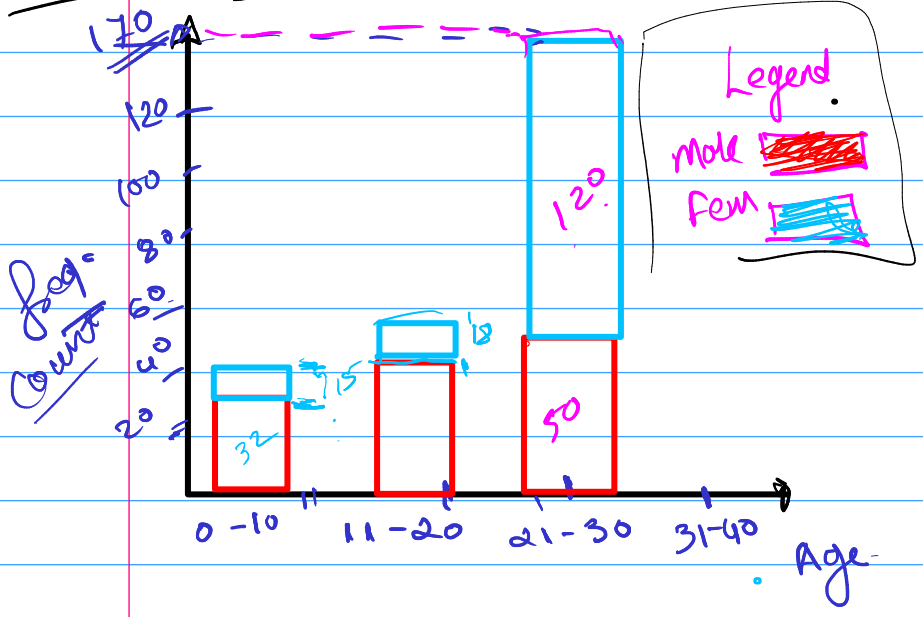
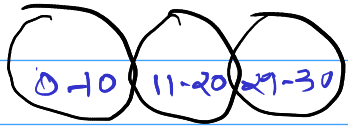
Stack Bar chart

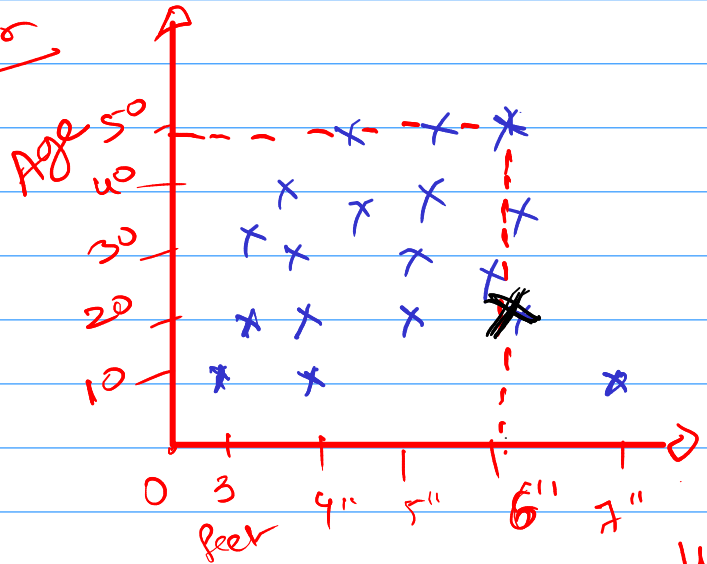


Bin

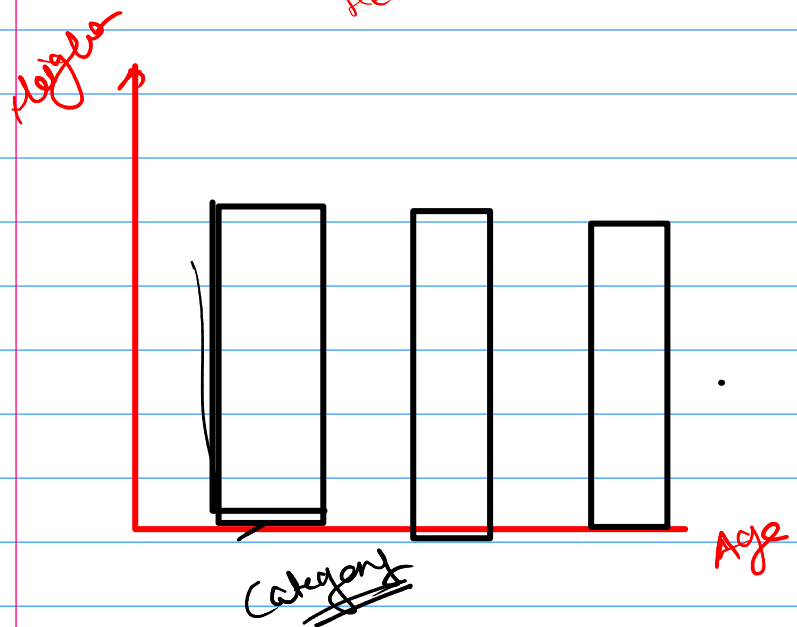


- Male 32 41 50
- Female 15 18 20

Scatter



Student Age Height
10 - 3
10 - 6
10 - 7
20 -



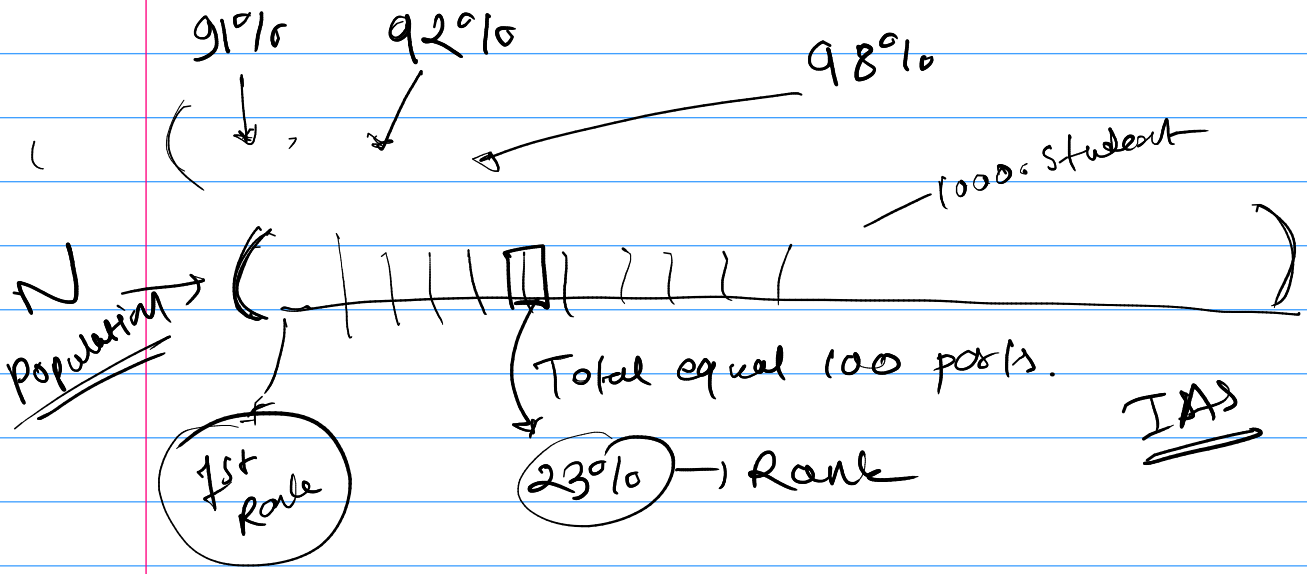
	x axis	Y axis
Bar	Category	freq. count
Pie	– Relative (%)	
Histo	Continuous	freq.
Scatt	num cont.	num cont.

Outlier Detection

5 - Number Summery

$$\text{Percentage} = \left(\frac{\text{Mark Obtain}}{\text{Total Mark}}\right) \times 100$$

BD . DM YCCE

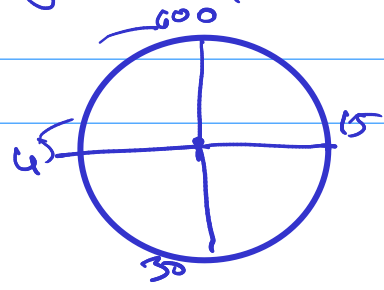


Quantile

→ Quantile are statistical measure use to divide a set of numerical data into equal-size group, each group contain equal no. of observation.

Quartile → Divide the Data into four equal part.

25%	25%	25%	25%



Quintile – Divide the data into five equal part

Decile :→ Divide the Data into Ten equal part

Percentile :– Divide data into 100 equal part.

Note :→ ① Data Should be sorted Low to high

② you are basically finding the Location or index position of observation.

③ They are not actual value in data.

Percentile :→ Divide the data into 100 equal parts.

for. ex – the 75th percentile is the value below which 75% of the observation in the dataset fall.

$$P_L = \frac{P}{100}(N+1)$$

P_L → Location P → percentile rank

N – no of observation.

1st 2nd 3 4 5 6 7 8 9 10

Data = 78, 82, 84, 88, 91, 93, 94, 96, 98, 99

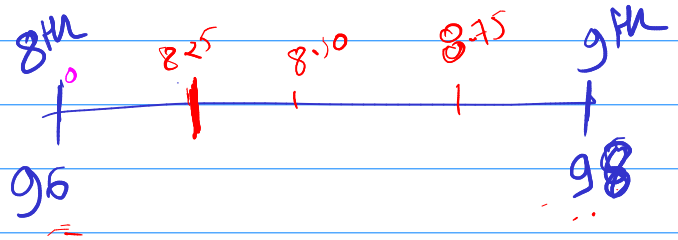
Q4- Find the 75th percentile Score.

N = 10 P = 75

$$PL = \frac{P}{100}(N+1)$$

$$PL = \frac{75}{100}(10+1) = \frac{75}{100}(11) = \frac{3}{4} \times 11 = \frac{33}{4}$$

$PL = 8.25$ — Location



$$\text{Rank} = 8.25 = 8^{th} + 0.25(9^{th} - 8^{th})$$

$$= 96 + 0.25(98-96)$$

$$= 96 + 0.25(2) = 96.0.50$$

75th percentile

Rank = 96.50

$$8.50 = 8 + 0.5(9^{th} - 8^{th})$$

$$8.75 = 8 + 0.75(9^{th} - 8^{th})$$

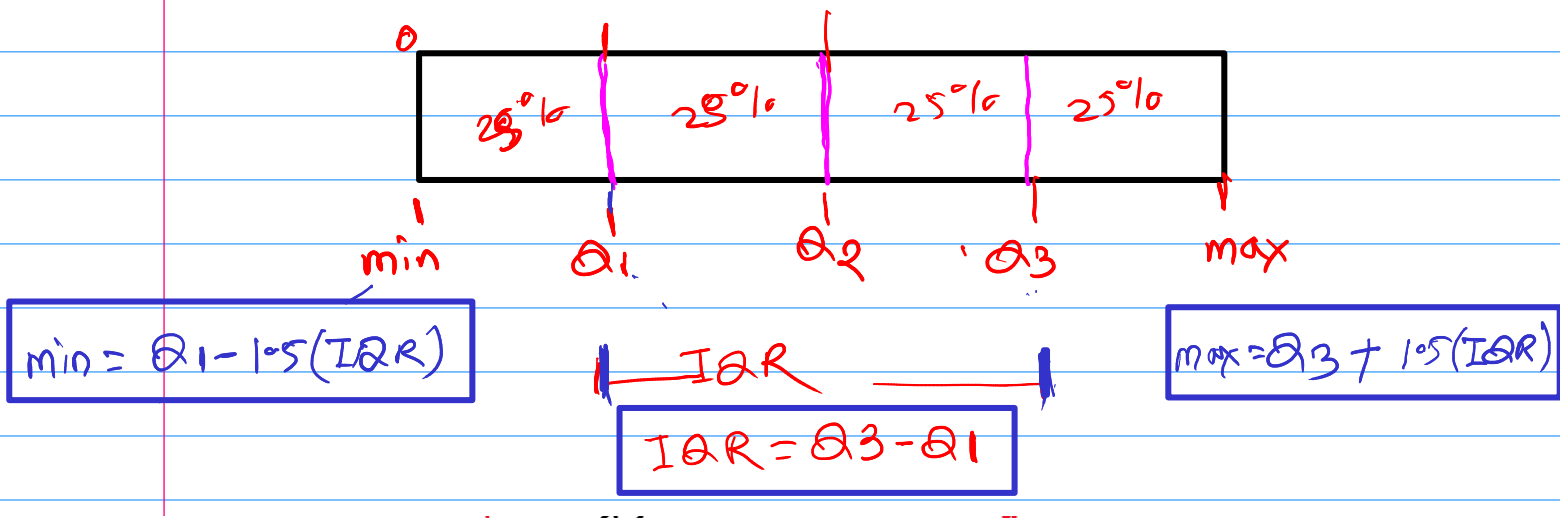
$$7.15 = 7^{th} + 0.15(8^{th} - 7^{th})$$

$$3.987 = 3^{rd} + 0.987(4^{th} - 3^{rd})$$

5-Number Summary

The five-number summary is a descriptive statistic that provides a summary of a dataset. It consists of five values that divide the dataset into four equal parts, also known as quartiles. The five-number summary includes the following values:

- 1-minimum Value - The smallest Value in dataset
- First Quartile (Q1) - The 25th percentile
- median / 2nd Quartile (Q2) = The 50th percentile
- Third Quatile (Q3) = The 75th percertile
- maximum Value = The max Value in data.



Interquartile Range] IQR

The interquartile range (IQR) is a measure of variability that is based on the five-number summary of a dataset. Specifically, the IQR is defined as the difference between the third quartile (Q3) and the first quartile (Q1) of a dataset.

Data

N = 10

x1 x2 x3 4 5 6 7 8 9 10

6, 213, 241, 260, 281, 290, 314, 321, 350, 1500

Q1, Q2, Q3, IQR, min, max ?