Unit - 1 Statistical Methods

* Measures of central tendencies (or Averages)

An average is a value that is typical, or representative, of a set of data. Here ~~typical values~~ tend to lie centrally within a set of data arranged according to averages are called measures of central tendency.

The commonly used measures of central values are Mean, Median, Mode, Geometric Mean, Harmonic Mean.

1) Mean (Arithmetic Mean)

* If $x_1, x_2, \ldots, x_n$ are a set of n values of a variable, then the mean is $\bar{x} = \frac{\Sigma x}{n}$

* In a frequency distribution, if $x_1, x_2, \ldots, x_n$ be the mid-values of the class intervals having frequencies $f_1, f_2, \ldots, f_n$ then $\bar{x} = \frac{\Sigma f x}{\Sigma f}$

(or) If $x_1, x_2, \ldots, x_n$ are n values & $f_1, f_2, \ldots, f_n$ are their weight (frequencies) then $\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \cdots + f_n x_n}{f_1 + f_2 + \cdots + f_n} = \frac{\Sigma f x}{\Sigma f} = \frac{\Sigma f x}{N}$

Here, $f_1 + f_2 + \cdots + f_n = N$ (total frequency)

① find the arithmetic mean of the numbers

a) 5, 3, 8, 7, 6, 1

Soln: Mean $= \frac{\Sigma x}{n} = \frac{5+3+8+7+6+1}{6} = \frac{30}{6} = 5$

b) 8, 3, 5, 12, 10

Soln: Mean $= \frac{8+3+5+12+10}{5} = \frac{38}{5} = 7.6$

—x—

② find the mean for the following table by 'direct method'.

Values of x	2	4	6	8	10
frequency (f)	3	5	6	4	2

Soln: Mean $= \frac{\Sigma fx}{\Sigma f} = \frac{2(3)+4(5)+6(6)+8(4)+10(2)}{20} = 5.7$

—x—

③ If 5, 8, 6 and 2 occur with frequencies 3, 2, 4 & 1 respectively then find the arithmetic mean by direct method.

Soln: Mean $= \frac{\Sigma fx}{\Sigma f} = \frac{5(3)+8(2)+6(4)+2(1)}{10} = 5.7$

—x—

* **Mean by Short-cut Method:**

Mean $= A + \frac{\Sigma fd}{\Sigma f}$

Where A - assumed mean
f - frequency
Σf - total frequency
d - deviation of x from A [ie) $d = x - A$]

① find the mean by short-cut method of the following table:

x	5	6	7	8	9	10
f	10	12	15	11	7	5

middle value of frequency table → (Assumed Mean) $A = 7$

Soln:

x	f	$d = x - A$	fd
5	10	−2	−20
6	12	−1	−12
7	15	0	0
8	11	1	11
9	7	2	14
10	5	3	15
	60		8

$\therefore$ Mean $= A + \frac{\Sigma fd}{\Sigma f}$
$= 7 + \frac{8}{60}$
$= 7.13$

—x—

Mean by Step-Deviation Method:

Mean $= A + \left(\frac{\Sigma fu}{\Sigma f}\right) C$

Where A = Assumed mean
$u = \frac{x - A}{C}$
C = class interval
Σf = Total frequency

(3) Find the mean of the following frequency table by step-deviation method:

Class	10-15	15-20	20-25	25-30	30-35	35-40
Frequency	5	6	8	12	6	3

Soln:

$\frac{x-27.5}{5}$

Class x	mid x	Frequency f	$u = \frac{x-A}{C}$	fu
10-15	12.5	5	-3	-15
15-20	17.5	6	-2	-12
20-25	22.5	8	-1	-8
*25-30	27.5	12	0	0
30-35	32.5	6	1	6
35-40	37.5	3	2	6
		40	Σ	-23

$\frac{25}{2} = 12.5 - 27.5 = 15/5 = -3$

$\frac{15+20}{2} - 27.5 = -2$

Assumed mean $(A) = \frac{25+30}{2} = \frac{55}{2} = 27.5$ | Class interval $(i) = 5$

Therefore mean $= A + \left(\frac{\Sigma fu}{\Sigma f}\right) i = 27.5 - \left(\frac{23}{40}\right) 5$

$= 24.63$

$\frac{10-27.5}{5}$

$\frac{25-27.5}{2}$

$12.5 - 27.5$

$\frac{15}{5}$

-x-

(8)

(3) find the mean for the following frequency distribution. (5)

Class	10 – 19	20 – 29	30 – 39	40 – 49	50 – 59	60 – 69	70 – 79	80 – 89
frequency	5	9	14	20	25	15	8	4

Soln:

class x	Middle x	frequency f	$U = \frac{x - A}{i}$ $\frac{x-44.5}{9}$	fU
10 – 19	14.5	5	−3.3	−16.5
20 – 29	24.5	9	−2.2	−19.8
30 – 39	34.5	14	−1.1	−15.4
40 – 49	44.5*	20	0	0
50 – 59	54.5	25	1.1	27.5
60 – 69	64.5	15	2.2	33
70 – 79	74.5	8	3.3	26.4
80 – 89	84.5	4	4.4	17.6
		100		52.8

$$\text{Mean } (\bar{x}) = A + \left(\frac{\Sigma fu}{\Sigma f}\right) i$$

$$= 44.5 + \left(\frac{52.8}{100}\right) 9$$

$$= 44.5 + \frac{475.2}{100} = 39.75$$

– x –

Median: The median of a set of numbers arranged in ascending (or descending) order of magnitude is the middle values. (or the mean of 2 middle values)

Median - Central value of distribution.

① find the median of the set of observations

27, 36, 28, 18, 35, 26, 20, 35, 40

Soln: The observations ~~are~~ in ascending order is

18, 20, 26, 26, 27, 28, 35, 35, 36, 40.

No. of observations = 10

$\therefore$ Median $= \frac{27+28}{2} = 27.5$

—x—

② find the median of the set of observations 1, 3, 5, 6, 8, 9, 10.

Soln: Median = 6

—x—

~~③ find the median for the following distribution.~~

* **Median for ungrouped data (or) discrete frequency distribution.**

a) If there are n observations ~~Σf~~ & N (Σf) is odd, then the median is $\left(\frac{N+1}{2}\right)^{th}$ observation.

b) If N is even, then the median is mean (average) of $\frac{N}{2}$ & $\left(\frac{N}{2}+1\right)^{th}$ observations.

Median for continuous frequency distribution.

$$\text{Median} = l + \frac{\frac{N}{2} - m}{f} \times c$$

Where l - lower limit of median class

N - Total frequency - ~~main~~

m - Cumulative frequency of pre-median class

f - frequency of median class

c - Uniform class interval

—x—

① find the median for following ~~frequency~~ distribution.

Weight ~~frequency~~	26	27	28	29	30	31
No. of students	3	5	6	7	6	4

Soln:

weights (x)	No. of students (f)	Cumulative frequency
26	3	3
27	5	8
28	6	14
29	7	21
30	6	27
31	4	31
	$\Sigma f = 31$	

Here $N = 31$ which is odd

$\therefore$ Mean $= \frac{N+1}{2} = \frac{31+1}{2} = 16$

Since 16 lies under cumulative frequency of 21

Median is 29.

2) Find the median for the following data

x	6	7	8	9	10	11	12
f	25	39	48	43	52	30	13

Soln:

x	f	Cumulative frequency
6	25	25
7	39	64
8	48	112
9	43	155 ✓
10	52	207
11	30	237
12	13	250
	$\Sigma f = 250$	

Here $N = 250$; which is even.

Median is the average of $\frac{N}{2}$th & $\left(\frac{N}{2}+1\right)$th items.

$\therefore$ Median $=$ Average of $\left(\frac{250}{2}\right)$ & $\left(\frac{250}{2}+1\right)$

$=$ Average of 125 & 126th observations.

$= 9$ $\left\{ \because 125 \text{ \& } 126 \text{ lie in the cumulative frequency of } 155 \right\}$

— x —

3) find the median for the following data:

class	0.5 - 10.5	10.5 - 20.5	20.5 - 30.5	30.5 - 40.5	40.5 - 50.5
frequency	3	7	13	17	12
class	50.5 - 60.5	60.5 - 70.5	70.5 - 80.5	80.5 - 90.5	90.5 - 100.5
frequency	10	8	8	6	6

Soln:

True class	frequency f	Mid x	Cumulative frequency (cf)
0.5 - 10.5	3	5.5	3
10.5 - 20.5	7	15.5	10
20.5 - 30.5	13	25.5	23
30.5 - 40.5	17	35.5	40 m
40.5 - 50.5	12	45.5	52
50.5 - 60.5	10	55.5	62
60.5 - 70.5	8	65.5	70
70.5 - 80.5	8	75.5	78
80.5 - 90.5	6	85.5	84
90.5 - 100.5	6	95.5	90
	Σf 90		

Here $\Sigma f = N = 90$ (Total frequency)

$N/2 = 90/2 = 45$ {Since 45 lies b/w ~~the class interval~~ 52 ~~40.5 - 50.5~~ }

Since l is the lower limit of median class

Here $l = 40.5$

m - Cumulative frequency of pre-median class

Here m = 40

c - class interval. So c = 10

Now, Median $= l + \dfrac{\frac{N}{2} - m}{f} \times c$

$= 40.5 + \dfrac{45-40}{12} \times 10$

$= 40.5 + \dfrac{50}{12}$

$= 40.5 + 4.17$

$= 44.67$

-x-

Geometric Mean (GM)

1) Geometric mean of the observations $x_1, x_2, \ldots, x_n$ is

$$GM = (x_1 \cdot x_2 \cdot \ldots \cdot x_n)^{1/n}$$

2) If $x_1, x_2, \ldots, x_n$ occur with frequency $f_1, f_2, \ldots, f_n$ respectively

then $GM = \left(x_1^{f_1} \cdot x_2^{f_2} \cdot \ldots \cdot x_n^{f_n}\right)^{1/N}$

where $N = f_1 + f_2 + \cdots + f_n$

Harmonic Mean (HM)

1) HM of a set of n observations $x_1, x_2, \ldots, x_n$ is

$$HM = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}$$

2) for frequency distribution, $HM = \dfrac{N}{\frac{f_1}{x_1} + \frac{f_2}{x_2} + \cdots + \frac{f_n}{x_n}}$; where $N = f_1 + f_2 + \cdots + f_n$

4) find the median for the following

class limits	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
frequency	129	115	241	117	52	10	6	3	2

Soln:

Class limits	frequency f	Cumulative frequency (cf)
0-5	129	129
5-10	115	224 (m)
10-15 (l)	241 (f)	465
15-20	117	582
20-25	52	634
25-30	10	644
30-35	6	650
35-40	3	653
40-45	2	655
	655	

Here $N = \frac{655}{2} = 327.5$

327.5 lies within class interval 10-15.

Since ~~K~~ 327.5 will be within cumulative frequency 465.

Here $l = 10$; m - premedian class $= 224$; $c = 5$; $f = 241$

$\therefore$ Median $= l + \dfrac{\frac{N}{2} - m}{f} \times c$

$= 10 + \dfrac{\frac{655}{2} - 224}{241} \times 5$

$= 10 + 2.15 = 12.15$

-x-

5) Calculate the GM & HM of the following data

3, 6, 24 and 48

Soln: a) $GM = (x_1 \cdot x_2 \cdot x_3 \cdot x_4)^{1/4}$

$= (3 \times 6 \times 24 \times 48)^{1/4}$

$= (20736)^{1/4}$

$= (3^4 \cdot 4^4)^{1/4} = 12$

b) $HM = \dfrac{4}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \frac{1}{x_4}}$

$= \dfrac{4}{\frac{1}{3} + \frac{1}{6} + \frac{1}{24} + \frac{1}{48}}$

$= \dfrac{4 \times 48}{16 + 8 + 2 + 1}$

$= \dfrac{192}{27}$

$= 7.11$

— x —

<u>Mode:</u> Mode is defined to be that value which occurs most often

<u>Mode for Continuous frequency distribution</u>:

$$\text{Mode} = l + \frac{f_1 - f_0}{2f_1 - (f_0 + f_2)} \times c$$

Where l - lower limit of the modal class.

f_1 - frequency of the modal class.

f_0 - frequency of the pre-modal class

f_2 - frequency of the post-modal class

c - class interval.

① find the mode of the following set of observations

3, 5, 7, 5, 9, 7, 5, 7, 6, 3, 9, 5, 6, 6, 3.

<u>Soln:</u> Here 5 appears maximum no. of times.

$\therefore$ Mode = 5.

—x—

② find the mode of the following distribution

	45.5-50.5	50.5-55.5	55.5-60.5	60.5-65.5	65.5-70.5	70.5-75.5	75.5-80.5
Class limits	46-50	51-55	56-60	61-65	66-70		
Frequency	2	3	5	7	9	11	7

2) find the mode of the following distribution:

Class limits	Frequency
45.5 - 50.5	2
50.5 - 55.5	3
55.5 - 60.5	5
65.5 - 70.5	9 (f_0)
70.5 - 75.5	11 → Modal class (f_1)
75.5 - 80.5	7 (f_2)
80.5 - 85.5	2
85.5 - 90.5	3
90.5 - 95.5	1

Soln: Here 11 = modal class {∵ the maximum largest no. in given frequency}

$l = 70.5$

$f_1 = 11$

$f_0 = 9$; $f_2 = 7$; $c = 5$

$$\text{Mode} = l + \frac{f_1 - f_0}{2f_1 - (f_0 + f_2)} \times c$$

$$= 70.5 + \frac{11-9}{2(11)-(9+7)} \times 5$$

$$= 70.5 + \frac{10}{6}$$

$$= 70.5 + 1.67$$

$$= 72.17$$

–x–

5) find the mode for the following data:

Class limits	Frequency
0.5 - 10.5	3
10.5 - 20.5	7
20.5 - 30.5	13 – f_0
30.5 - 40.5	17 → modal class f_1
40.5 - 50.5	12 – f_2
50.5 - 60.5	10
60.5 - 70.5	8
70.5 - 80.5	8
80.5 - 90.5	6
90.5 - 100.5	6

Soln:

$$\text{Mode} = l + \frac{f_1 - f_0}{2f_1 - (f_0 + f_2)} \times c$$

$$= 30.5 + \frac{17-13}{2(17)-(13+12)} \times 10$$

$$= 30.5 + \frac{40}{9}$$

$$= 34.94$$

–x–

Measures of Dispersion :- Describe how the observations are distributed

The measures of dispersion are

* Range * Quartile deviation
* Mean deviation * Standard deviation.

Range :

Range is defined to be the difference between the largest and the smallest of the observations.

i.e) Range $= L - S$ where L - Largest of observations
S - Smallest of observations

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

① find the range and the coefficient of range for the following data : 35, 40, 52, 29, 51, 46, 27, 30, 30, 23

Soln : Largest value $L = 52$
Smallest value $S = 23$

$$\text{Range} = L - S = 52 - 23 = 29$$

$$\text{Coefficient of Range} = \frac{L-S}{L+S} = \frac{52-23}{52+23} = \frac{29}{75}$$

—x—

② find the coefficient of range for the following

Wages	35 - 45 (40)	45 - 55 (50)	55 - 65 (60)	65 - 75 (70)	75 - 85 (80)
No. of workers	18	22	30	6	4

Soln: L = Mid value of last class = 80 (Largest value)

S = Mid value of first class = 40 (smallest value)

$$\text{Coefficient of Range} = \frac{L-S}{L+S} = \frac{80-40}{80+40} = 0.33$$

— x —

* **Quartile Deviation**

Quartiles are position values similar to the median.

Quartile deviation is defined as $\frac{Q_3 - Q_1}{2}$

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

Where Q_3 = upper quartile

Q_1 = lower quartile

$$\text{coefficient of Quartile Deviation} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

*) for Continuous frequency distribution:

$$Q_3 = l_3 + \frac{3\frac{N}{4} - m_3}{f_3} \times c$$

$$Q_1 = l_1 + \frac{\frac{N}{4} - m_1}{f_1} \times c$$

Where l_3 & l_1 = lower and upper limit of Q_3 & Q_1 class 18

m_3 & m_1 = cumulative frequency of the preceding class

f_3 & f_1 = frequency of Q_3 & Q_1 class

c = class interval

N = Total frequency.

Quartile for discrete frequency distribution:

$$Q_1 = \left(\frac{N+1}{4}\right)^{th} \text{ observation}$$

$$Q_3 = \text{value of } \frac{3(N+1)}{4}^{th} \text{ observation.}$$

① find the quartile deviation & the quartile coefficient of dispersion for the following data

class	Frequency
0-10	8
10-20	20
20-30	34
30-40	46
40-50	28
50-60	14
60-70	10

Soln:

17

class	Frequency	Cumulative frequency
0-10	8	8
10-20	20	28
20-30	34	62
30-40	46	108
40-50	28	136
50-60	14	150
60-70	10	160
	$\Sigma f = 160$	

$\frac{N}{4} = \frac{160}{4} = 40$

40 lies within cumulative frequency 62.

$\therefore l_1 = 20$

$m_1 = 28$

$f_1 = 34$

$c = 10$

Now, $Q_1 = l_1 + \left(\frac{\frac{N}{4} - m_1}{f_1}\right) \times c$

$= 20 + \left(\frac{40-28}{34}\right) \times 10 = 20 + \frac{120}{34} = 23.53$

$Q_3 = l_3 + \left(\frac{3\frac{N}{4} - m_3}{f_3}\right) \times c$; $3\frac{N}{4} = 3 \times 40 = 120$

120 lies within cf 136

$\therefore l_3 = 40$;

$m_3 = 108$

$f_3 = 28$

$= 40 + \left(\frac{120-108}{28}\right) \times 10$

$= 40 + \frac{120}{28}$

$= 44.29$

Now, Quartile Deviation $= \frac{Q_3 - Q_1}{2} = \frac{44.29 - 23.53}{2}$

$= 10.38$

Quartile coefficient $= \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{44.29 - 23.53}{44.29 + 23.53} = \frac{20.76}{67.82} = 0.$

-x-

Mean Deviation:

If ~~let~~ $x_1, x_2, \ldots, x_n$ be n values then the mean deviation about the mean ~~of~~ these values is

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

In a frequency distribution with frequencies $f_1, f_2, \ldots, f_n$ against the values $x_1, x_2, \ldots, x_n$ then

$$M.D \text{ about the mean} = \frac{\Sigma f_i |x_i - \bar{x}|}{N} \quad \text{where } \bar{x} = \frac{\Sigma f x}{N}$$

Where N = Total frequency * for class interval (continuous) $\Sigma \frac{f|x-\bar{x}|}{N}$ where $\bar{x} = A + \left[\frac{\Sigma fd}{N} \times C\right]$ *

Illy

~~M.D~~ M.D about Median $= \frac{\Sigma f_i |x_i - M|}{N}$ Where M=median

Relative Measure:

Coefficient of M.D about Mean $= \frac{M.D}{Mean}$

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

—x—

① find the mean deviation about the mean for the following data:

18, 20, 12, 14, 19, 22, 26, 16, 19, 24

Soln: Mean $\bar{x} = \frac{18+20+12+14+19+22+26+16+19+24}{10}$

$= \frac{190}{10}$

$= 19$

Mean deviation about mean $= \dfrac{\Sigma|x-\bar{x}|}{n}$

$$= \frac{1+1+7+5+0+3+7+3+0+5}{10} = \frac{32}{10} = 3.2$$

{ Since $|18-19| = |-1| = 1$; $|20-19| = |1| = 1$; $|12-19| = |-7| = 7$; $|14-19| = |-5| = 5$; .... ; $|24-19| = |5| = 5$ } .

— x —

② Find the mean deviation about mean for the following data:

Value (x)	10	11	12	13	14
frequency (f)	3	12	18	12	3

Soln :

x	f	fx	$\lvert x-\bar{x}\rvert$	$f\lvert x-\bar{x}\rvert$
10	3	30	2	6
11	12	132	1	12
12	18	216	0	0
13	12	156	1	12
14	3	42	2	6
	48	576		36

$$\bar{x} = \frac{\Sigma fx}{N} = \frac{576}{48} = 12$$

Mean deviation about mean $= \dfrac{\Sigma f|x-\bar{x}|}{N} = \dfrac{36}{48}$

$= 0.75$

— x —

3) find the mean deviation about mean for the following data 22

class	0-5	5-10	10-15	15-20	20-25
frequency	3	5	12	6	4

Soln:

class	Mid x	f	$d = \frac{x-A}{c} = \frac{x-12.5}{5}$	fd	$\lvert x-\bar{x}\rvert = \lvert x-13\rvert$	$f\lvert x-\bar{x}\rvert$
0-5	2.5	3	-2	-6	10.5	31.5
5-10	7.5	5	-1	-5	5.5	27.5
10-15	12.5	12	0	0	0.5	6.0
15-20	17.5	6	1	6	4.5	27.0
20-25	22.5	4	2	8	9.5	38.0
		30		$\Sigma fd = 3$		130.0

Here $A = 12.5$

Mean $\bar{x} = A + \frac{\Sigma fd}{N} \times c$

$= 12.5 + \frac{3}{30} \times 5$

$= 13$

$\therefore$ Mean deviation about mean $= \frac{\Sigma f\lvert x-\bar{x}\rvert}{N}$

$= \frac{130}{30}$

$= 4.33$

—x—

4) find the mean deviation about the mean for the following data

class	0-10	10-20	20-30	30-40	40-50	50-60	60-70
frequency	8	12	10	8	3	2	7

Soln:

Class	Mid x	f	$d = \frac{x-A}{c} = \frac{x-35}{10}$	fd	$\lvert x-\bar{x}\rvert = \lvert x-29\rvert$	$f\lvert x-\bar{x}\rvert$
0-10	5	8	-3	-24	24	192
10-20	15	12	-2	-24	14	168
20-30	25	10	-1	-10	4	40
30-40	* 35	8	0	0	6	48
40-50	45	3	1	3	16	48
50-60	55	2	2	4	26	52
60-70	65	7	3	21	36	252
		50		-30		800

$\bar{x} = A + \frac{\Sigma fd}{N} \times c$

$= 35 + \left(\frac{-30}{50}\right) \times 10$

$= 29$

Mean deviation about mean $= \frac{\Sigma f\lvert x-\bar{x}\rvert}{N}$

$= \frac{800}{50}$

$= 16$

5) Calculate the mean deviation about the median and Coefficient of mean deviation about the median for the following data

Class	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
frequency	18	16	15	12	10	5	2	2

Soln:

Class	frequency (f)	cf	mid x	$\lvert x-M\rvert$	$f\lvert x-M\rvert$
0-10	18	18	5	19	342
10-20	16	34	15	9	144
20-30	15	49✓	25	1	15
30-40	12	61	35	11	132
40-50	10	71	45	21	210
50-60	5	76	55	31	155
60-70	2	78	65	41	82
70-80	2	80	75	51	102
	80				1182

Here $N=80$; $N/2=40$ {40 lies within Cf of 49} ; $c=10$

$\therefore l=20$; $f=15$; $m=34$

$$\text{Median}(M) = l + \frac{\frac{N}{2}-m}{f} \times c$$

$$= 20 + \frac{40-34}{15} \times 10$$

$$= 20 + \frac{60}{15} = 24$$

Mean deviation about median $= \dfrac{\Sigma f|x-M|}{N}$

$$= \frac{1182}{80}$$

$$= 14.775$$

Coefficient of mean deviation about Median $\Big\} = \dfrac{M.D}{Median} = \dfrac{14.775}{2[illegible]}$

$$= 0.0156$$

-x-

Standard Deviation : (ungrouped data)

① for the set of values $x_1, x_2, \ldots, x_n$ such that

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

② If an assumed value A is taken from mean & $d = x - A$

then $\sigma = \sqrt{\dfrac{\Sigma d^2}{n} - \left(\dfrac{\Sigma d}{n}\right)^2}$

③ for frequency distribution,

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

where $d = x - A/c$

C = Class interval

N = Total frequency

① find the SD of the set of nos / 3, 8, 6, 10, 12, 9, 11, (26) 10, 12, 7

Soln: $\Sigma x = 3+8+6+10+12+9+11+10+12+7$

$= 88$

$\Sigma x^2 = 9+64+36+100+144+81+121+100+144+49$

$= 848$

$$\sigma^2 = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2 = \frac{848}{10} - \left(\frac{88}{10}\right)^2 = 7.36$$

$$\sigma = \sqrt{7.36} = 2.71$$

—x—

② find the SD of the following set of observations 45, 36, 40, 37, 39, 42, 45, 35, 40, 39.

Soln:

x	d	d^2
45	5	25
36	-4	16
40	0	0
37	-3	9
39	-1	1
42	2	4
45	5	25
35	-5	25
40	0	0
39	-1	1
	-2	106

$A = 40$

$d = x - A$

$= x - 40$

Now,

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

$$= \sqrt{\frac{106}{10} - \left(\frac{-2}{10}\right)^2}$$

$$= \sqrt{10.6 - .04}$$

$$= \sqrt{10.56}$$

$$= 3.25$$

—x—

③ The weekly salaries of a group of employees given in the following data. find the standard deviation of the salaries.

Salary (in Rs)	75	80	85	90	95	100
No. of Persons	3	7	18	12	6	4

Soln:

x	f	$d = \frac{x-A}{c} = \frac{x-85}{5}$	fd	fd^2
75	3	-2	-6	12
80	7	-1	-7	7
85	18	0	0	0
90	12	1	12	12
95	6	2	12	24
100	4	3	12	36
	50		23	91

Here

$A = 85$

$c = 5$

$\Sigma f = 50$

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

$$= \sqrt{\frac{91}{50} - \left(\frac{23}{50}\right)^2} \times 5$$

$$= 6.34.$$

—x—

Find the standard deviation for the following data:

Wages	70-80	80-90	90-100	100-110	110-120
No. of Persons	12	18	35	42	50
Wages	120-130	130-140	140-150		
No. of persons	45	20	8		

Soln:-

Wages in Rs	Mid x	f	$d = \frac{x-A}{c} = \frac{x-105}{10}$	fd	fd^2
70 - 80	75	12	-3	-36	108
80 - 90	85	18	-2	-36	72
90 - 100	95	35	-1	-35	35
100 - 110	105	42	0	0	0
110 - 120	115	50	1	50	50
120 - 130	125	45	2	90	180
130 - 140	135	20	3	60	180
140 - 150	145	8	4	32	128
		230		125	753

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

$$= \sqrt{\frac{753}{230} - \left(\frac{125}{230}\right)^2} \times 10 = \cancel{17.2583} \quad \cancel{1.724}$$

$$17.248$$

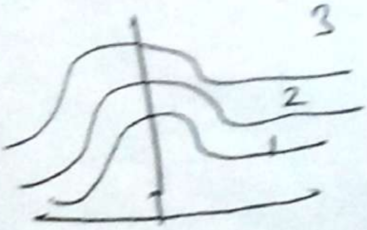
Relative Measure

Relative measure of dispersion based on standard deviation. It is also called Coefficient of variation (CV).

$$CV = \frac{S.D}{mean} \times 100 = \frac{\sigma}{\bar{x}} \times 100$$

Kurtosis :

It is a measure of peakness of a distribution



Normal curve is called as Mesokurtic. i.e) $\beta_2 = 3$

The curve which is more plate opned than the normal curve is called platykurtic i.e) $\beta_2 < 3$

The curve which is more peaked than the normal curve is called leptokurtic i.e) $\beta_2 > 3$.

* Kurtosis is measured by the coefficient.

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

* Measure of skewness based on moment is

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

The 1st four moments of a distribution are ~~about~~ 0, 2.5, 0.7 and 8.75. find skewness and kurtosis.

Soln: $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0.031$

$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3$

—x—

A factory produces 2 types of electric bulbs A and B. In an experiment relating to their life, the following results were obtained.

Length of life (in hours)	500-700	700-900	900-1100	1100-1300	1300-1500
No. of bulbs A	5	11	26	10	8
No. of bulbs B	4	30	12	8	6

Compare the variability of the two varieties using the coefficient of variation.

Soln: WKT, Coeff of Variation $= \frac{\sigma}{\bar{x}} \times 100$

for ~~A~~ A

① To find $\bar{x} = A + \left[\frac{\Sigma fd}{N}\right] \times C$

② To find $\sigma = \left(\sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}\right) \times C$

For A - Bulbs

Class	Mid Value of class x	f	$d = \frac{x-A}{c}$ $= \frac{x-1000}{200}$	fd	fd^2
500-700	600	5	-2	-10	20
700-900	800	11	-1	-11	11
900-1100	1000 ✓	26	0	0	0
1100-1300	1200	10	1	10	10
1300-1500	1400	8	2	16	32
		60		5	73

$\therefore \bar{x} = 1000 + \left[\frac{5}{60}\right] \times 200$

$\boxed{\bar{x} = 1016.67}$

$\sigma = \left[\sqrt{\frac{73}{60} - \left(\frac{5}{60}\right)^2}\right] \times 200$

$\boxed{\sigma = 219.98}$

Coeff of variation for A $= \frac{219.98}{1016.67} \times 100$

$= 21.64$

For B-bulbs

Class	Mid x	f	$d = \frac{x-A}{c}$	fd	fd^2
	600	4	-2	-8	16
	800	30	-1	-30	30
	1000✓	12	0	0	0
	1200	8	1	8	8
	1400	6	2	12	24
				-18	78

$$\bar{x} = 100 + \left[\frac{-18}{60}\right] \times 200$$

$$\Rightarrow \boxed{\bar{x} = 940}$$

$$\sigma = \sqrt{\frac{78}{60} - \left(\frac{-18}{60}\right)^2} \times 200$$

$$\Rightarrow \boxed{\sigma = 220}$$

coefficient of variation for B $= \frac{220}{940} \times 100$

$= 23.40$.

$\because$ C.V for B is more than that of A.

$\therefore$ Type B bulbs is more Variable.

— x —