# Skewness

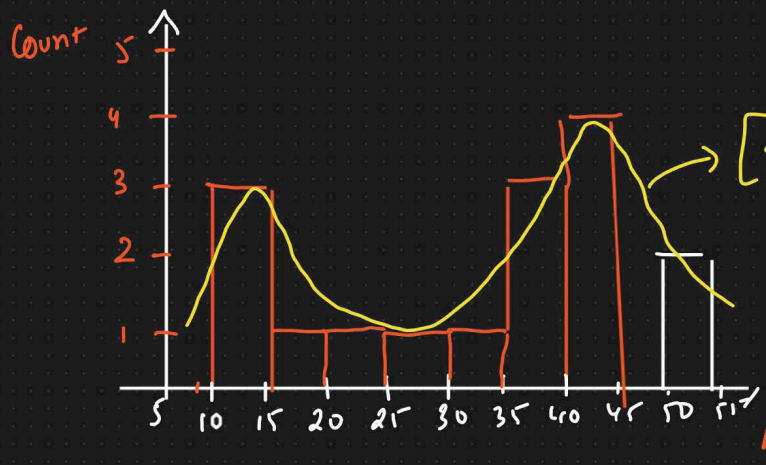
Skewness is a statistical measure that describes the asymmetry of the distribution of values in a dataset around its mean. It indicates whether the data points are concentrated more on one side of the mean than the other.

Histograms And Skewness → [Frequency]

Ages = {10, 12, 14, 18, 24, 26, 30, 35, 36, 37, 40, 41, 42, 48, 50, 51}

$\frac{50}{10} = 5$ → bin size  {No. of bins = 10}

$\frac{50}{20} = 2.5$ → bin size  {No. of bins = 20}




→ [probability distribution function]

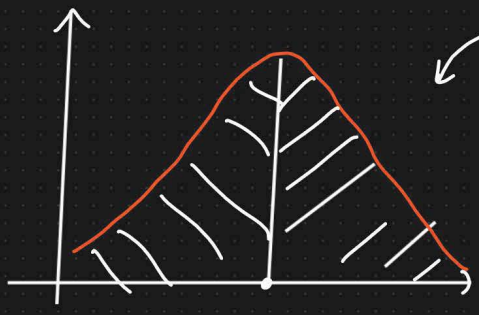
↓

Pdf = {probability density function}

Skewness

→ Normal / Gaussian Distribution

Symmetrical Distribution



① No skewness

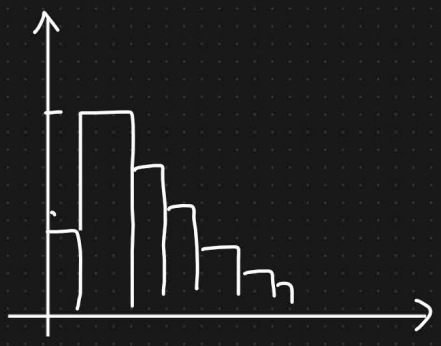
Box plot


$Q_3 - Q_2 \approx Q_2 - Q_1$

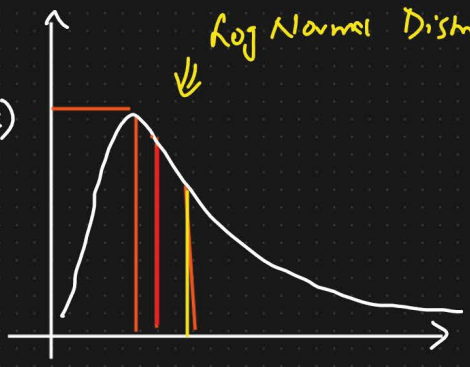
The mean, median, and mode all are perfectly at the center

mean = median = mode

## ② Right Skewed

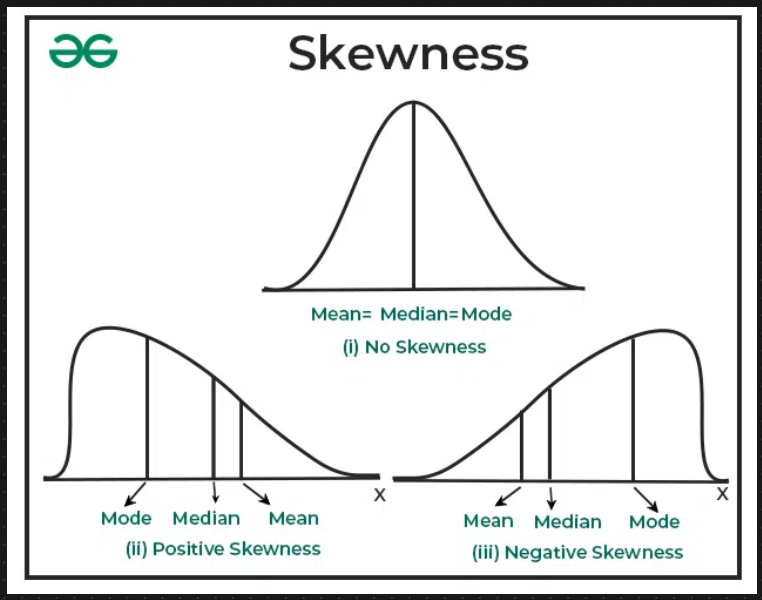


⇒ Positive Skewed ⇒




mean ≥ median ≥ mode

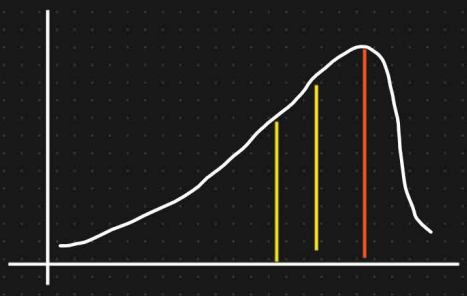
Relationship between Mean, Median, Mode



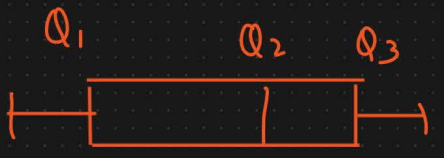

Box plot


$Q_3 - Q_2 \geq Q_2 - Q_1$

## ③ Left Skewed Distribution



⇒ Negative Skewed




$Q_2 - Q_1 \geq Q_3 - Q_2$

Relationship : mean ≤ median ≤ mode