STATISTICS: UNIT - II

Mode = 3 Median - 2 Mean

Mean: Some what of Central Data. [Average Value]

Median: Devides the entire Data into two equal Parts.

Mode: Highest Frequency of Data.

Ungrouped Data	Discrete Data	Continous Data
Mean: $\bar{x} = \frac{\Sigma x}{n}$	Mean: $\bar{x} = \frac{\Sigma fx}{\Sigma f}$ (or) $\frac{\Sigma fx}{N}$ $N = \Sigma f$	Mean: $\bar{x} = \frac{\Sigma fm}{\Sigma f}$ m → mid point of class, f → frequency
Median: odd [Median]: $\left(\frac{N+1}{2}\right)^{th}$ value even ["]: Avg $\left(\frac{n}{2}, \frac{n}{2}+1\right)^{th}$ val	Median: $\left(\frac{N+1}{2}\right)^{th}$ value	Median: $l + \frac{h}{f}\left(\frac{N}{2} - c\right)$ l - lower limit h - length of class f - frequency of class c - frequency preceding class.
Mode: No. of highest occurance.	Mode: Highest frequency	Mode: $l + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$

1. Calculate Mean from the following Data.

R. No	1	2	3	4	5	6	7	8	9	10
Marks	40	50	55	78	58	60	73	35	43	48

Sol: $\bar{x} = \frac{\Sigma x}{n} = \frac{540}{10} = 54.$

2. An average Rain fall of a city from Monday - Saturday is 0.3 inch due to heavy rain fall on Sunday. The avg Rain on the week incread to 0.5 inch. What was the Rain fall on Sunday?

Monday - Saturdy (6days) The average is 0.3 $\bar{x} = 0.3$

Sunday - Saturday (7days) The average is 0.5 $\bar{y} = 0.5$

$\frac{\Sigma x}{n} = 0.3 \Rightarrow \frac{\Sigma x}{6} = 0.3 \Rightarrow \boxed{\Sigma x = 1.8}$

$\frac{\Sigma y}{n} = 0.5 \Rightarrow \frac{\Sigma y}{7} = 0.5 \Rightarrow \boxed{\Sigma y = 3.5}$

Rain fall on Sunday $= \Sigma x - \Sigma y$
$= 3.5 - 1.8$
$= 1.75$

3. Mean of twenty values is 45. If one of these values is to be taken 64 instead of 46 Find the Corrected Mean.

Sol: $\bar{x} = 45 \Rightarrow \frac{\Sigma x}{n} = 45 \Rightarrow \frac{\Sigma x}{20} = 45 \Rightarrow \Sigma x = 900.$

$\Sigma x - 64 + 46 \Rightarrow 900 - 64 + 46 = 882$

Corrected Mean $= C.\frac{\Sigma x}{n} = \frac{882}{20} \Rightarrow 44.1$

4. The Mean Salary Paid to Thousand works of an establishment was found to be R.s 180.40 Later on, After dispasment of salary it was discovered that the salary of two employee was wrongly entered as R.s 297 And R.s 165. Their correct salary were R.s. 197 & R.s 185 Find the correct A. Mean?

$\bar{x} = 180.40 \Rightarrow \frac{\Sigma x}{n} = 180.40 \Rightarrow \frac{\Sigma x}{1000} = 180.40 \Rightarrow$ ~~Σx=180~~

$\Sigma x = 180400.$

Sol: $\Sigma x - 297 - 165 + 197 + 185 = 180320$

Arithmetic Mean $= C\frac{\Sigma x}{n} = \frac{180320}{1000} = 180.32$

5. The Mean Age of Group of 100 children was 9.35 years. The Mean Age of 25 of them is 8.75 And that of 65 was 10.51 years What was the Mean Age for remainder?

$\bar{x} = 9.35 \Rightarrow \frac{\Sigma x}{n} = \frac{\Sigma x}{100} = 9.35 = 93.5$

The Mean for remainder $= -9.35 + [8.75 + 10.51]$
$= 9.91$

6. In class of 50 students 10 have failed. The avg mark 2.5 the total marks secured by the entire class were 281. Find the avg marks of those who have passed.

$N = 50$ Pass Total $= 281 - 25$

$\bar{x} = 2.5$ $= 256$

$\frac{\Sigma x}{n} = 2.5$ Pass Average $= \frac{256}{40} = 6.4$

$\frac{\Sigma x}{10} = 2.5$

$\Sigma x = 25$

Discrede Data.

1. Find Arithmetic Mean for following Data

x :	5	10	15	20	25	30	35	40	45	50
f :	20	43	75	67	72	45	39	9	8	6

Sol:

x	f	fx
5	20	100
10	43	430
15	75	1125
20	67	1340
25	72	1800
30	45	1350
35	39	1365
40	9	360
45	8	360
50	6	300
	384	8530

$\bar{x} = \frac{\Sigma fx}{\Sigma f}$

$= \frac{8530}{384} = 22.21$

2. Find Mean.

x :	1	2	3	4	5	15
f :	3	5	9	6	2	25
fx :	3	10	27	24	10	74

$\bar{x} = \frac{\Sigma fx}{\Sigma f}$

$= \frac{74}{25} = 2.96$

3. Compute the Mean.

(cm) x :	219	216	213	210	207	204	201	198	195	
No. of person :	2	4	6	10	11	7	5	4	1	50
Σfx :	438	864	1278	2100	2277	1428	1005	792	195	10377

$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{10377}{50} = 207.54$

CONTINOUS DATA.

①

x	f	m	fm
0-10	5	5	25
10-20	10	15	150
20-30	40	25	1000
30-40	20	35	700
40-50	25	45	1125
	100		3000

$\bar{x} = \frac{\Sigma fm}{\Sigma f} = \frac{3000}{100}$

$= 30$

②

x	f	m	fm
0-10	12	5	60
10-20	18	15	270
20-30	27	25	675
30-40	20	35	700
40-50	17	45	765
50-60	6	55	330
	100		2800

$\bar{x} = \frac{\Sigma fm}{\Sigma f} = \frac{2800}{100}$

$= 28.$

③

Find the avg marks of students from following data.

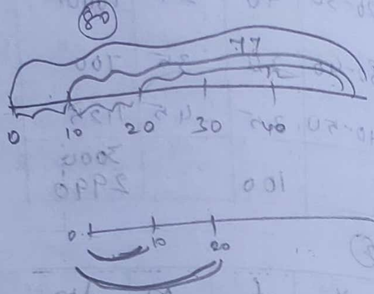
Marks	Above 0	Above 10	Above 20	Above 30	Above 40	Above 50	Above 60	Above 70	Above 80	Above 90	Above 100
No. of stu	80	77	72	65	55	43	28	16	10	8	0

Marks	No. of students	x	f	m	fm
Above 0	80	0-10	3	5	15
Above 10	77	10-20	5	15	75
Above 20	72	20-30	7	25	175
Above 30	65	30-40	10	35	350
Above 40	55	40-50	12	45	540
Above 50	43	50-60	15	55	825
Above 60	28	60-70	12	65	780
Above 70	16	70-80	6	75	450
Above 80	10	80-90	2	85	170
Above 90	8	90-100	8	95	760
Above 100	0				
			80		4140

$$\bar{x} = \frac{\Sigma fm}{\Sigma f}$$

$$= \frac{4140}{80}$$

$$= 51.75$$



④ Find the mean from the following Data.

Marks	No. of students	x	f	m	fm
Below 10	5	0-10	5	5	25
" 20	9	10-20	4	15	60
" 30	17	20-30	8	25	200
" 40	29	30-40	12	35	420
" 50	45	40-50	16	45	720
" 60	60	50-60	15	55	825
" 70	70	60-70	10	65	650
" 80	78	70-80	8	75	600
" 90	83	80-90	5	85	425
" 100	85	90-100	2	95	190
			85		4115

$$\bar{x} = \frac{\Sigma fm}{\Sigma f}$$

$$= \frac{4115}{85}$$

$$= 48.41$$

⑤ Find the Mean for following Data.

Marks	No. of Candidates	x	f	m	fm
1-5	7	0.5 - 5.5	7	3	21
6-10	10	5.5 - ~~9~~ 10.5	10	8	80
11-15	16	10.5 - 15.5	16	13	208
16-20	32	15.5 - 20.5	32	18	576
21-25	24	20.5 - 25.5	24	23	552
26-30	18	25.5 - 30.5	18	28	504
31-35	10	30.5 - 35.5	10	33	330
35-40	5	35.5 - 40.5	5	38	190
41-45	1	40.5 - 45.5	1	43	43
			123		2504

$$\bar{x} = \frac{\Sigma fm}{\Sigma f}$$

$$= \frac{2504}{123}$$

$$= 20.35$$

⑥ Find the Mean for following Data.

Mid Values :	115	125	135	145	155	165	175	185	195	
f :	6	25	48	72	116	60	38	22	3	390
fm	690	3125	6480	10440	17980	9900	6650	4070	585	59920

$$\bar{x} = \frac{\Sigma fm}{\Sigma f} = \frac{59920}{390} = 153.64$$

MEDIAN : Median is a Central Measure that divides the entire data into two equal Parts.

RAW DATA :

1. Obtain the Median for following Data.

7, 8, 10, 9, 18, 15, 12, 11, 14, 17

Sol: Ascending order. 7, 8, 9, 10, 11, 12, 14, 15, 17, 18

$N = 10$ (even)

$$\text{Median} = \text{Avg}\left(\frac{n}{2}, \frac{n}{2}+1\right)^{th} \text{ value} \Rightarrow \frac{11+12}{2}$$

$$= \text{Avg}\left(\frac{10}{2}, \frac{10}{2}+1\right)^{th} \text{ value} = \frac{23}{2}$$

$$= \text{Avg}(5, 6)^{th} \text{ value} \quad \boxed{= 11.5}$$

$$= \text{Avg } 11, 12$$

2. Obtain the Median : 5, 7, 9, 11, 13, 15, 17

Ascending order : 5, 7, 9, 11, 13, 15, 17

$n = 7$ (odd)

Median $= \left(\frac{n+1}{2}\right)^{th}$ value $\Rightarrow \left(\frac{7+1}{2}\right)^{th}$ value $= 4^{th}$ value $= \boxed{11}$

3. Obtain the Median : 30, 12, 25, 38, 2, 10, 4, 8, 15, 20

Ascending order : 2, 4, 8, 10, 12, 15, 20, 25, 30, 38

$= n = 10$ (even)

Median $=$ Avg $\left(\frac{n}{2}, \frac{n}{2}+1\right)^{th}$ value $=$ Avg$(5, 6)^{th}$ value $= \frac{27}{2}$

$=$ Avg $\left(\frac{10}{2}, \frac{10}{2}+1\right)^{th}$ value $= \frac{12+15}{2}$ $\boxed{= 13.5}$

4. Find Median : 2, 30, 12, 25, 20, 8, 10, 4, 15

Ascending Order : 2, 4, 8, 10, 12, 15, 20, 25, 30

$n = 9$ (odd)

Median $= \left(\frac{n+1}{2}\right)^{th}$ value $\Rightarrow \left(\frac{9+1}{2}\right)^{th}$ value $= \cancel{10}\, 5^{th}$ value $= \boxed{12}$

Discrete Data :

1. Calculate Median for following Data.

x :	1	2	3	4	5	6	7	8	9
f :	8	10	11	16	20	25	15	9	6
c.f	8	18	29	45	$\boxed{65}$	90	105	114	120

$N = 120$

Median $= \left(\frac{N+1}{2}\right)^{th}$ value $= \left(\frac{120+1}{2}\right)^{th}$ value $= (60.5)^{th}$ value

The Next highest value is 65

So the Median is $\boxed{5}$

2. Find the Median for following Data.

x :	5	7	9	11	13	15	17	19
f :	1	2	7	9	11	8	5	4
c.f :	1	3	10	19	$\boxed{30}$	38	43	47

$N = 47$

Median $= \left(\frac{N+1}{2}\right)^{th}$ value $= \left(\frac{47+1}{2}\right)^{th}$ value $= \left(\frac{48}{2}\right)^{th}$ value $= 24^{th}$ value

The Next highest value is 30

So the median is $\boxed{13}$

Continous Data.

1. Find the Median for following Data.

Marks	No. of Students	c.f
0 - 10	2	2
10 - 20	18	20
20 - 30	30	50
30 - 40	45	95
40 - 50	35	130
50 - 60	20	150
60 - 70	6	156
70 - 80	3	159

Median : $\left(\frac{N+1}{2}\right)^{th}$ value $= \left(\frac{159+1}{2}\right)^{th}$ value $= 80^{th}$ value

The Next highest value is 95.

So the Class interval is 30 - 40.

Formual : $l + \frac{h}{f}\left(\frac{N}{2} - c\right)$ $\quad l = 30,\ h = 10,\ N = 159,\ f = 45,\ c = 50$

$= 30 + \frac{10}{45}\left(\frac{159}{2} - 50\right)$

$= 30 + \cancel{\frac{10}{45}}\, 0.22\, (79.5 - 50)$

$= \cancel{30.54}(29.5)$

$= 30 + 6.49$

$= 36.49$

2. Find the Median:

Wages on Rs	No. of Workers	Cf
0-10	22	22
10-20	38	60
20-30	46	$\boxed{106}$
30-40	35	141
40-50	20	161

Median = $\left(\frac{N+1}{2}\right)^{th}$ value = $\left(\frac{161+1}{2}\right)^{th}$ value

= 81^{th} value ∴ The Next highest value is 106. The class interval 20-30

Median: $l + \frac{h}{f}\left(\frac{N}{2} - cf\right)$

$= 20 + \frac{10}{46}\left(\frac{161}{2} - 60\right)$

$= 20 + 4.44 = \boxed{24.44}$

3. Find the Median

Mid value	frequency	class interval	cf
115	6	110-120	6
125	25	120-130	31
135	48	130-140	79
145	72	140-150	151
155	$\boxed{116}$	150-160	267
165	60	160-170	327
175	38	170-180	365
185	22	180-190	387
195	3	190-200	390

Median = $\left(\frac{N+1}{2}\right)^{th}$ value = $\left(\frac{391}{2}\right)^{th}$ value

= 195.5^{th} value: The Next highest value is 116 so the class 150-160

Median: $l + \frac{h}{f}\left(\frac{N}{2} - cf\right)$

$= 150 + \frac{10}{116}\left(\frac{390}{2} - \cancel{72}\,151\right)$

$= 150 + 0.86\,(44)$

$= 153.793$

4. Find the Median

Marks	No. of stu	Class inte	f	cf
Below 10	15	0-10	15	15
" 20	35	10-20	20	35
" 30	60	20-30	25	60
" 40	84	30-40	24	84
" 50	94	40-50	10	94
" 60	124	50-60	33	$\boxed{127}$
" 70	198	60-70	71	198
" 80	249	70-80	51	249

Median = $\left(\frac{N+1}{2}\right)^{th}$ value = $\left(\frac{249+1}{2}\right)^{th}$ value

= 125^{th} value. The Next highest Value is 124 ∴ the class 50-60

Median: $l + \frac{h}{f}\left(\frac{N}{2} - cf\right)$

$= 50 + \frac{10}{33}\left(\frac{249}{2} - 94\right)$

$= 50 + \frac{10}{33}\left(30.5\right)$

$= 50 + 9.242$

$= 59.242$

5. Find the Median

Marks	frequency	Class Inter	cf
10-19	7	9.5-19.5	7
20-29	15	19.5-29.5	22
30-39	18	29.5-39.5	40
40-49	25	39.5-49.5	65
50-59	30	49.5-59.5	95
60-69	20	59.5-69.5	115
70-79	16	69.5-79.5	131
80-89	7	79.5-89.5	138
90-99	2	89.5-99.5	140

Median = $\left(\frac{N+1}{2}\right)^{th}$ value = $\left(\frac{141}{2}\right)^{th}$ value

= 70.5 The Next highest value is 90 ∴ The class interval 49.5-59.5

Median: $l + \frac{h}{f}\left(\frac{N}{2} - cf\right)$

$= 49.5 + \frac{10}{30}\left(\frac{140}{2} - 65\right)$

$= 49.5 + \frac{10}{30}\,(5)$

$= 49.5 + 1.6$

$= 51.16$

6. Find the Median

Marks	f	C. Inter	f	cf
Above 0	50	0-10	4	4
" 10	46	10-20	6	10
" 20	40	20-30	20	30
" 30	20	30-40	10	40
" 40	10	40-50	7	47
" 50	3	50-60	3	50

Median = $\left(\frac{N+1}{2}\right)^{th}$ value = $\left(\frac{51}{2}\right)^{th}$ value

= 25.5 The Next highest value is 40 ∴ the class interval 20-30

Median: $l + \frac{h}{f}\left(\frac{N}{2} - cf\right)$

$= 20 + \frac{10}{20}\left(25 - \cancel{46}\,10\right)$

$= 20 + 7.5$

$= 27.5$

MODE: Mode is the value it occurs most frequently in set of observations. It is a point of maximum frequency or the point of greater density.

RAW DATA.

1. If 7 Men are receiving Daily wages of Rs 5, 6, 7, 7, 7, 8, 10. Find Model Wage.

The Mode is 7

2. Determine Mode from the following.

25, 15, 23, 40, 27, 25, 23, 25, 20

The Mode is 25

Discree Data

1. Find Mode

X	1	2	3	$\boxed{4}$	5	6	7	8
f	4	9	16	25	22	16	8	3

The highest frequency is 25 so the Mode is 4.

~~1~~ 1. Find Mode [CONTINUOUS DATA]

Class Interval	frequency
0 - 7	~~4~~ 19
7 - 14	9 25
14 - 21	~~16~~ 36
21 - 28	~~25~~ 72
28 - 35	~~22~~ 51
35 - 42	~~16~~ 43
42 - 49	8 28

Model class : 21-28

$l = 21,\ h = 7,\ \Delta_1 = f_m - f_1 = 72 - 36 = 36,\ \Delta_2 = f_m - f_2 = 72 - 51 = 21$

$\text{Mode} = l + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$

$= 21 + \left(\frac{36}{36+21}\right) \times 7 = 21 + 4.42 = 25.42$

2. Find Mode.

Marks	No. of std
0-10	2
10-20	18
20-30	30
30-40	45
40-50	35
50-60	20
60-70	6
70-80	3

Model class : 30-40

$l = 30$

$h = 10$

$\Delta_1 = 45 - 30 = 15$

$\Delta_2 = 45 - 35 = 10$

$\text{Mode} = l + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$

$= 30 + \left(\frac{15}{15+10}\right) \times 10 = 30 \left(\frac{25}{25}\right) \times 10$

$= 36$

3. Find the Mode from following frequency Distribution

X	f(I)	II	III	IV	V	VI
4	2					
5	5	7 (4,5)		15 (4,5,6)		
6	8		13 (5,6)		22 (5,6,7)	
7	9	17 (6,7)				29 (6,7,8)
8	12		21 (7,8)	35 (7,8,9)		
9	14	26 (8,9)			$\boxed{40}$ (8,9,10)	
10	14		$\boxed{28}$ (9,10)			$\boxed{43}$ (9,10,11)
11	$\boxed{15}$	$\boxed{29}$ (10,11)		$\boxed{40}$ (10,11,12)		
12	11		26 (11,12)		39 (11,12,13)	
13	13	24 (12,13)				

I - frequency

II - Add two

III - Leave first & Add two

IV - Add Three

V - Leave first & Add three

VI - Leave two & Add three

Coloms	x value
C - I	11
C - II	10, 11
C - III	9, 10
C - IV	10, 11, 12
C - V	8, 9, 10
C - VI	9, 10, 11

10 has occured More times So $\boxed{\text{Mode is 10}}$

4. Find the Mode

Size	freq	II	III	IV	V	VI
5	48					
6	52	100 (5,6)		156 (5,6,7)		
7	56		108 (6,7)		$\boxed{168}$ (6,7,8)	
8	60	116 (7,8)				$\boxed{163}$ (7,8,9)
9	53		113 (8,9)	$\boxed{170}$ (8,9,10)		
10	57	110 (9,10)			165 (9,10,11)	
11	55		112 (10,11)			162 (10,11,12)
12	50	105 (11,12)		157 (11,12,13)		
13	52		102 (12,13)		143 (12,13,14)	
14	41	93 (13,14)				150 (13,14,15)
15	57		98 (14,15)	161 (14,15,16)		
16	$\boxed{63}$	$\boxed{120}$ (15,16)			152 (15,16,17)	
17	52		$\boxed{115}$ (16,17)			163 (16,17,18)
18	48	100 (17,18)		140 (17,18,19)		
19	40		88 (18,19)			

columns	x value.
C - I	16
C - II	15, 16
C - III	16, 17
C - IV	8, 9, 10
C - V	6, 7, 8
C - VI	7, 8, 9 / 16, 17, 18

16 has occured More times so the Mode is 16

1. Find Mean, Median, Mode for following Data.

Size	frequency	fx	cf
1	3	3	3
2	8	16	11
3	15	45	26
4	23	92	49
5	35	175	84
6	40	240	124
7	32	224	156
8	28	224	184
9	20	180	204
10	45	450	249
11	14	154	263
12	6	72	269
	269	1875	

Mean: $\frac{\Sigma fx}{N}$

$= \frac{1875}{269} = 6.97$

Median: $\left(\frac{N+1}{2}\right)^{th}$ value $= \left(\frac{269+1}{2}\right)^{th}$ value

$=$ 135th value ∴ The Next highest value 156 so $= 7$

Mode: 6 has occured More times so The Mode is 6

Mean	:	6.97
Median	:	4
Mode	:	6

X	I	II	III	IV	V	VI
1	3					
2	8	11 (1,2)		26 (1,2,3)		
3	15		23 (2,3)		46 (2,3,4)	
4	23	38 (3,4)		$\boxed{98}$ (4,5,6)		73 (3,4,5)
5	35		58 (4,5)			
6	40	$\boxed{75}$ (5,6)			$\boxed{107}$ (5,6,7)	
7	32		$\boxed{72}$ (6,7)			$\boxed{100}$ (6,7,8)
8	28	60 (7,8)		70 (7,8,9)		
9	20		48 (8,9)		93 (8,9,10)	
10	$\boxed{45}$	65 (9,10)				79 (9,10,11)
11	14		59 (8,11)	65 (10,11,12)		
12	6	20 (11,12)				

Columns	x value
I	10
II	5,6
III	6,7
IV	4,5,6
V	5,6,7
VI	6,7,8

1. calculate the Mean, Median, Mode for the following.

Age	No. of People	m	fm	cf	I	II	III	IV	V	VI
20-25 A	14	22.5	315	14	14					
25-30 B	28	27.5	770	42	28	42 (a,b)		$\boxed{75}$ (a,b,c)		
30-35 C	33	32.5	1072.5	75	$\boxed{33}$		$\boxed{61}$ (b,c)		$\boxed{91}$ (b,c,d)	
35-40 D	30	37.5	1125	105	30	$\boxed{63}$ (c,d)				$\boxed{88}$ (c,d,e)
40-45 E	20	42.5	850	125	20		50 (d,e)	65 (d,e,f)		
45-50 F	15	47.5	712.5	140	15	35 (e,f)			48 (e,f,g)	
50-55 G	13	52.5	682.5	153	13		28 (f,g)			35 (f,g,h)
55-60 H	4	57.5	402.5	160	4	20 (g,h)				
	160		5929.5							

Mean: $\frac{\Sigma fm}{N} = \frac{5929.5}{160} = \boxed{37.05}$

Median: $\left(\frac{N+1}{2}\right)^{th}$ value $= \left(\frac{160+1}{2}\right)^{th}$ value $= 80.5^{th}$ value

So the median class interval is 35-40

$l + \frac{h}{f}\left(\frac{N}{2} - c\right) \Rightarrow 35 + \frac{5}{30}\left(\frac{160}{2} - 75\right) = \boxed{35.83}$

Mode: $l + \frac{\Delta_1}{\Delta_1+\Delta_2} \times h = 30 + \frac{5}{5+3} \times 5 = 30 + \frac{5}{8} \times 5 = \boxed{33.125}$

So the modal class interval is 30-35

Columns	X Value
I	c
II	c,d
III	b,c
IV	a,b,c
V	b,c,d
VI	c,d,e

Mean:	37.05
Median:	35.83
Mode:	~~30-35~~ 33.125

Measure of Dispersion:

(I) Range (II) Quartile Deviation (or) Semi inter Quatile Range

(III) Standard Deviation

Range:

Range is the Difference Between the extreme Values of the Variable.

$L-S$ where $L \to$ largest value, $S \to$ Smallest Value

1. Find the Range of weights of 7 Students from the following Data. 27, 30, 35, 36, 38, 40, 43

Range: $L-S = 43-27 = 16$

Coefficient of Range: $\frac{L-S}{L+S} = \frac{43-27}{43+27} = \frac{16}{70} = 0.23$

Merits:

* It is simple to compute And Understand.

* It gives Rough but Quick Answer.

Demerits:

* It is not reliable because it is affected by the extreme Values.

~~Demerits~~:

* Usually frequency Distribution may be Concentrated in the middle of the Series, But range depends on extreme Values, & It is an unsatisfactoring measure.

STANDARD DEVIATION:

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

1. Calculate the Standard deviation.

14, 22, 9, 15, 20, 17, 12, 11

x	14	22	9	15	20	17	12	11	
$x-\bar{x}$ ($x-15$)	-1	7	-6	0	5	2	-3	-4	
$(x-\bar{x})^2$	1	49	36	0	25	4	9	16	140

$\sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$ $\bar{x} = \frac{120}{8} = 15$

$\sigma = \sqrt{\frac{140}{8}} = \sqrt{17.5} = 4.18$

2. Calculate Standard Deviation.

x	12	21	21	23	27	28	30	34	37	39	39	39	40	49	54	
$x-\bar{x}$ ($x-33$)	-21	-12	-12	-10	-6	-5	-3	1	4	6	6	6	7	16	21	
$(x-\bar{x})^2$	441	144	144	100	36	25	9	1	16	36	36	36	49	256	441	1770

$\bar{x} = \frac{493}{15} = 32.8$ $\sigma = \sqrt{\frac{1770}{15}} = \sqrt{118} = 10.86$

Discrete Data: $\boxed{\sigma = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{N}}}$

1. Calculate standard deviation.

Marks	No. of stud	fx	$x-\bar{x}$ ($x-31$)	$(x-\bar{x})^2$	$f(x-\bar{x})^2$
10	8	80	-21	441	3528
20	12	240	-11	121	1452
30	20	600	-1	1	20
40	10	400	9	81	810
50	7	350	19	361	2527
60	3	180	29	841	2523
	60	1850			10860

$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{1850}{60} = 30.83 \simeq 31$

$\sigma = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{N}} = \sqrt{\frac{10860}{60}} = \sqrt{181} = 13.45$

2. Calculate Standard Deviation.

Size	frequency	fx	$(x-\bar{x})$ $x-9$	$(x-\bar{x})^2$	$f(x-\bar{x})^2$
6	3	18	-3	9	~~162~~ 27
7	6	42	-2	4	~~168~~ 24
8	9	72	-1	1	~~72~~ 9
9	13	117	0	0	~~0~~ 0
10	8	80	1	1	~~80~~ 8
11	5	55	2	4	~~220~~ 20
12	4	36	3	9	~~324~~ 36
	48	420			~~1026~~ 124

$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{420}{48} = 8.75$ $\sigma = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{N}} = \sqrt{\frac{\cancel{1026}}{\cancel{48}}} = \sqrt{\cancel{21.37}} = \cancel{4.62}$

$= \sqrt{\frac{124}{48}} = \sqrt{2.58} = 1.60$

3. Calculate standard deviation.

x	f	fx	$(x-\bar{x})$ $x-27$	$(x-\bar{x})^2$	$f(x-\bar{x})^2$
15	3	45	-12	144	432
20	25	500	-7	49	1225
25	19	475	-2	4	76
30	16	480	3	9	144
35	4	140	8	64	256
40	5	200	13	169	845
45	6	270	18	324	1944
	78	2110			4922

$\bar{x} = \frac{2110}{78} = 27$ $\sigma = \sqrt{\frac{4922}{78}} = \sqrt{63.10} = 7.94$

CONTINOUS DATA: $C.V = \frac{\sigma}{\bar{x}} \times 100$, $\sigma = \sqrt{\frac{\Sigma f(x_m - \bar{x})^2}{N}}$

1. Find S.D & coefficient of Variance.

Marks	No. of stu	xm	fm	$x-\bar{x}$ $x-39$	$(x-\bar{x})^2$	$f(x-\bar{x})^2$
0-10	5	5	25	-34	1156	5780
10-20	10	15	150	-24	576	5760
20-30	20	25	500	-14	196	3920
30-40	40	35	1400	-4	16	640
40-50	30	45	1350	6	36	1080
50-60	20	55	1100	16	256	5120
60-70	10	65	650	26	676	6760
70-80	4	75	300	36	1296	5184
	139		5475			34244

$\bar{x} = \frac{\Sigma fm}{N} = \frac{5475}{139} = 39.39$, $\sigma = \sqrt{\frac{34244}{139}} = \sqrt{246.35} = 15.69$

$C.V = \frac{15.69}{39.39} \times 100 = 39.84$

2. Find S.D & C.V

Wage	No. of workers	m	fm	$x-\bar{x}$ $x-5$	$(x-\bar{x})^2$	$f(x-\bar{x})^2$
1-3	15	2	30	-3	9	135
3-5	18	4	72	-1	1	18
5-7	27	6	162	1	1	27
7-9	10	8	80	3	9	90
9-11	6	10	60	5	25	~~30~~ 150
	76		404			420

$\bar{x} = \frac{404}{76} = 5.3$ $\sigma = \sqrt{\frac{420}{76}} = \sqrt{5.52} = 2.3$

$C.V = \frac{2.3}{5.3} \times 100 = 43.3$

1. The scores of 2 Golfers for 10 rounds each hour

A:	58	59	60	54	65	66	52	75	69	52
B:	84	56	92	65	86	78	44	54	78	68

which may be regarded as the more consisten [c.v]

Player.

Player A

$\bar{x} = 61$, $\sigma_1 = 7.253$

C.V = 11.8

player B

$\bar{x} = 70.5$, $\sigma = 14.89$

C.V = 21.12

Player A is more consistent.

2. In two sample where the variates x_1, x_2 are measured in same units from their respective means $n_1 = 36$, $n_2 = 49$, $\Sigma x_1^2 = 49428$, $\Sigma x_2^2 = 71258$ compute S.D of two samples. what additional information is require to calculate C.V of above TWO samples.

$n_1 = 36$, $\Sigma x_1^2 = 4942$, $n_2 = 49$, $\Sigma x_2^2 = 71258$

$\sigma_1 = \sqrt{\frac{49428}{36}} \Rightarrow 1.171$, $\sigma_2 = \sqrt{\frac{71258}{49}} = 38.13$

3. Calculate Quartile Deviation

Marks	No. of Students	cf
10	4	4
20	7	11
30	15	26
40	8	34
50	7	41
60	2	43

$N = 43$

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

$= \left(\frac{44}{4}\right)^{th}$ value

$Q_1 = 20$

$Q_3 = 3\left(\frac{N+1}{4}\right)^{th}$ value

$= 3\ (11^{th})$ value

$= 33^{th}$ value

$Q_3 = 40$

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

Coefficient $= \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{40-20}{40+20} = 0.333$

4. Calculate Semi inter Quartile Range.

Class	frequency	cf
0-10	4	4
10-20	15	19
20-30	28	47
30-40	16	63
40-50	7	70
	70	

$Q_1 = l + \frac{h}{f}\left(\frac{N}{4} - cf\right)$

$Q_3 = l + \frac{h}{f}\left(3\frac{N}{4} - cf\right)$

Q_1 class $= \left(\frac{N+1}{4}\right)^{th}$ value $= \left(\frac{71}{4}\right)^{th} = 17.75^{th}$ value $= 10-20$.

$Q_1 = l + \frac{h}{f}\left(\frac{N}{4} - cf\right) = 10 + \frac{10}{15}\left(\frac{70}{4} - 4\right) = 19.$

Q_3 class $= 3\left(\frac{N+1}{4}\right)^{th}$ value $= 3\ (17.75)^{th}$ value $= 53^{th}.25$ value $= 30-40$

$Q_3 = l + \frac{h}{f}\left(\frac{3N}{4} - cf\right) = 10 + \frac{30}{16}\left(\frac{3(70)}{4} - 47\right) = 33.44$

$Q.D = \frac{Q_3 - Q_1}{2} = \frac{33.44 - 19}{2} = \frac{14.44}{2} = 7.22$

5. Find Semi inter Quartile Range & coefficient.

Marks	No. of stu	cf
10-20	60	60
20-30	45	105
30-40	120	225
40-50	25	250
50-60	90	340
60-70	80	420
70-80	120	540
80-90	60	600

Q_1 class $= \left(\frac{N+1}{4}\right)^{th}$ value $= \left(\frac{601}{4}\right)^{th}$ value $= 30-40$

$Q_1 = l + \frac{h}{f}\left(\frac{N}{4} - cf\right) = 30 + \frac{10}{120}\left(\frac{600}{4} - 105\right)$

$= 30 + 0.833(150 - 105)$

$= \cancel{67.485}\ 33.75$

$Q_3 = l + \frac{h}{f}\left(3\frac{N}{4} - cf\right) = 70 + \frac{10}{120}\left(3\frac{600}{4} - 420\right)$

$= 72.5$

Q_3 Class $= 3\left(\frac{N+1}{4}\right) = 450.75^{th} = 70-80$.

$Q.D = \frac{Q_3 - Q_1}{2} = 19.375$, coefficient $= \frac{Q_3 - Q_1}{Q_3 + Q_1} = 0.364$

TESTING HYPOTHESIS

Null: Hypothesis of No difference. [H_0]

Alternative: Hypothesis of difference. [H_1]

Type I - Reject Ho When it is True. [Produces Risk]

Type II - Accept H_1 When it is False: [Consomer Risk]

Critical Region: Rejion of Rejection

Formula: $z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}}$

1.

$\frac{x}{n}$ p - Sample proportion.

P - Population proportion

$Q = 1 - P$

n - Sample Size

1. Experience has shown that 20% of manufactured Product is a top Quality. In 1 day's production of 400 articals only 50 are top Quality. Test Weather the production of dat chooren was not representative Sample.

$P = 20\% = \frac{20}{100} = 0.2$

$p = \frac{50}{400} = 0.125$

$n = 400$, $Q = 1 - P = 1 - 0.2 = 0.8$

Testing Hypothesis:

Null Hypothesis $H_0 = P = P$

Alternative Hypothesis: $H_1 = p \neq P$

Level of significance: $5\% = 1.96$

Test statistics $z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}} = \dfrac{0.125 - 0.2}{\sqrt{\dfrac{(0.2)(0.8)}{400}}} = -3.75$

∴ Conclusion:

$C.V = |z| = 3.75$

$C.V > T.V$

So Reject H_0

2. A Fertality Typoid Patient is belived to be 17.26%. in a certain year 640 patients suffering from typoid were treatd in Hospital and only 63 patient died can you consider the hospital efficient.

1. Null hypothesis: $P = P$
2. Alternative ": $p \neq P$
3. α: 0.5 : T.V = 1.96
4. Test statics: $z = -4.96$
5. Conclusion: C.V > T.V

∴ Reject H_0

α - 0.5	~~One~~ Two Tail	~~Two~~ One Tail
5%	1.96	1.645
2%	2.33	2.055
1%	2.58	2.33

ANOVA → Analysis of Variance.

Steps:

1. find T, N $T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3$
2. find $\frac{T^2}{N}$
3. find $TSS = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 \dots - \frac{T^2}{N}$ [Total Sum of Squares]
4. find $SSC = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \dots - \frac{T^2}{N}$ [Sum of squares of column]
5. find $SSE = TSS - SSC$ [Sum of Square of errors]

Null hypothesis: There is No significance d/b columns

Alternative ": There is a significance d/b columns.

x_1	x_2	x_3		x_1^2	x_2^2	x_3^2
5	8	7		25	64	49
6	10	3		36	100	9
8	11	5		64	121	25
9	12	4		81	144	16
7	4	1		49	16	1
35	45	20	100 T	255	445	100

STEP:

1. $T = 100$, $N = 15$

2. $\frac{T^2}{N} = \frac{100^2}{15} = 666.67$

3. $TSS = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 - \frac{T^2}{N} = 255 + 445 + 100 - 666.67 = 133.33$

4. $SSC = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - \frac{T^2}{N} = \frac{1225}{5} + \frac{2025}{5} + \frac{400}{5} - 666.67 = 63.33$

5. $SSE = TSS - SSC = 133.33 - 63.33 = 70$

Source of variation	Sum of square	Degrees of freedom	Mean sum of squares	F ratio	Table F
Between columns.	$SSC = 63.33$ $= \cancel{666}$	$C-1=2$	$MSC = \frac{SSC}{C-1} = \frac{63.33}{2} = 31.66$	$F_c = \frac{MSC}{MSE} = \frac{31.66}{5.833}$	Dof (N_1, N_2) $(2, 12)$ T.v = 3.88
Between Errors.	$SSE = 70$	$E-1=12$	$MSE = \frac{SSE}{E-1} = \frac{70}{12} = 5.833$	$F_c = 5.427$	
	$TSS = 133.33$	$N-1=14$			

N.H H_0: There is No significance b/w columns.

A.H H_1: " " a " " "

L.O.S = 5%

T.v = 3.88

C.v = 5.427

Conclusion: C.v > T.v Reject Ho

Small Sample:

T-test [Mean Test]: $t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n-1}}}$

1. Test made on the breaking strength of 10 pieces of a metal wire gave the results 578, 572, 570, 568, 572, 570, 570, 572, 596, 584 kg. Test if the mean breaking strength of the wire can be assumed as 577 kg.

$\bar{x} = 575.2$ $S = 8.255$, $\mu = 577$, $n = 10$.

Sahitha

Null hypothesis: $\bar{x} = \mu$

Alternative " : $\bar{x} \neq \mu$

Level of significance: 5%.

Degrees of freedom: $n-1 = 9$

Test statistic: $t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n-1}}} = \frac{575.2 - 577}{\frac{8.256}{\sqrt{9}}} = -0.65$

$|t| = 0.65$

Conclusion : C.v < T.v Accept Ho

Sahitha.S

2. A machinest is expected to make engine parts with axcel diameter of 1.75 cm. A random sample of 10 parts shows as mean diameter 1.85 cm with the S.D of 0.1 cm. On the basis of this sample would you say the work of machinest is interior. (one tail)

$n = 10$, $\bar{x} = 1.85$, $S = 0.1$, $\mu = 1.75$

Null hypothesis: $\bar{x} = \mu$

Alternative " : $\bar{x} > \mu$

L.O.S : 5%

D.O.F : $n-1 = 9$

Test statistics: ~~A 1.9~~ $t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n-1}}} = \frac{1.85 - 1.75}{\frac{0.1}{\sqrt{9}}} = \frac{0.1}{\frac{0.1}{3}} = 3.00$

Conclusion : ~~to C.v < T.v~~ C.v > T.v

~~Reject Ho~~ ~~Accept Ho~~ Reject Ho.

3. A certain injection adminstered to each of 12 Patient resulted in the following increases of B.P 5, 2, 8, -1, 3, 0, 6, -2, 1, 5, 0, 4. Can it be concluded that the injection will be in general accompanied by an increase of bp?

$\bar{x} = 2.58$, $\mu = 0$, $n = 12$, $s = 2.95$

Null hypothesis : $\bar{x} = \mu$

Alternative " : $\bar{x} \neq \mu$

L.O.S : 5% 1.796

D.O.F : $n-1 = 11$

Test statistics : $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} = \frac{2.58 - 0}{\frac{2.95}{\sqrt{11}}} = \frac{2.58}{0.889} = 2.902$

Conclusion :

C.V > T.V Reject H_0

---

4. Two Sample : $t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

D.O.F : $n_1 + n_2 - 2$

1. The Mean height and S.D height of 8 Randomly chosen soldiers are 166.9 cm and 8.29 cm respectivily The corresponding values of 6 randomly chosen sailors are 170.3 and 8.50 cm. Based on this data can we conclude that soldier are in general shorter than sailors.

$n_1 = 8$, $\bar{x}_1 = 166.9$

$n_2 = 6$, $\bar{x}_2 = 170.3$

$S_1 = 8.29$, $S_2 = 8.50$.

Null hypothesis : $\bar{x}_1 = \bar{x}_2$

Alternative " : $\bar{x}_1 < \bar{x}_2$

L.O.S : 5% = 1.782

D.O.f : $n_1 + n_2 - 2 = 8 + 6 - 2 = 12$

Test statistics : $t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

$t = \frac{166.9 - 170.3}{\sqrt{\frac{8(8.29)^2 + 6(8.50)^2}{8+6-2}\left(\frac{1}{8} + \frac{1}{6}\right)}}$

$t = \frac{-3.4}{\sqrt{\frac{549.79 + 4433.5}{12}(0.125 + 0.166)}}$

$t = \frac{-3.4}{\sqrt{81.94(0.291)}} = \frac{-3.4}{\sqrt{23.844}} = \cancel{0.142}$

$= \frac{-3.4}{4.88} = -0.696 \Rightarrow |t| = 0.696$

Conclusion :

C.V < T.V Accept H_0

2. Two independent samples of sises 8 and 7 containg following values.

Sample I	19	17	15	21	16	18	16	14
Sample II	15	14	15	19	15	18	16	

is difference b/w is significant

Null hypothesis : $\bar{x}_1 = \bar{x}_2$

Alternative " : $\bar{x}_1 \neq \bar{x}_2$

L.O.S : 5% = 2.160

D.O.F : n_1+n_2-2 = 8+7-2 = 13

Test statistics :

$\bar{x}_1 = 17$, $S_1 = 2.12$, $\bar{x}_2 = 16$, $S_2 = 1.69$, $n_1 = 8$, $n_2 = 7$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{17-16}{\sqrt{\frac{8(2.12)^2+7(1.69)^2}{8+7-2}\left(\frac{1}{8}+\frac{1}{7}\right)}}$$

$$t = \frac{1}{\sqrt{\frac{35.95+19.92}{13}(0.125+0.14)}} = \frac{1}{\sqrt{4.29(0.265)}} = \frac{1}{\sqrt{1.13}} = \frac{1}{1.06}$$

$= 0.94$

Conclusion : C.V < T.V Accept Ho.

3. Samples of 2 types of electric bulbs were tested for length of life following data are obtaind.

	Size	Mean	S.D
Sample I	8	1234 hrs	36
" II	7	1036 hrs	40

is the difference in mean sufficient to warrant that type I bulbs are superior to type II.

Null hypothesis : $\bar{x}_1 = \bar{x}_2$

Alternative hypothesis : $\bar{x}_1 > \bar{x}_2$

L.O.S : 5% = 1.771

D.O.F : n_1+n_2-2 = 8+7-2 = 13

Test statistics : $\bar{x}_1 = 1234$, $S_1 = 36$, $n_1 = 8$, $\bar{x}_2 = 1036$, $S_2 = 40$, $n_2 = 7$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{1234-1036}{\sqrt{\frac{8(36)^2+7(40)^2}{8+7-2}\left(\frac{1}{8}+\frac{1}{7}\right)}}$$

$$= \frac{198}{\sqrt{\frac{10368+11200}{13}(0.265)}} = \frac{198}{\sqrt{439.65}} = \frac{198}{20.96} = 9.44$$

Conclusion : C.V > T.V Reject Ho

Normal Population [both T, F Test] $\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$

F-TEST [Variance Test] $\frac{\hat{\sigma}_1^2}{\sigma_2^2}$ (or) $\frac{\sigma_2^2}{\sigma_1^2}$ $\sigma_1^2 = \frac{n_1 s_1^2}{n_1-1}$ $\sigma_2^2 = \frac{n_2 s_2^2}{n_2-1}$

1. A sample of size 13 gave an estimated population variance of 3.0 while another sample of size 15 gave an estimate of 2.5 could both samples be from population with same variance.

Null hypothesis (Ho) : $\sigma_1^2 = \sigma_2^2$

Alternative " (H1) : $\sigma_1^2 \neq \sigma_2^2$

L.O.S : 5% = 2.53

D.O.F : (13-1, 15-1) = (12, 14)

Test Statics : $\frac{\sigma_1^2}{\sigma_2^2} = \frac{3}{2.5} = 1.2$

Conclusion : C.V < T.V ∴ Accept Ho

2. Two samples of sizes 9 & 8 gave the sum of squares of deviations from their respective means = 160 and 91 respectively. can they be regarded as drawn from the same Normal population.

T Test & F Test

Sample 1 $n_1 = 9$, $\Sigma(x_1-\bar{x}_1)^2 = 160$, $n_1 s_1^2 = 160$, $\sigma_1^2 = \frac{160}{8} = 20$

Sample 2 $n_2 = 8$, $\Sigma(x_2-\bar{x}_2)^2 = 91$, $n_1 s_1^2 = 91$, $\sigma_2^2 = \frac{91}{7} = 13$

Null Hypothesis (H_0) : $\sigma_1^2 = \sigma_2^2$

Alternative " (H_1) : $\sigma_1^2 \neq \sigma_2^2$

L.O.S : 5% = 3.64

D.O.F : $(9-1, 8-1) = (8, 7)$

Test Statistics : $\frac{\sigma_1^2}{\sigma_2^2} = \frac{20}{13} = 1.53$

Conclusion : C.V < T.V Accept H_0

3. Two independent samples of 8 & 7 items respectively had the following values of the variable

Sample 1 : 9 11 13 11 15 9 12 14

Sample 2 : 10 12 10 14 9 8 10 Do the two estimates of population variance differs significantly at 5% level of significance?

F-Test

Sample I	Sample II
$n_1 = 8$	$n_2 = 7$
$S_1 = 2.046$	$S_2 = 1.840$
$S_1^2 = 4.186$	$S_2^2 = 3.387$
$n_1 S_1^2 = 33.48$	$n_2 S_2^2 = 23.709$
$\sigma_1^2 = \frac{33.48}{7}$	$\sigma_2^2 = \frac{23.709}{6}$
$\sigma_1^2 = 4.782$	$\sigma_2^2 = 3.95$

Null hypothesis : $\sigma_1^2 = \sigma_2^2$

Alternative " : $\sigma_1^2 \neq \sigma_2^2$

L.O.S : 5% = 4.06

D.O.F : $(8-1, 7-1) = (7, 6)$

Test Statistics : $\frac{\sigma_1^2}{\sigma_2^2} = 1.210$

Conclusion : C.V < T.V

Accept H_0

Two Random samples gave the following data.

	Size	Mean	Variance
Sample I	8	9.6	1.2
Sample II	11	16.5	2.5

Can we conclude that the two samples have been drawn from Normal population.

F-TEST :

Null hypothesis (H_0) : $\sigma_1^2 = \sigma_2^2$

Alternative " (H_1) : $\sigma_1^2 \neq \sigma_2^2$

L.O.S : 5% ~~2.98~~ 3.64

D.O.F : ~~(7, 10)~~ (10, 7)

Test Statistics : $\frac{\sigma_2^2}{\sigma_1^2} = \frac{2.5}{1.2} = 2.08$

Conclusion : C.V > T.V ~~Accept~~ Reject H_0.

T-TEST :

Null hypothesis (H_0) : $\bar{x}_1 = \bar{x}_2$

Alternative " (H_1) : $\bar{x}_1 < \bar{x}_2$

L.O.S : 5% = 2.110

D.O.F : $n_1 + n_2 - 2 = 8 + 11 - 2 = 17$

Test Statistics : $\bar{x}_1 = 9.6$, $S_1 =$ ~~1.04~~ 1.2, $n_1 = 8$

$\bar{x}_2 = 16.5$, $S_2 =$ ~~2.5~~ ~~6.25~~ 2.5, $n_2 = 11$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{9.6 - 16.5}{\sqrt{\frac{8(1.\overset{2}{44}) + 11(\overset{2.5}{6.25})}{8 + 11 - 2}\left(\frac{1}{8} + \frac{1}{11}\right)}}$$

$$= \frac{-6.9}{\cancel{2.???}\ 0.} = \cancel{2.76} = -10.05 \quad |t| = 10.05$$

Conclusion: C.V > T.V Reject H_0

CHISQUARE: To test the fit. and attributes.

UNIT-V

Independent means $r = 0$.

$-1 \le r \le 1$

Correlation: Study of relation between variables.

Positive: ↑↑ ↓↓ , Negative: ↑↓

$r = -1$ ∴ x & y are perfectly negative

$r = 1$ ∴ " " " Positive

$$r = \frac{N\Sigma UV - \Sigma U \Sigma V}{\sqrt{N\Sigma U^2 - (\Sigma U)^2}\sqrt{N\Sigma V^2 - (\Sigma V)^2}} \qquad U = x - \bar{x}, \quad V = y - \bar{y}$$

Sahana

1. Calculate Correlation coefficient fro the following height in inches of fathers x and sons y.

x	y	U ($x-\bar{x}$)	V ($y-\bar{y}$)	uv	u^2	v^2
65	67	65−68= −3	67−69 −2	6	9	4
66	68	−2	−1	2	4	1
67	65	−1	−4	4	1	16
67	68	−1	−1	1	1	1
68	72	0	3	0	0	9
69	72	1	3	3	1	9
70	69	2	0	0	4	0
72	71	4	2	8	16	4
$\bar{x}=68$	$\bar{y}=69$			24	36	44

$$r = \frac{N\Sigma uv - \Sigma u \Sigma v}{\sqrt{N\Sigma u^2 - (\Sigma u)^2}\sqrt{N\Sigma v^2 - (\Sigma v)^2}}$$

$$= \frac{8(24) - 0(0)}{\sqrt{8(36) - 0^2}\sqrt{8(44) - 0^2}}$$

$$= \cancel{8}\ 0.6030$$

2. Find the correlation coefficient

x	y	u	v ($y-23$)	uv	u^2	v^2
5	16	-10	-7	70	100	49
10	19	-5	-4	20	25	16
15	23	0	0	0	0	0
20	26	5	3	15	25	9
25	30	10	7	70	100	49
15	22.8	0	-1	175	250	123

$$r = \frac{N\Sigma uv - \Sigma u \Sigma v}{\sqrt{N\Sigma u^2 - (\Sigma u)^2}\sqrt{N\Sigma v^2 - (\Sigma v)^2}}$$

$$= \frac{5(175) - 0(-1)}{\sqrt{5(\cancel{175}\,250) - (0)^2}\sqrt{5(123) - (-1)^2}}$$

$$= \frac{875}{35.35 \times 24.779}$$

$$= 0.998$$

Regression: Study of average relation between variables

x on y $= x - \bar{x} = r \cdot \frac{\sigma_x}{\sigma_y}(y - \bar{y})$ $\quad b_{xy}$ y on x : $y - \bar{y} = r \cdot \frac{\sigma_y}{\sigma_x}(x - \bar{x})$ $\quad b_{yx}$

$$b_{xy} \cdot b_{yx} = r \cdot \frac{\sigma_x}{\sigma_y} \times r \cdot \frac{\sigma_y}{\sigma_x} = r^2 \quad \therefore r = \pm\sqrt{b_{xy} \cdot b_{yx}}$$

Properties :

* One of the regression equ is greater than one and another is less than one.

* If regression is positive r should be positive

1. obtain the eqn of the lines of regression from the following data.

x	y	u	v	uv	u^2	v^2
1	9	-3	-2	6	9	4
2	8	-2	-3	6	4	69
3	10	-1	-1	1	1	1
4	12	0	1	0	0	1
5	11	1	0	0	1	0
6	13	2	-2	4	4	4
7	14	3	-3	9	9	9
4	11	0	0	26	28	28

$$r = \frac{N\Sigma uv - \Sigma u \Sigma v}{\sqrt{N\Sigma u^2 - (\Sigma u)^2}\sqrt{N\Sigma v^2 - (\Sigma v)^2}}$$

$$= \frac{7(27) - 0}{\sqrt{7(28) - 0^2}\sqrt{7(28) - 0^2}}$$

$$r = \cancel{0.034} \quad \frac{182}{\sqrt{196}\sqrt{196}} = \frac{182}{(14)(14)}$$

$$= 0.9285$$

$\sigma_x = 2$, $\sigma_y = 2$

Regression x on y

$$x - \bar{x} = r \cdot \frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

$$x - 4 = 0.9285\left(\frac{2}{2}\right)(y - 11)$$

$$x = 0.92854y - 10.2135 + 4$$

$$x = 0.92854 - 6.2135$$

Regression y on x

$$y - \bar{y} = r \cdot \frac{\sigma_x}{\sigma_y}(x - \bar{x})$$

$$y - 11 = 0.9285\left(\frac{2}{2}\right)(x - 4)$$

$$y = 0.9285x - 3.914 + 11$$

$$y = 0.9285x + 7.286$$

2. Obtain the Regression lines.

x	y	u	v	uv	u^2	v^2
67	126	-6	-29	174	36	841
64	125	-4	-30	120	16	900
65	139	-3	-16	48	9	256
69	145	+1	-10	-10	1	100
70	165	2	10	20	4	100
71	152	3	-3	-9	9	9
72	180	4	25	100	16	625
74	208	6	53	318	36	2809
68.3	155	3	0	761	127	5640

$$r = \frac{N\Sigma uv - \Sigma x \Sigma v}{\sqrt{N\Sigma u^2 - (\Sigma u)^2}\sqrt{N\Sigma v^2 - (\Sigma v)^2}}$$

$$= \frac{8(761) - 1(0)}{\sqrt{8(127) - (3)}\sqrt{8(5640) - 0}}$$

$$= \frac{6088}{(31.82)(212.4)}$$

$$= \frac{6088}{6760.4} \quad \cancel{0.899} = 0.9031$$

$\sigma_x = 3.98$, $\sigma_y = 26.55$

Rregression x on y

$$x - \bar{x} = r \cdot \frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

$$x - 68.3 = 0.9031\left(\frac{3.98}{26.55}\right)(y - 155)$$

$$x - 68.3 = 0.1349y - 20.983$$

$$x = 0.13474 + 47.49$$

Regressio y on x

$$y - \bar{y} = r \cdot \frac{\sigma_x}{\sigma_y}(x - \bar{x})$$

$$y - 155 = 0.9031\left(\frac{26.55}{3.96}\right)(x - 68.3)$$

$$y - 155 = 6.0548x - 413.5$$

$$y = 6.0548x - 258.54$$

3. From following data two regression eqn, the coefficient correlation b/w Maths & Stat

And Most likely maths & stat when maths in maths are 30.

x	y	u	v	u^2	v^2	uv
25	43	-7	5	49	25	-35
28	46	-4	8	16	64	-32
35	49	3	11	9	121	33
32	41	0	3	0	9	0
31	36	-1	-2	1	4	2
36	32	4	-6	16	36	-24
29	31	-3	-7	9	49	21
38	30	6	-8	36	64	-48
34	33	2	-5	4	25	-10
32	39	0	1	0	1	0
32	38			140	~~418~~ 398	-93

$r = \frac{N\Sigma uv - \Sigma u \Sigma v}{\sqrt{N\Sigma u^2 - (\Sigma u)^2}\sqrt{N\Sigma v^2 - (\Sigma v)^2}}$

$r = \frac{10(-93) - 0}{\sqrt{10(140) - 0}\sqrt{10(398) - 0}}$

$r = -930$

$r = -0.3939$

$\sigma_x = 3.74$, $\sigma_y = 6.308$

Regression x on y

$x - \bar{x} = r \cdot \frac{\sigma_x}{\sigma_y}(y - \bar{y})$

$x - 32 = -0.3939\left(\frac{3.74}{6.308}\right)(y - 38)$

$x - 32 = -0.23354y + 8.8746$

$x = -0.2335y + 40.8799$

$x = 40.8799 - 0.2335y$

Regression y on x

$y - \bar{y} = r \cdot \frac{\sigma_y}{\sigma_x}(x - \bar{x})$

$y - 38 = -0.3939\left(\frac{6.308}{3.74}\right)(x - 32)$

$y - 38 = -0.6643x + 21.2596$

$y = -0.6643x + 59.2577 \Rightarrow y = 59.2577 - 0.6643x$

When $x = 30$ then y

$y = 59.2577 - 0.6643(30)$

$y = 39.3287 \simeq 39$

4. For ten observation on price x and supply y the following data obtained.

$\Sigma x = 130$, $\Sigma y = 220$, $\Sigma x^2 = 2288$, $\Sigma y^2 = 5506$, $\Sigma xy = 3467$

Obtain the lines of Regression y on x and estimate supply when the price is 16 units.

$r = \frac{N\Sigma uv - \Sigma u \Sigma v}{\sqrt{N\Sigma u^2 - (\Sigma u)^2}\sqrt{N\Sigma v^2 - (\Sigma v)^2}}$

$= \frac{10(3467) - (130)(220)}{\sqrt{10(2288) - (130)^2} - \sqrt{10(5506) - (220)^2}}$

$= 6070$

5. The 2 lines of regression are $8x - 10y + 66 = 0$, $40x - 18y - 214 = 0$. The variance is 9, find (i) the mean value of x & y (ii) correlation coefficient b/w x & y.

Mode 3 times. Eqn 1, Unknown 2.

Sol: (i)

$8\bar{x} - 10\bar{y} = -66 \rightarrow (1)$, $40\bar{x} - 18\bar{y} = 214 \rightarrow (2)$

$\bar{x} = 13$, $\bar{y} = 17$

(ii)

	x on y $\frac{-\text{coeff of } y}{\text{coeff of } x}$	y on x $\frac{-\text{coeff of } x}{\text{coeff of } y}$
$8x - 10y + 66 = 0$	✗ $b_{xy} = \frac{-(-10)}{8} = \frac{10}{8} > 1$	$b_{yx} = \frac{-8}{-10} = \frac{8}{10} < 1$
$40x - 18y - 214 = 0$	✓ $b_{xy} = \frac{-(-18)}{40} = \frac{18}{40} < 1$	✗ $b_{yx} = \frac{-40}{-18} = \frac{40}{18} > 1$

$b_{xy} = \frac{18}{40}$

$b_{yx} = \frac{8}{10}$

$r = \pm\sqrt{\frac{18}{40} \times \frac{8}{10}}$ $r = +0.6$

6. $3x + 12y - 19 = 0$, $9x + 3y - 46 = 0$ are the lines of regression. Find the mean of x & y and the correlation coefficient b/w x & y also obtain estimate of x for given $y = 1$.

(i) $\bar{x} = 5$ $\bar{y} = 0.33$

(ii)

	x on y	y on x
$3x + 12y - 19 = 0$	✗ $b_{xy} = \frac{-12}{3} = -4$	✓ $b_{yx} = \frac{-3}{12} = -0.25$
$9x + 3y - 46 = 0$	✓ $b_{xy} = \frac{-3}{9} = -0.33$	✗ $b_{yx} = \frac{-9}{3} = -3$

$b_{xy} = -0.33$

$b_{yx} = -0.25$

$r = \pm\sqrt{0.0825}$

$r = -0.287$

(iii) $9x + 3y - 46 = 0 \Rightarrow 9x + 3(1) - 46 = 0$

$9x + 3 - 46 = 0$

$9x - 43 = 0$

$x = \frac{43}{9} = 4.777$

7. The regression eqn of x on y is $3y - 5x + 108 = 0$. If the mean of y is 44 and the variance of x where $\left(\frac{9}{16}\right)^{th}$ of variance of y. Find mean of x and r.

$3y - 5x + 108 = 0$ $\bar{y} = 44$, $v(x) = \frac{9}{16} v(y)$

$3(44) - 5(\bar{x}) = -108$

$132 - 5\bar{x} = -108$

$-5\bar{x} = -108 - 132$

$-5\bar{x} = -240 \Rightarrow \boxed{\bar{x} = 48}$

$\sigma_x^2 = \frac{9}{16}\sigma_y^2$

$\sigma_x = \frac{3}{4}\sigma_y$

$\frac{\sigma_x}{\sigma_y} = \frac{3}{4} \rightarrow (1)$

$3y - 5x + 108 = 0$

$5x = 3y + 108$

$x = \frac{3}{5}y + \frac{108}{5}$

$b_{xy} = \frac{3}{5}$

$r \cdot \frac{\sigma_x}{\sigma_y} = \frac{3}{5} \Rightarrow r \cdot \frac{3}{4} = \frac{3}{5}$ $r = \frac{4}{5} = \boxed{0.8}$

8. $y = 0.516\bar{x} + 33.73$, $x = 0.5124 + 32.52$

$y = 68.63$

$x = 67.63$

9. The regression lines are $3x+2y=26$, $6x+y=31$ find $\bar{x}, \bar{y}, r$ $\bar{x}=4$, $\bar{y}=7$

	x on y	y on x
$3x+2y=26$	$b_{xy}=\frac{-(2)}{3}=-0.66$	$b_{yx}=\frac{-3}{2}=-1.5$
$6x+y=31$	$b_{xy}=\frac{-1}{6}=-0.166$	$b_{yx}=\frac{-6}{1}=-6$

$b_{xy}=-0.166$

$b_{yx}=-1.5$

$r=\pm\sqrt{-0.166\times-1.5}$

$r=-0.49$

$r=-0.5$

y on x $= 3x+2y=26$

$b_{yx}=\frac{-3}{2} \Rightarrow r\cdot\frac{\sigma_y}{\sigma_x}=-\frac{3}{2} \Rightarrow (-0.5)\frac{\sigma_y}{1}=\frac{-3}{2}$

$\boxed{\sigma_y=3}$

ONE WAY

A Completely randomist design experiment with 10 plots and 3 treatments gave the following results.

Plot No:	1	2	3	4	5	6	7	8	9	10
Treatment:	A	B	C	A	C	C	A	B	A	B
Yield:	5	4	3	7	5	1	3	4	1	7

Analyse the result for treatment effects.

Sol:

$A\ x_1$	$B\ x_2$	$C\ x_3$		x_1^2	x_2^2	x_3^2
5	4	3		25	16	9
7	4	5		49	16	25
3	7	1		9	49	1
1	-	-		1	-	-
16	15	9	40	84	81	35

1. Find $T, N = 40, 10$
2. Find $\frac{T^2}{N} = \frac{40^2}{10} = 160$
3. $TSS = \Sigma x_1^2+\Sigma x_2^2+\Sigma x_3^2 - \frac{T^2}{N} = 84+81+35-160$
$=40$
4. $SSC = \frac{(\Sigma x_1)^2}{n_1}+\frac{(\Sigma x_2)^2}{n_2}+\frac{(\Sigma x_3)^2}{n_3}-\frac{T^2}{N} \Rightarrow \frac{(16)^2}{4}+\frac{(15)^2}{3}+\frac{(9)^2}{3}-160$
$=64+75+27-160=6$
5. $SSE = TSS - SSC = 40-6=34$

Source of Variation	Sum of square	Degree of freedom	Mean sum of square	F ratio	Table F
Between columns	$SSC=6$	$C-1=2$	$MSC=\frac{SSC}{C-1}=\frac{6}{2}=3$		Dof (7, 2)
				$F_c=\frac{MSE}{MSC}$	T.V = 19.40
Between Errors.	$SSE=34$	$E-1=7$	$MSE=\frac{SSE}{E-1}=\frac{34}{7}$ $=4.857$	$=\frac{4.857}{3}$ C.V $=1.619$	
	$TSS=40$	$N-1=9$			

Null hypothesis: There is no significance d/b colums.

A. H H_1 : " " " difference b/w ".

α: 0.5 : 5%

Conclusion: C.V < T.V.E Accept H_0

2WAY classification (or) Randomised Block Design:

1. 4 doctors each test four treaments for a certain dieases and observed the number of day each patient takes to recover. The results are as folows (Recovory Time in days).

Doctor	Treatment 1	2	3	4
A	10	14	19	20
B	11	15	17	21
C	9	12	16	19
D	8	13	17	20

Discuss the differ b/w (i) Doctors and (ii) Treatements.

Subtract the data from the Number 10

	x_1	x_2	x_3	x_4		x_1^2	x_2^2	x_3^2	x_4^2
y_1	0	4	9	810	23 Σy_1	0	16	81	100
y_2	1	5	7	11	24 Σy_2	1	25	49	121
y_3	-1	2	6	9	16 Σy_3	1	4	36	81
y_4	-2	3	7	10	18 Σy_4	4	9	49	100
	-2	14	29	40	81	6	54	215	402

1. Find $T, N = 81, 16$

2. Find $\frac{T^2}{N} = \frac{81^2}{16} = 410.06$

3. $TSS = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 + \Sigma x_4^2 - \frac{T^2}{N} = 6 + 54 + 25 + 402 - 410.06$

$= 266.94$

4. $SSC = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} + \frac{(\Sigma x_4)^2}{n_4} - \frac{T^2}{N} = \frac{4}{4} + \frac{196}{4} + \frac{841}{4} + \frac{1600}{4} - \cancel{28}\ 410$

$= 1 + 49 + 210.25 + 400 - 410$

$= 250.19$

5. $SSr = \frac{(\Sigma y_1)^2}{m_1} + \frac{(\Sigma y_2)^2}{m_2} + \frac{(\Sigma y_3)^2}{m_3} + \frac{(\Sigma y_4)^2}{m_4} - \frac{T^2}{N} = \frac{529}{4} + \frac{576}{4} + \frac{256}{4} + \frac{324}{4}$

$- 410.06$

$= 11.19$

6. $SSE = TSS - SSC - SSr = 266.94 - 250.19 - 11.19$

$= 5.56$

Source variable	Sum of square	D.o.f	Mean sum of square	F ratio	Table F
Between columns	$SSC = 250.19$	$C-1 = 3$	$MSC = \frac{250.19}{3} = 83.396$	$F_c = \frac{\cancel{250.19}\ 83.396}{0.617} = 135.163$	$F_c(3,9) = 3.14$
b/w Rows	$SSr = 11.19$	$R-1 = 3$	$MSr = \frac{11.19}{3} = 3.73$	$F_r = \frac{\cancel{11.19}\ 3.73}{0.617} = 6.045$	$F_r(3,9) = 3.14$
b/w errors	$SSE = 5.56$	$E-1 = 9$	$MSE = \frac{5.56}{9} = 0.617$		
	$TSS = 266.94$	$N-1 = 15$			

Null Hypothesis (H_0): There is no significance difference b/w Colums & rows.

Alternative " (H_1): There is a " " " "

L.o.s : 5%

Conclusion : T.v = 3.14 , C.V.F_c > T.V so Reject H_0;
C.V F_r > T.v

2. A Company appoints 4 sales man A B C D and observes their sales in 3 seasons summer, winter, Monsoon. The figures (in lakh of rs) are given. carry out an analysis of variance.

Season	A	B	C	D
	sales			
Summer	36	36	21	35
Winter	28	29	31	32
Monsoon	26	28	29	29

Subtract the data from the Number 25

	x_1	x_2	x_3	x_4		x_1^2	x_2^2	x_3^2	x_4^2
y_1	11	11	-4	10	28	121	121	16	100
y_2	3	4	6	7	20	9	16	36	49
y_3	1	3	4	4	12	1	9	16	16
	15	18	6	21	60	131	146	68	165

1. Find $T, N = 60, 12$

2. $\frac{T^2}{N} = \frac{60^2}{12} = 300$

3. $SSC = \frac{225}{3} + \frac{324}{3} + \frac{36}{3} + \frac{441}{3} - 300 = 42$

4. $TSS = 131 + 146 + 68 + 165 - 300 = 210$

5. $SSr = \frac{784}{4} + \frac{400}{4} + \frac{144}{4} - 300 = 32$

6. $SSE = 210 - 42 - 32 = 136$

Source of Variable	Sum of Square	D.o.f	Mean sum of square	F ratio	Table F
b/w columns	SSC = 42	c-1 = 3	$MSC = \frac{42}{3} = 14$	$F_C = \frac{22.7}{14} = 1.6$	$F_C(6,3) = 8.79$
b/w rows	SSr = 32	r-1 = 2	$MSr = \frac{32}{2} = 16$	$F_r = \frac{22.7}{16} = 1.4$	$F_r = (6,2) = 19.40$
b/w errors	SSE = 136	E-1 = 6	$MSE = \frac{136}{6} = 22.7$		
	TSS = 210	N-1 = 11			

Null hypothesis (H_0): There is No significance d/b rows & columns

Alternative (H_1): " " a " " "

α.o.s : 5%

Conclusion : C.V F_C < T.V, C.V F_r < T.V ∴ Accept H_0.

Three Way ANOVA:

1. Analyse the variance in the following Latin square of yield (in kgs) of paddy where ABCD denote the different methods of cultivation.

D122	A121	C123	B122
B124	C123	A122	D125
A120	B119	D120	C121
C122	D123	B121	A122

Examine whether different method of cultivation have given significantly different yields.

Subtract data from 120

	x_1	x_2	x_3	x_4		x_1^2	x_2^2	x_3^2	x_4^2
y_1	2	1	3	2	8 Σy_1	4	1	9	4
y_2	4	3	2	5	14 Σy_2	16	9	4	25
y_3	0	-1	0	1	0 Σy_3	0	1	0	1
y_4	2	3	1	2	8 Σy_4	4	9	1	4
	8	6	6	10	30 T	24	20	14	34

Table 2:

A	B	C	D
1	2	3	2
2	4	3	5
0	-1	1	0
2	1	2	3
5 Σk_1	6 Σk_2	9 Σk_3	10 Σk_4

1. Find $T, N = 30, 16$

2. $\frac{T^2}{N} = \frac{900}{16} = 56.25$

3. $TSS = 24+20+14+34-56.3 = 35.75$

4. $SSC = \frac{64}{4} + \frac{36}{4} + \frac{36}{4} + \frac{100}{4} - 56.25 = 2.8$

5. $SSr = \frac{64}{4} + \frac{196}{4} + \frac{0}{4} + \frac{64}{4} - 56.25 = \cancel{12.8}$ 24.75

6. $SSk = \frac{25}{4} + \frac{36}{4} + \frac{81}{4} + \frac{100}{4} - 56.25 = 4.25$

7. $SSE = 35.75 - 2.8 - 24.75 - 4.25 = 4$

Source of Variable	Sum of Square	D.o.f	Mean of Square	F ratio	Table F
b/w Column	SSC = 2.75	c-1 = 3	$MSC = \frac{2.75}{3} = 0.92$	$F_c = \frac{0.9}{0.7} = 1.3$	$F_c(3,6) = 4.06$
b/w rows	SSr = 24.75	r-1 = 3	$MSr = \frac{24.75}{3} = 8.3$	$F_r = \frac{8.3}{0.7} = 11.9$	$F_r(3,6) = 4.06$
b/w letters	SSk = 4.25	k-1 = 3	$MSk = \frac{4.25}{3} = 1.4$	$F_k = \frac{1.4}{0.7} = 2$	$F_k(3,6) = 4.06$
b/w error	SSE = 4	E-1 = 6	$MSE = \frac{4}{6} = 0.7$		
	TSS = 35.75	N-1 = 15			

Null hypothesis (H_0): There is No significance d/b rows colums & letters.

Alternative " (H_1): " " a " " "

L.O.S : 5%.

Conclusion : C.V F_c < T.V, C.V F_r > T.V, C.V F_k < T.V Reject H_0

2. Analyse the Variance in the following Latin Square.

20 B	17 C	25 D	34 A
23 A	21 D	15 C	24 B
24 D	26 A	21 B	19 C
26 C	23 B	27 A	22 D

Subtract Data from 20

	x_1	x_2	x_3	x_4		x_1^2	x_2^2	x_3^2	x_4^2
y_1	0	-3	5	14	16	0	9	25	196
y_2	3	1	-5	4	3	9	1	25	16
y_3	4	6	1	-1	10	16	36	1	1
y_4	6	3	7	2	18	36	9	49	4
	13	7	8	19	47	61	55	100	217

A	B	C	D
14	0	-3	5
3	4	-5	1
6	1	-1	4
7	3	6	2
30	8	-3	12

1. T, N = 47, 16

2. $\frac{T^2}{N} = 138.1$

3. TSS = 61 + 55 + 100 + 217 = 294.9

4. $SSC = \frac{169}{4} + \frac{49}{4} + \frac{64}{4} + \frac{361}{4} - 138.1 = 22.7$

5. SSr = 34.2

6. SSK = 141.2

7. SSE = 96.5

Source of Variable	Sum of Square	D.o.f	Mean of Square	F ratio	Table F
b/w columns	SSC = 22.7	C-1 = 3	$MSC = \frac{22.7}{3} = 7.6$	$F_c = \frac{16.1}{7.6} = 2.1$	$F_c(6,3) = 8.79$
b/w rows	Sr = 34.2	r-1 = 3	$MSR = \frac{34.2}{3} = 11.4$	$F_r = \frac{16.1}{11.4} = 1.4$	$F_c(6,3) = 8.79$
b/w letters	SSk = 141.2	k-1 = 3	$MSk = \frac{141.2}{3} = 47.1$	$F_k = \frac{47.1}{16.1} = 2.9$	$F_k(3,6) = 4.06$
b/w error	SSE = 96.5	E-1 = 6	$MSE = \frac{96.5}{6} = 16.1$		
	TSS = 294.9	N-1 = 15			

Null hypothesis (H_0): There is No significance d/b rows, columns, letters

Alternative (H_1): " " a " " "

L.O.S : 5%.

Conclusion : C.V F_c < T.V, C.V F_r < T.V, C.V F_k < T.V

Accept H_0.

CONTROL CHART: is a kind of graphical representation used for presenting a sequence of suitable sample characteristics.

It is used to detect any unusual, acquiring in the process of production that put the process out of control.

1. A control chart has 3 horizontal line from the right side of the vertical lines and parallel to the base line of the chart. A vertical line represent the quality statistic of each sample. The base line shows the sample Number.

The 3 horizontal lines are control linese. They are

(i) Central line (CL)

(ii) Upper Control limit (UCL) $\bar{\bar{X}} + A_2\bar{R}$

(iii) Lower Control limit (LCL) $\bar{X} - A_2\bar{R}$

Types: (i) $\bar{X}$ chart, (ii) R chart

$\bar{X}$-chart: $LCL = \bar{\bar{x}} - A_2\bar{R}$, $UCL = \bar{\bar{x}} + A_2\bar{R}$

1. The following data gives reading for 10 samples of size 6 each in the production of a certain component.

Sample:	1	2	3	4	5	6	7	8	9	10
Mean $\bar{x}$:	383	508	505	582	557	337	514	614	707	753
Range R:	95	128	100	91	68	65	148	28	37	80

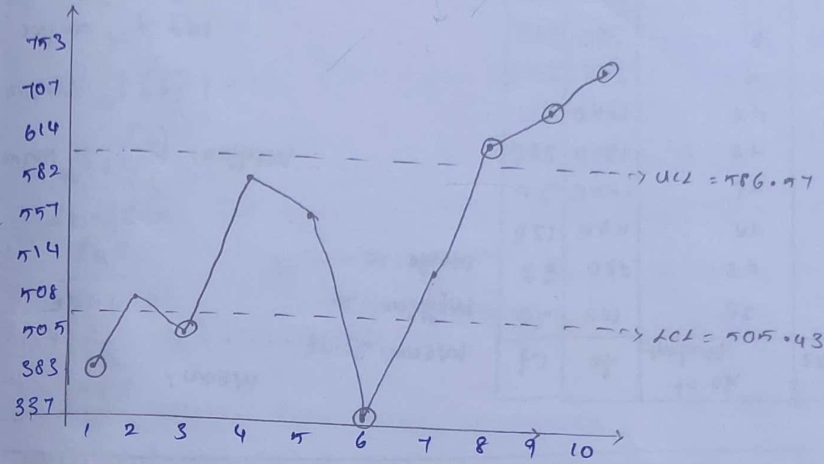
Draw a control chart for $\bar{X}$.

Sol: $\bar{\bar{x}} = \frac{\Sigma\bar{x}}{10} = \frac{5460}{10} = 546$, $\bar{R} = \frac{\Sigma R}{10} = \frac{840}{10} = 84$

$n = 6$, $A_2 = 0.483$

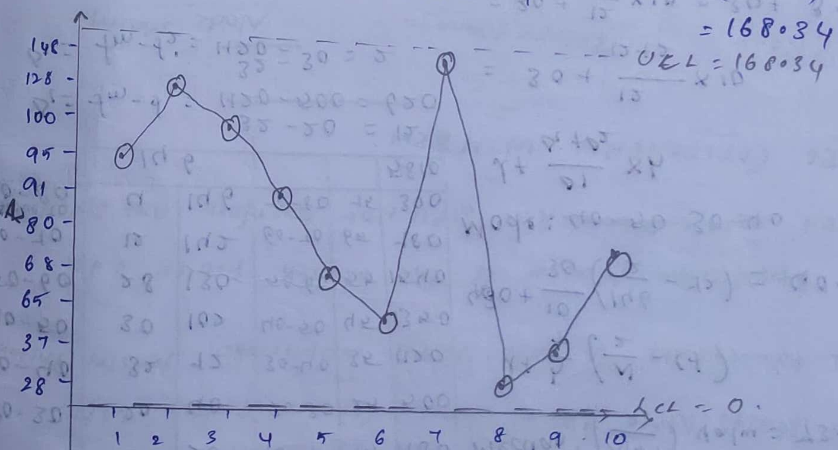
$LCL = \bar{\bar{x}} - A_2\bar{R} \Rightarrow 546 - (0.483)84 = 507.43$

$UCL = \bar{\bar{x}} + A_2\bar{R} \Rightarrow 546 + (0.483)84 = 586.57$



As some of the points fall outside the control, the process is out of control.

R-Chart: $LCL = D_3\bar{R} \Rightarrow 84(0) = 0$, $UCL = D_4\bar{R} \Rightarrow (2.004)(84) = 168.34$



The R chart shows that all the values of sample range lies within control limits. hence the process variability is under control.

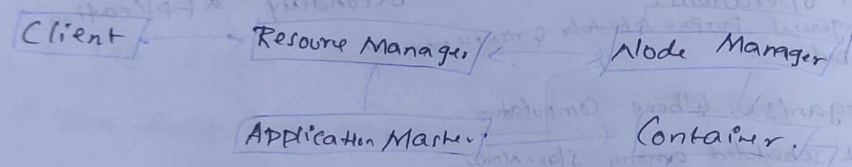
YARN Application Components.

Three : The client, application Master, Container.

=> Starts with YARN client communicating with RM to create a new YARN application Master instance.

=> Application Master is the master process of a YARN application. AM will not perform any specific work.

CORE COMPONENTS IN YARN:



The shape - ((1000, 2), (1000))

1.

Marks	No. of students	fx	cf
5	20	100	20
10	43	430	63
15	75	1125	138
20	76	1520	214
25	72	1800	286
30	45	1350	331
40	9	360	340
45	8	360	348
50	50	2500	398
	398	9545	

Mean: 23.98
Median: 20
Mode: 20

Mean: $\frac{\Sigma fx}{\Sigma f} = \frac{9545}{398} = 23.98$

Median: $\left(\frac{N+1}{2}\right)^{th}$ value

$= \left(\frac{399}{2}\right)^{th}$ value

$= 199.5^{th}$ value

$= \cancel{76}\ 20.$

Mode: Highest freque = 20.

Marks	No. of Student f	c.f	class in	m	fm
Below 10	8	8	0-10	5	40
10-20	12	20	10-20	15	180
20-30	20	40	20-30	25	500
30-40	32	72	30-40	35	1120
40-50	30	102	40-50	45	1350
50-60	28	130	50-60	55	1540
60-70	12	142	60-70	65	780
Above 70 ~~70-80~~	4	146	70-80	75	300
	146				5810

$\Delta_1 = f_m - f_1 = \cancel{1120 - 500 = 620}\ 32 - 20 = 12$

$\Delta_2 = f_m - f_2 = \cancel{1120}\ 32 - 30 = 2$

Mean: $\frac{\Sigma fm}{\Sigma f} = \frac{5810}{146} = 39.79$

Median: $\left(\frac{N+1}{2}\right)^{th}$ value $= 73.5^{th}$ value

$l + \frac{h}{f}\left(\frac{N}{2} - cf\right)$

$30 + \frac{10}{30}\left(\frac{146}{2} - 72\right) = 40.3$

Mode: ~~40-50~~ 30-40

$l + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$

$= 30 + \frac{12}{12+2} \times 10$

$= 30 + \frac{12}{14} \times 10 = 30 + 8.57$

$= 38.57$