$n = 150$			$\Sigma fd = -390$	$\Sigma fd^2 = 17,637.5$