Standard Deviation:

Without frequency

If variable X takes values x_1 , x_2 , x_3 … … … x_n with frequencies then

Variance $\sigma_x^2 = \frac{\sum(x-\bar{x})^2}{n}$

Standard Deviation $\sigma_x = \sqrt{\frac{\sum (x-\bar{x})^2}{n}}$

With frequency

If variable X takes values x_1 , x_2 , x_3 … … … x_n with frequencies f_1 , f_2 , f_3 … … … f_n then

Variance $\sigma_x^2 = \frac{\sum f (x-\bar{x})^2}{\sum f}$

Standard Deviation $\sigma_x = \sqrt{\frac{\sum f (x-\bar{x})^2}{\sum f}}$

Computational Formulae

Without Frequency

Standard Deviation $\sigma_x = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2} = \sqrt{\frac{\sum x^2}{n} - (\frac{\sum x}{n})^2}$

With Frequency

Standard Deviation $\sigma_x = \sqrt{\frac{\sum f x^2}{\sum f} - (\bar{x})^2} = \sqrt{\frac{\sum f x^2}{\sum f} - (\frac{\sum f x}{\sum f})^2}$

Coefficient of Variation = C.V. = $\frac{\sigma_x}{\bar{x}}$ x 100

Q1. Calculate Standard Deviation for the following data

	X	X^2
	8	64
	9	81
	15	225
	23	529
	5	25
	11	121
	19	361
	8	64
	10	100
	12	144
Total	**120**	**1714**

$$\sigma_x = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2} \qquad \bar{X} = \frac{\sum x}{n} = \frac{120}{10} = 12$$

$$= \sqrt{\frac{1714}{10} - (12)^2}$$

$$= \sqrt{171.4 - 144}$$

$$= \sqrt{27.4} = 5.23$$

Q2. Calculate Standard Deviation and coefficient of variation for the following data

No.of Decayed Teeth	No. of Children	fx	fx^2
0	8	0	0
1	4	4	4
2	2	4	8
3	2	6	18
4	1	4	16
5	1	5	25
6	0	0	0
7	0	0	0
8	0	0	0
9	1	9	81
10	1	10	100
	20		252

$\sigma_x = \sqrt{\frac{\sum fx^2}{\sum f} - (\bar{x})^2}$ $\bar{X} = \frac{\sum fx}{n} = \frac{42}{20} = 2.1$

$= \sqrt{\frac{252}{20} - (2.1)^2}$

$= \sqrt{12.6 - (2.1)^2}$

$= \sqrt{8.19} = 2.86$

Coefficient of Variation = C.V. = $\frac{\sigma_x}{\bar{x}}$ x 100

$= \frac{2.86}{2.1}\ x\ 100$

= 136.19%

Q3. Calculate Standard Deviation and coefficient of variation for the following data

x	f	fx	fx^2
20	5	100	2000
30	8	240	7200
40	12	480	19200
50	9	450	22500
60	7	420	25200
70	5	350	24500
80	2	160	12800
90	2	180	16200
	50	2380	129600

$\sigma_x = \sqrt{\frac{\sum fx^2}{\sum f} - (\bar{x})^2}$ $\bar{X} = \frac{\sum fx}{n} = \frac{2380}{50} = 47.6$

$= \sqrt{\frac{129600}{50} - (47.6)^2}$

$= \sqrt{2592 - (47.6)^2}$

$= \sqrt{326.24} = 18.062$

Coefficient of Variation = C.V. = $\frac{\sigma_x}{\bar{x}}$ x 100

$= \frac{18.062}{47.6}\ x\ 100$

$= 37.94\%$

Q4. Calculate Standard Deviation and coefficient of variation for the following data

Age in years	No.of persons	x	fx	fx^2
0 -10	1	5	5	25
10 – 20	2	15	30	450
20 – 30	3	25	75	1875
30 – 40	2	35	70	2450
40 - 50	2	45	90	4050
Total	10		270	8850

$\sigma_x = \sqrt{\frac{\sum fx^2}{\sum f} - (\bar{x})^2}$ $\bar{X} = \frac{\sum fx}{n} = \frac{270}{10} = 27$

$= \sqrt{\frac{8850}{10} - (27)^2}$

$= \sqrt{885 - (27)^2}$

$= \sqrt{156} = 12.48$

Coefficient of Variation = C.V. = $\frac{\sigma_x}{\bar{x}}$ x 100

$= \frac{12.48}{27}\ x\ 100$

= 46.22%

Q5. Calculate Standard Deviation and coefficient of variation for the following data

Marks	No. of students	x	fx	fx^2
0 - 5	2	2.5	5	12.5
5 – 10	5	7.5	37.5	281.5
10 – 15	7	12.5	87.5	1093.75
15 – 20	13	17.5	227.5	3981.25
20 – 25	21	22.5	472.5	10631.25
25 – 30	16	27.5	440	12100
30 – 35	8	32.5	260	8450
35 – 40	3	37.5	112.5	4218.75
Total			1642.5	40768.75

$\sigma_x = \sqrt{\frac{\sum f x^2}{\sum f} - (\bar{x})^2}$ $\bar{X} = \frac{\sum f x}{\sum f} = \frac{1642.50}{75} = 21.9$

$= \sqrt{\frac{40768.75}{75} - (21.9)^2}$

$= \sqrt{543.58 - 479.61}$

$= \sqrt{63.97} = 7.99$

Coefficient of Variation = C.V. = $\frac{\sigma_x}{\bar{x}}$ x 100

$= \frac{7.99}{21.9}\ x\ 100$

= 36.48%

Q6. The Scores of 2 batsmen in an over is recorded as follows. Find which one has consistent scores.

Balls	Scores of Batsman A	Scores of Batsman B		
	X	Y	X^2	Y^2
1	4	3	16	9
2	6	4	36	16
3	6	2	36	4
4	1	3	1	9
5	0	4	0	16
6	6	2	36	4
Total	23	18	125	58

$\sigma_x = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}$ $\bar{X} = \frac{\sum x}{n} = \frac{23}{6} = 3.83$

$$= \sqrt{\frac{125}{6} - (3.83)^2}$$

$$= \sqrt{20.83 - 14.67}$$

$$= \sqrt{6.16} = 2.48$$

Coefficient of Variation = C.V. = $\frac{\sigma_x}{\bar{x}}$ x 100

$= \frac{2.48}{3.83}\, x\, 100$

= 64.75%

$\sigma_Y = \sqrt{\frac{\sum Y^2}{n} - (\bar{Y})^2}$ $\bar{Y} = \frac{\sum y}{n} = \frac{18}{6} = 3$

$= \sqrt{\frac{58}{6} - (3)^2}$

$= \sqrt{9.66 - 9}$

$= \sqrt{0.66} = 0.81$

Coefficient of Variation = C.V. = $\frac{\sigma_x}{\bar{Y}}$ x 100

$= \frac{0.81}{3}\ x\ 100$

= 27.0%

As Coefficient of Variation for Batsman B is less , Variable Y , Score of Batsman B is more consistent.

Q7. The Sales of 2 stores for a week is recorded as follows. Find which of these stores has consistent sales.

Weekday	Sales in Sore I	Sales in Sore II		
	X	Y	X^2	Y^2
1	50	90	2500	8100
2	30	80	900	6400
3	40	40	1600	1600
4	60	10	3600	100
5	20	10	400	100
6	50	20	2500	400
Total	250	250	11500	16700

$\sigma_x = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}$ $\bar{X} = \frac{\sum x}{n} = \frac{250}{6} = 41.66$

$$= \sqrt{\frac{11500}{6} - (41.66)^2}$$

$$= \sqrt{1916.66 - 1735.55}$$

$$= \sqrt{181.11} = 13.46$$

Coefficient of Variation = C.V. = $\frac{\sigma_x}{\bar{x}}$ x 100

$$= \frac{13.46}{41.66}\ x\ 100$$

$$= 32.31\%$$

$\sigma_Y = \sqrt{\frac{\sum Y^2}{n} - (\bar{Y})^2}$ $\bar{Y} = \frac{\sum y}{n} = \frac{250}{6} = 41.66$

$$= \sqrt{\frac{16700}{6} - (41.66)^2}$$

$$= \sqrt{2783.33 - 1735.55}$$

$$= \sqrt{1047.78} = 32.37$$

Coefficient of Variation = C.V. = $\frac{\sigma_x}{\bar{Y}}$ x 100

$$= \frac{32.37}{41.66}\ x\ 100$$

$$= 77.7\%$$

As Coefficient of Variation for X is less , Variable X , Sales in Sore I are more consistent.

Standard Deviation for the Combined Group

If we have two groups of n_1 and n_2 observations , with means $\overline{x_1}$ and $\overline{x_2}$ and standard deviations $\sigma_1$and $\sigma_2$respectively, then we know that the combined mean is given by

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

Let d1 = $\bar{x} - \overline{x_1}$ and d2 = $\bar{x} - \overline{x_2}$

$$\sigma = \sqrt{\frac{n_1\left(\sigma_1^2+d_1^2\right) \;\; +n_2(\sigma_1^2+d_2^2)}{n_1 + \; n_2}}$$

Q1. The following information about two factories is given below.

	Factory A	Factory B
Number	50	100
Means	120	85
Variance	9	16

i. Which factory has larger wage bill ?
ii. Which factory has greater variation ?
iii. Calculate the S.D. of wages of employees of both the factories taken together

i. <u>Wage Bill</u>

Wage Bill = Mean Wages * No. of employees
Factory A = 120 * 50 = 6000
Factory B = 85 * 100 = 8500
Factory B has larger Wage Bill

ii. <u>Variation</u>

C.V. = $\frac{S.D}{Mean} \times 100$

Factory A C.V. = $\frac{3}{120} \times 100$ = 2.5%

Factory B C.V. = $\frac{4}{85} \times 100$ = 4.7 %

Factory B has greater variation

iii. **<u>Combined S.D.</u>**

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

$$= \frac{50 * 120 + 100 * 85_2}{50+100}$$

$$= \frac{6000+8500}{50+100} = \frac{14500}{150} = 96.66$$

d1 = $\bar{x} - \overline{x_1}$ = -23.33 and d2 = $\bar{x} - \overline{x_2}$ = 11.67

$$\sigma = \sqrt{\frac{n_1\left(\sigma_1^2+d_1^2\right) + n_2(\sigma_1^2+d_2^2)}{n_1 + n_2}}$$

$$= \sqrt{\frac{50(9+23.33^2)+100(16+11.67^2)}{50+100}}$$

$$= \sqrt{\frac{50(553.29)+100(152.19)}{50+100}}$$

$$= \sqrt{\frac{27664.5 + 15219}{50 + 100}}$$

$$\sqrt{\frac{42883.5}{150}} = \sqrt{285.89} = 16.9$$

Q2. The mean and S.D. of group of 100 items are 80 and 5 respectively . In 2nd group consisting of 25 observations , where each value is 60 , Calculate mean and S.D. of 2 groups taken together .

	Group A	Group B
Number	100	25
Means	80	60
Variance	5	0

Combined Mean

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

$$= \frac{100 * 80 + 25 * 60}{100+25}$$

$$= \frac{8000+1500}{100+25} = \frac{9500}{125} = 76$$

Combined S.D

d1 = $\bar{x} - \overline{x_1}$ = - 4 and d2 = $\bar{x} - \overline{x_2}$ = 16

$$\sigma = \sqrt{\frac{n_1\,(\sigma_1^2+d_1^2)\ +n_2(\sigma_1^2+d_2^2)}{n_1 +\ n_2}}$$

$$= \sqrt{\frac{100(25+16)+25\,(0+256)}{125}}$$

$$= \sqrt{\frac{4100+6400}{125}} = \sqrt{84} = 9.17$$

Q3. From the group containing 100 observations with mean 8 and S.D. $\sqrt{10.5}$, 50 observations were selected. Mean and S.D. of these 50 observations were recorded as 10 & 2 respectively . Calculate mean and S.D. of remaining 50 observations .

Combined Mean

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

$$8 = \frac{50 * 10 + 50 * \overline{x_2}}{100}$$

$$800 = 500 + 50 * \overline{x_2}$$

$$\overline{x_2} = \frac{300}{50} = 6$$

Combined S.D

d1 = $\bar{x} - \overline{x_1}$ = 8 - 10 =-2 and d2 = $\bar{x} - \overline{x_2}$ = 8 – 6 = 2

$$\sigma = \sqrt{\frac{n_1\left(\sigma_1^2 + d_1^2\right) + n_2\left(\sigma_1^2 + d_2^2\right)}{n_1 + n_2}}$$

$$\sqrt{10.5} = \sqrt{\frac{50(4+4) + 50\left(\sigma_2^2 + 4\right)}{100}}$$

Squaring both sides

$10.5 * 100 = 50 * 8 + 50 * \sigma_2^2 + 50 * 4$

$1050 = 400 + 50 * \sigma_2^2 + 200$

$50 * \sigma_2^2 = 1050 - 400 - 200$

$50 * \sigma_2^2 = 450$

$\sigma_2^2 = 9$

$\sigma_2 = 3$

Q4. There are two groups containing 400 & 500 observations respectively. Mean and variance of the first group are 50 & 25 respectively and Mean for the second group is 41. Calculate S.D. of the second group , given the combined variance is 37 .

Combined Mean

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

$$= \frac{400 * 50 + 500 * 41}{900}$$

$$= \frac{20000+20500}{900} = \frac{40500}{900} = 45$$

Combined S.D

d1 = $\bar{x} - \overline{x_1}$ = 45 – 50 =-5 and d2 = $\bar{x} - \overline{x_2}$ = 45 – 41 = 4

$$\sigma = \sqrt{\frac{n_1\,(\sigma_1^2+d_1^2) \;\; +n_2(\sigma_1^2+d_2^2)}{n_1 + \; n_2}}$$

$$37 = \frac{400*50+500*\,\sigma_2^2\,+500*16}{400+500}$$

$37 * 900 = 20000 + 500 * \sigma_2^2 + 8000$

$333000 = 28000 + 500 * \sigma_2^2$

$500 * \sigma_2^2 = 5300$

$$\sigma_2^2 = \frac{5300}{500} = 10.6$$

$$\sigma_2 = 3.25$$