Revision

Describe → Stats

{ Cleaning ↓ Transforming } Programs → R prog

← Modeling → Algorithms
↓
Insight → Reports → Excel | Tableau

Population → Descriptive) Results

Sample → Inferential) Battery

Describe ⟶ Central tendency → Mean, Median, Mode
Quantile, Quartile, percentile

→ Variability → Variance, S.D
→ Range

→ Shape → Skewness, Kurtosis, Symmetry

Mean → Average → 1, 2, 3, 4, 5 ⟹ $\frac{15}{5}$ = ③

$\underline{\text{Mean}} \rightarrow$ Average $\rightarrow$ 1, 2, 3, 4, 5 $\Rightarrow \dfrac{15}{5} = \boxed{3}$

$\underline{\text{Median}} \Rightarrow$ Middle $\underline{no.} = 1, 1, 2, 2, 3, 3, 4, 4 = \left(\dfrac{n+1}{2}\right)^{th}$ element

$$\dfrac{2+3}{2} = \boxed{2.5}$$

$$\dfrac{9}{2} = 4.5^{th} \text{ element}$$

$\underline{\text{Mode}} \rightarrow$ Repeating many times

$$= 1, 1, 1, 1, 2, 2, 3, 3, 3, 3$$

$\rightarrow 3 \rightarrow$ Unimoda

$(3, 1) =$ bimodal

$> 3 \, no's \rightarrow$ Multimodel

$\underline{\text{Quantile}} \Rightarrow$

μ

process of Dividing data

$D_1 \qquad\qquad D_2 \qquad$ Tertile

25%   25%   25%   25%   Inter

$\boxed{D_1}$ D_2 50% $\boxed{D_3}$ Quantile Range

$$IQR = Q_3 - Q_1$$
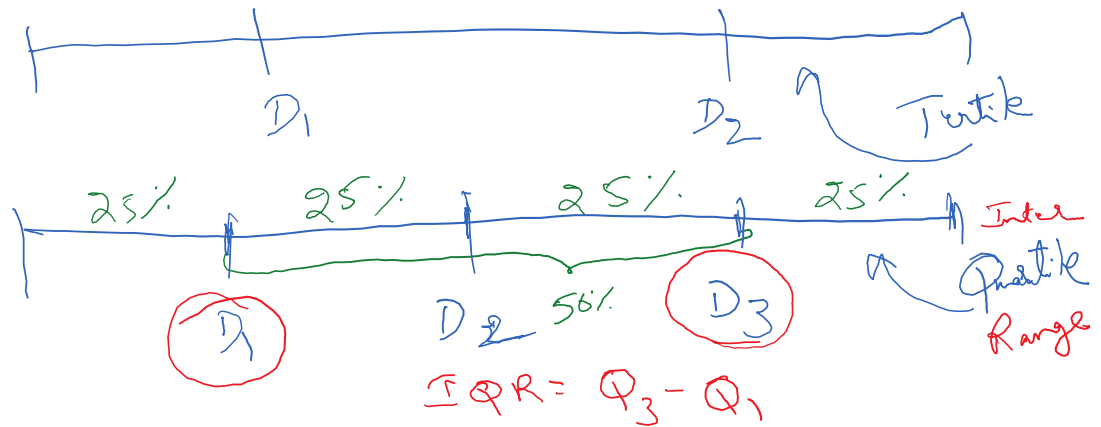
Percentile

Centile

$\underline{\text{Range}} \Rightarrow$

1 $\leftarrow$ Range $\longleftarrow$ 10

$(10 - 1) = \boxed{9}$
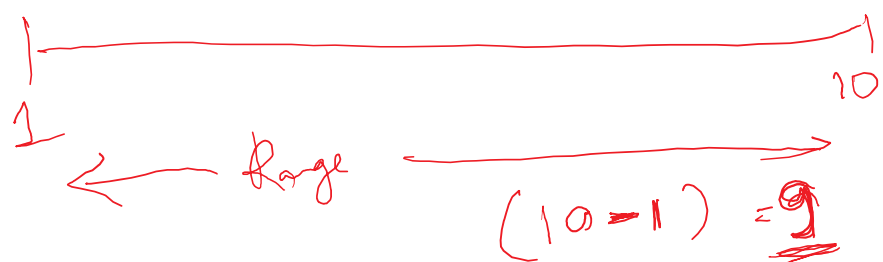
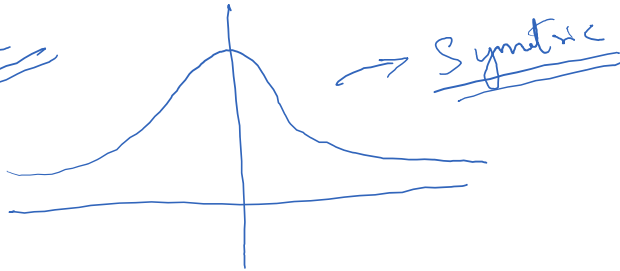**Variance** $\sigma^2 = \dfrac{\Sigma(x_i - M)^2}{N}$ $\Rightarrow$ Squared units $\rightarrow$ Remove
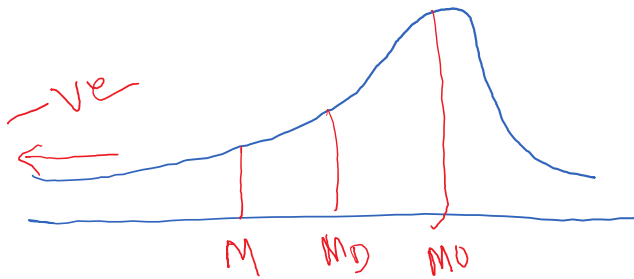
Standard dev $\sqrt{\sigma^2} = \sigma \rightarrow$ standard deviation $= \sqrt{\dfrac{\Sigma(x_i - M)^2}{N}} = \sqrt{\sigma^2}$
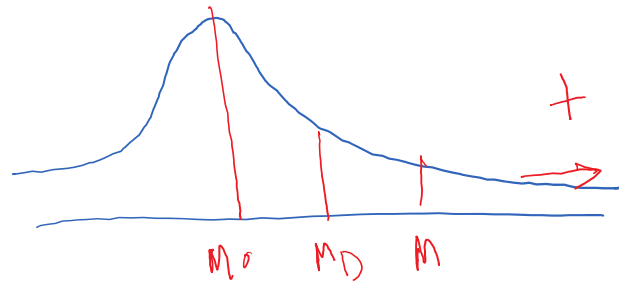

Shape

 $\rightarrow$ Symetric

Skewed $\rightarrow$ X symetry

$-ve$

M   MD   MO

Mo   MD   M $+$

Pearson's Co.efficit of $S_k = \left(\dfrac{Mean - Md}{SD}\right) 3$

Kurtosis $\rightarrow$ **Peakedness**
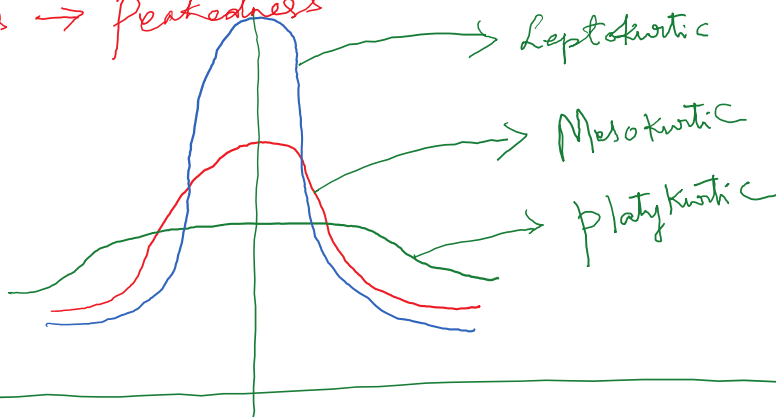
$\rightarrow$ Leptokurtic

$\rightarrow$ Mesokurtic

$\rightarrow$ Platykurtic

Grouped data

0 - 35	25, 26 29
35 - 75	43, 56,
	10 98

Ungrouped data.

26, 48, 35, 78, 83, 89
64, 96, 91, 83

35-75	47,...
75-100	28,98

64, 96, 91, 85

Raw data

Class MP
Relative freq
Cumulative freq

Freq dist		No's
0-35	25, 26, 27, 28, 29	5
35-75	20 no's	20
75-100	78, 79, 90, 96.	4

Frequency

Class mp → 0-35 ⇒ $\left(\dfrac{35-0}{2}\right)$ = $\boxed{17.5}$

	Freq dist		No's
C①	0-35	25, 26, 27, 28, 29	5
C②	35-75	20 no's	20
C③	75-100	78, 79, 90, 96.	4
			29

$C_1 = \dfrac{5}{29} =$

$C_2 = \dfrac{20}{29}$

$C_3 = \dfrac{4}{29}$

Relative freq = $\dfrac{\text{class freq}}{\text{Total freq}}$

Cumulative

$CF_1 = \dfrac{5}{29} = C_1$

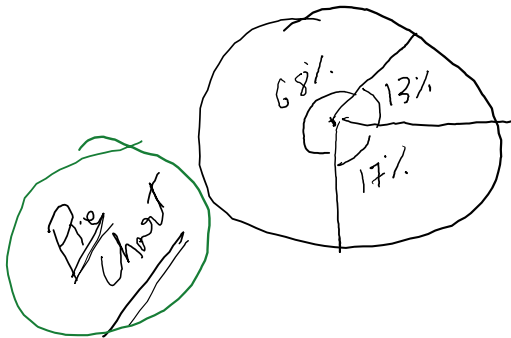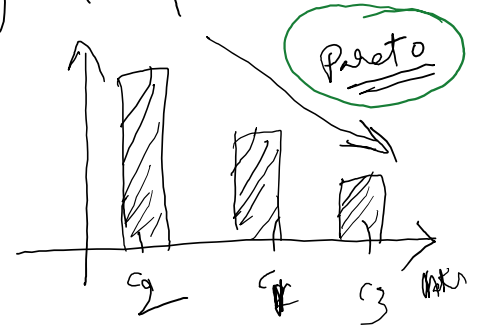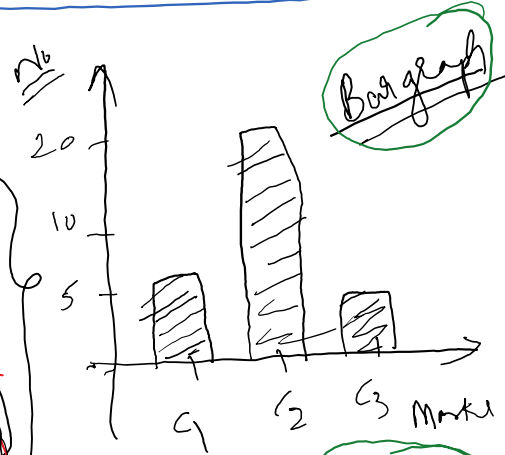$CF_2 = \boxed{C_0 + C_1}$

$CF_{C_3} = \text{Total freq}$

New Section 3 Page 4

$$CF_2 = C_2 + C_1$$

$$CF_3 = C_3 + C_2 + C_1$$

Graphs

Freal dist		No's
C_1	0-35	25, 26, 27, 28, 29	5
C_2	35-75	20 no's	20
C_3	75-100	78, 79, 90, 96.	4
			29

Bar graph

C_1 C_2 C_3 Marks

Pareto

C_1 C_2 C_3 Marks

Pie Chart

68%   13%   17%

$C_1 - 17\%$
$C_2 - 68\%$
$C_3 - 13\%$
$= 100\%$

Discrete data.
(Qualitative)

(Quantitative)
Continuous data

R Freq 40 20 10

X   Y   Z X_1
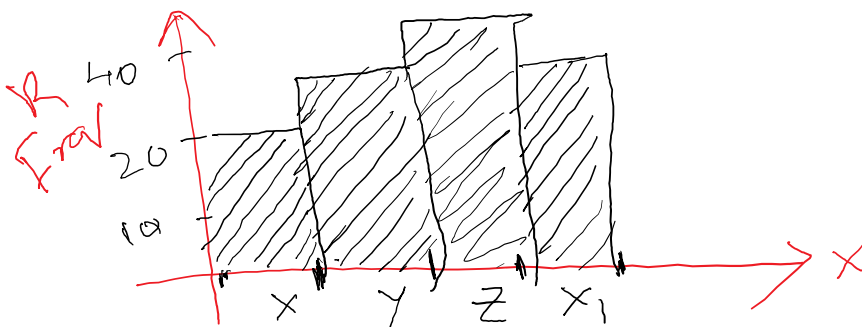
Histogram

Freq Polygon

IMP

	Event	$P(x)$
x	X	20
y	Y	30
z	Z	40
x_1	X_1	35

R Freq

x y z x_1 → class M.P

F_r

class M.P

x   y   Z x_1

Freq Cumulation

y

x y z x_1 m.P

X Y Z X_1

X

C_1 C_2 C_3 C_4

Ojive

$\uparrow$ curve

Scatterplot

y

75 —
50 —
25 —

0   1 2 3 4 5 6 7 8 x
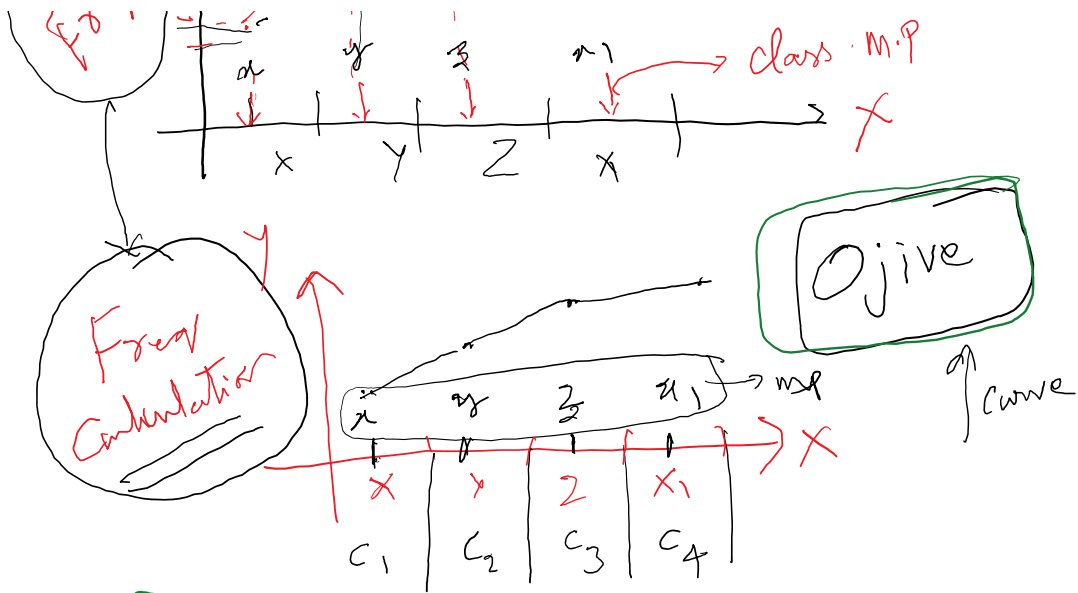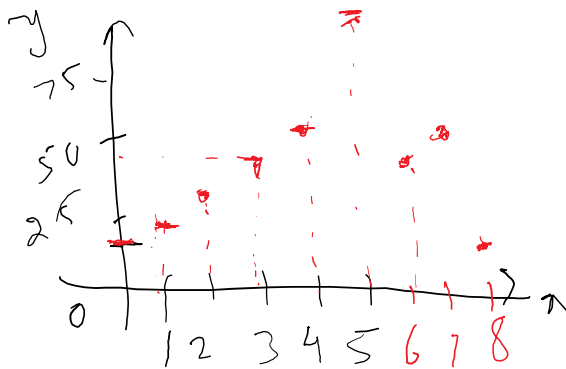
x	y
0	16
1	18
2	23
3	36
4	48
5	96
6	36
7	43
8	12

Chart Title

120
100
80
60
40
20
0
0   2   4   6   8   10