

Histogram And Skewness

A histogram is a graphical representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous variable and is used to visualize the shape, central tendency, and variability of a dataset.

Ages = { 11, 12, 14, 18, 24, 26, 30, 35, 36, 37, 40, 41, 42, 43, 50 } \Rightarrow Histogram

using this we will make :

0-50

① No. of bins = 10 $\rightarrow \frac{50}{10} = 5 \rightarrow$ bin size
(or no. of elements)

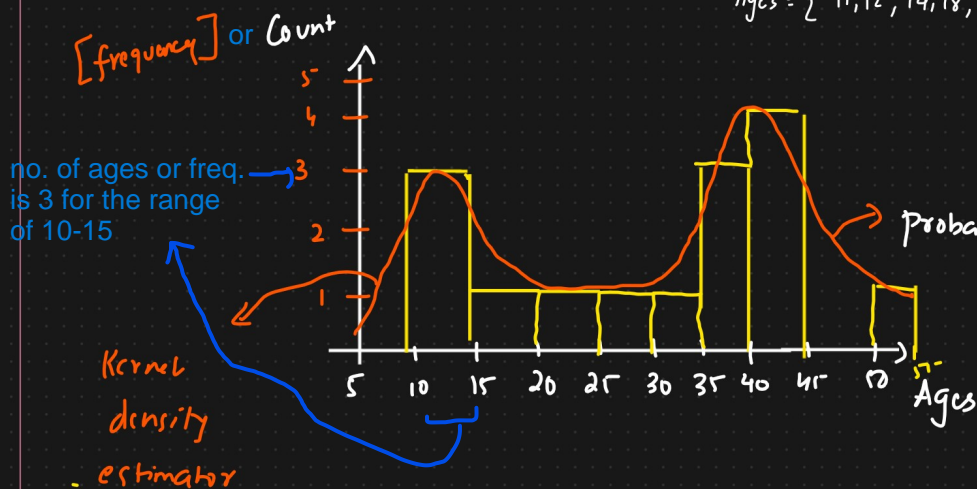
using bin size:

Bins \rightarrow [0-5, 5-10, 10-15, 15-20, 20-25, 25-30, 30-35, ... 45-50]

Ages = { 11, 12, 14, 18, 24, 26, 30, 35, 36, 37, 40, 41, 42, 43, 50 }

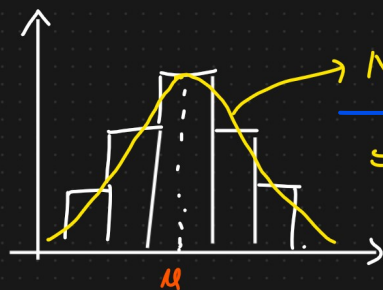
\downarrow shows the :

Distribution of Numerical DATA

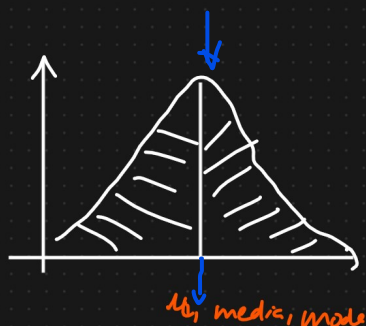


Ages = { - - - - - }

Skewness

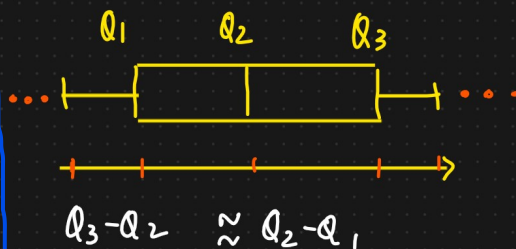


\Rightarrow Symmetrical Distribution \Rightarrow No Skewness



The mean, median and mode are all perfectly at the center.

Box plot

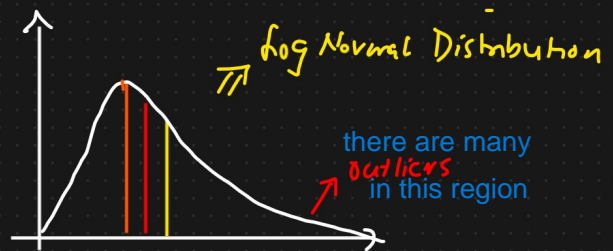


$$\boxed{\text{Mean} = \text{Median} = \text{Mode}}$$

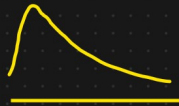
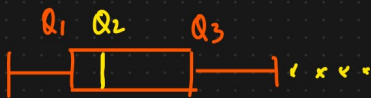
② Right Skewed



\Rightarrow Positive Skewed \Rightarrow



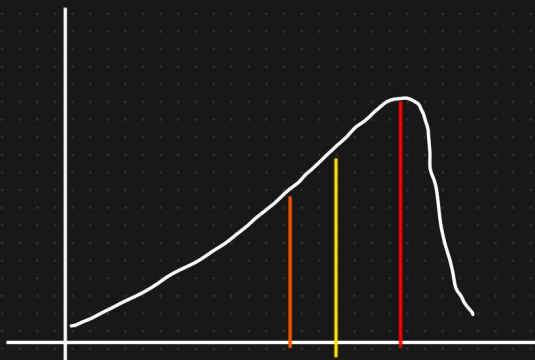
Box plot will be :



$$\boxed{\text{mean} > \text{median} > \text{mode}}$$

$$Q_3 - Q_2 > Q_2 - Q_1$$

③ Left Skewed Distribution



\Rightarrow Negative Skewed

Box plot



$$Q_2 - Q_1 > Q_3 - Q_2$$

Relationship

$$\text{mean} < \text{median} < \text{mode}$$