Descriptive Statistics

Basic Metrics :

#Salary data : 30, 30, 35, 40, 40, 300 $n=6$
(in LPA)

- **Mean** $= \left(\frac{30+30+35+40+40+300}{6}\right) = 79$

- **Median** $= \left(\frac{35+40}{2}\right) = 37.5$

10, 20, 30, (40), 50, 60, 70
↓
Central Value is (Median)

90, 90, 90, 80, 90, 75, 95 → Mode = 90

- **Mode** : most frequently occuring numer

20, 25, 60, 100

- Range = (max − min) = (100 − 20) = 80

data:
- array of numbers
- To find median → sort the data first.

Sorted data: 66, 67, 67, 68, 68, 68, 68, 69, 69, 69, 69, 70, 70, 71, 71, 72, 73, 73

- Median = 50th percentile (P_{50})

median

50% of data less than median value

66, 67, 67, 68, 68, 68, 68, 69, 69, 69, 69, 70, 70, 71, 71, 72, 73, 73

median of first half (25% data less than this) → P_{25} → Q_1

median → P_{50} → Q_2

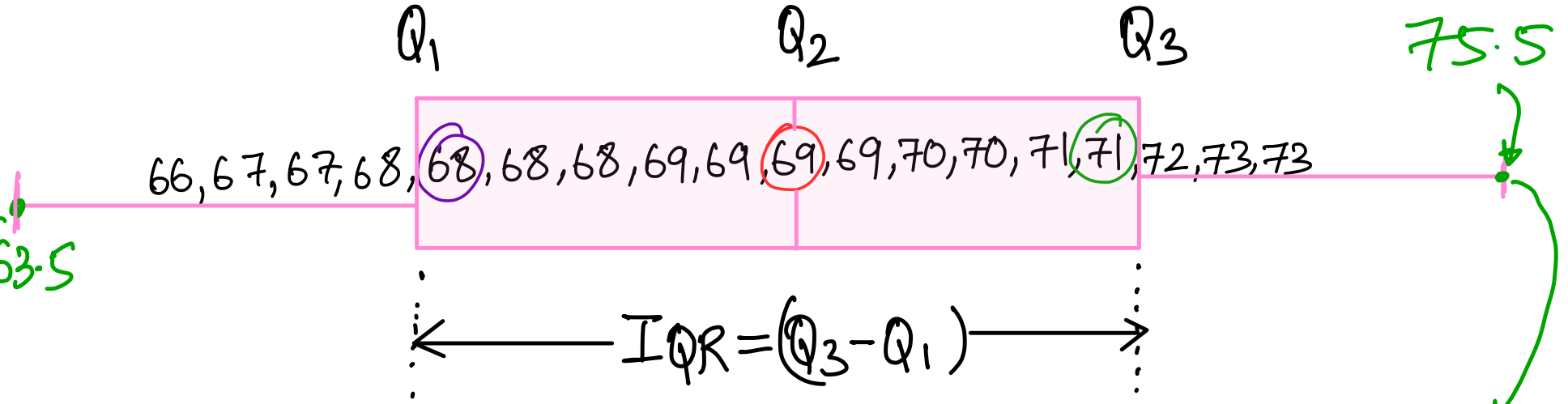
median of second half (75% of data less than this) → P_{75} → Q_3

Percentiles → P_{25}, P_{50}, P_{75}

Quartiles → Q_1, Q_2, Q_3

Inter Quartile Range (IQR) → $Q_3 - Q_1 = (71 - 68) = 3$

Box-Plot



71 + 1.5 x 3

lower end of whisker

$Q_1 - 1.5 * IQR$

upper end of whisker

$Q_3 + 1.5 * IQR$

$[Q_1 - 1.5 \times IQR, Q_3 + 1.5 \times IQR]$

values outside this range ≡ Potential Outliers

Variance & standard deviation

– measurement of spread around the mean value

- (data)

x_1	110
x_2	105
x_3	95
x_4	90

Mean $(\bar{x}) = 100$

- **Variance**: Average squared distance from mean value

$$\text{Variance} = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2+(x_4-\bar{x})^2}{4}$$

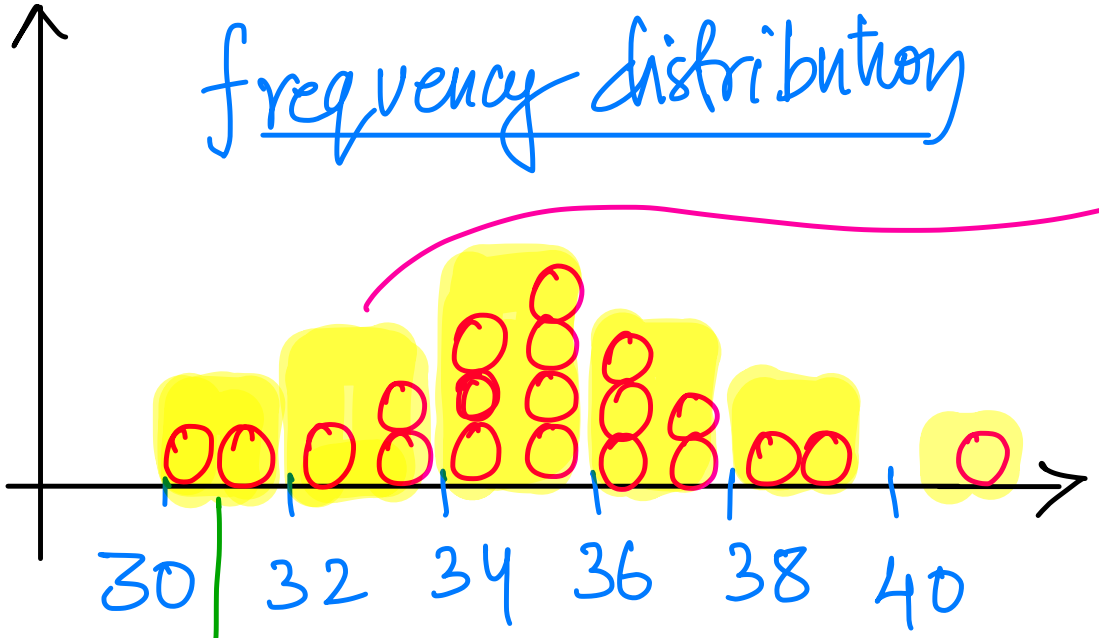
$$= \frac{(110-100)^2+(105-100)^2+(95-100)^2+(90-100)^2}{4}$$

$$\text{Variance} = \frac{250}{4} = \sigma^2$$

- **Std. deviation** $= \sqrt{\text{Variance}} \Rightarrow \sigma = \sqrt{\frac{250}{4}}$

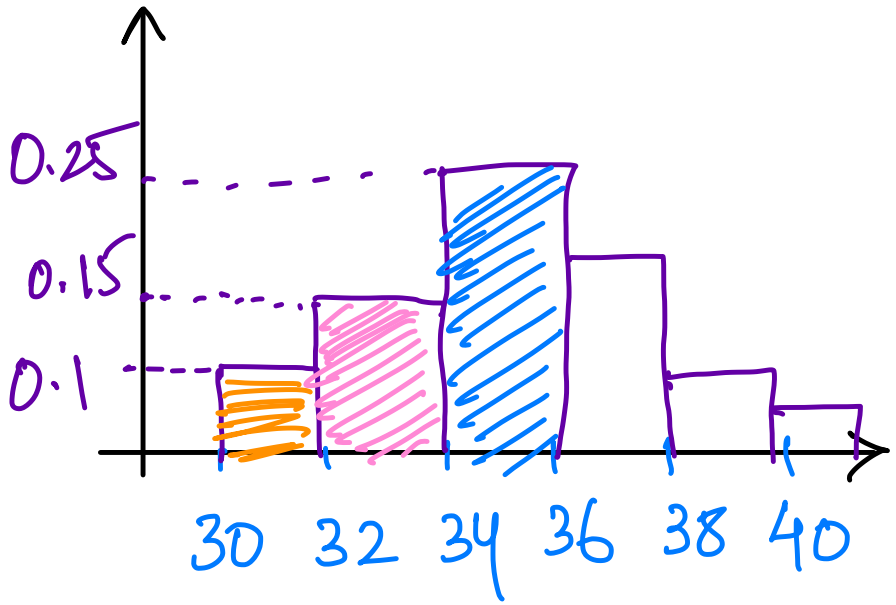
Histogram:

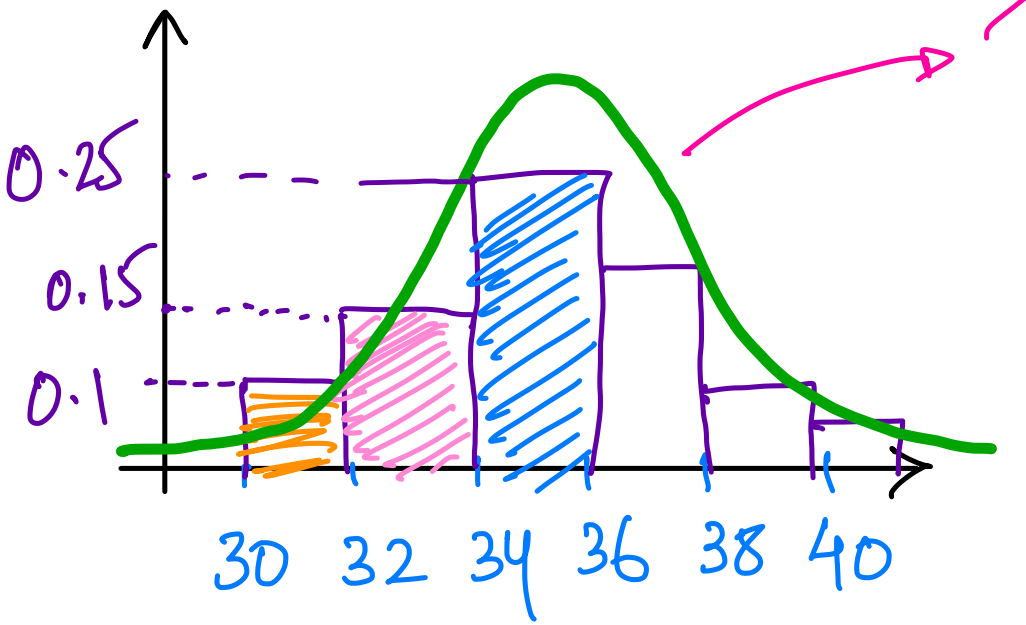
32, 36, 33, 34, - - - (n=20)



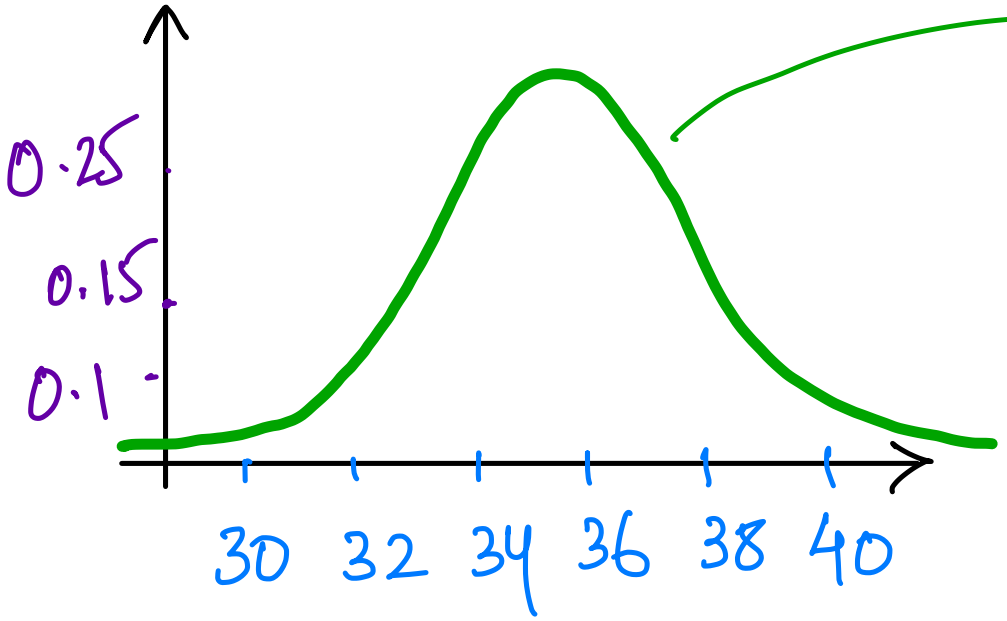
what fraction of data is in this bar $\Rightarrow \frac{2}{20} = 0.1$

fraction of data [32,34) $= \frac{3}{20} = 0.15$





The curve that estimates this trend is called probability density function (PDF)

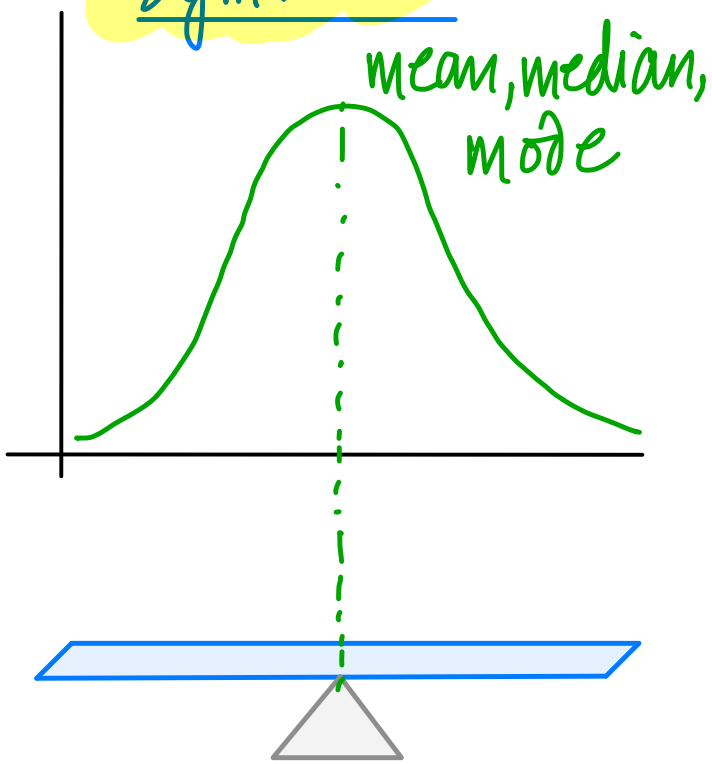


PDF

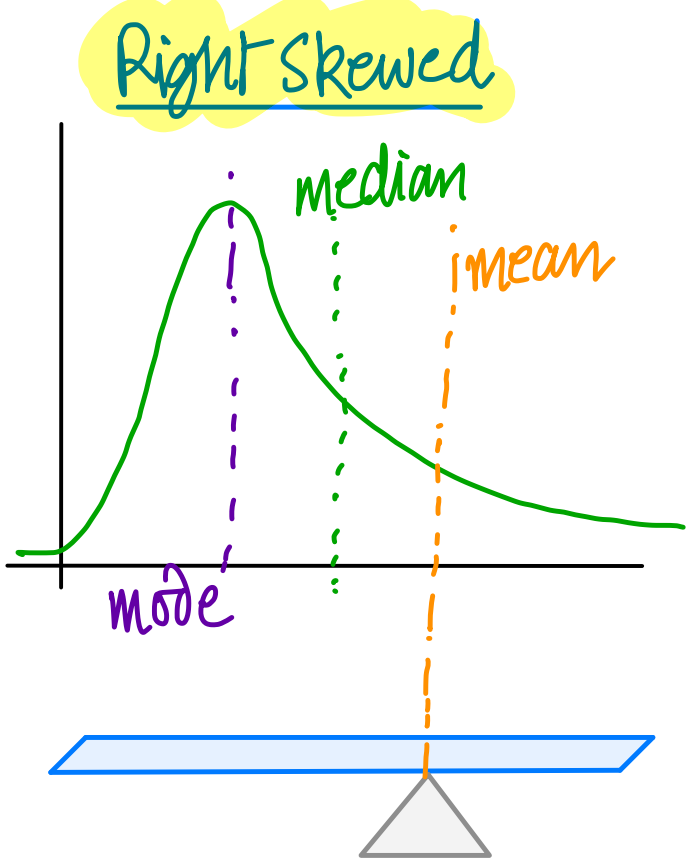
↓

A mathematical function

Symmetric

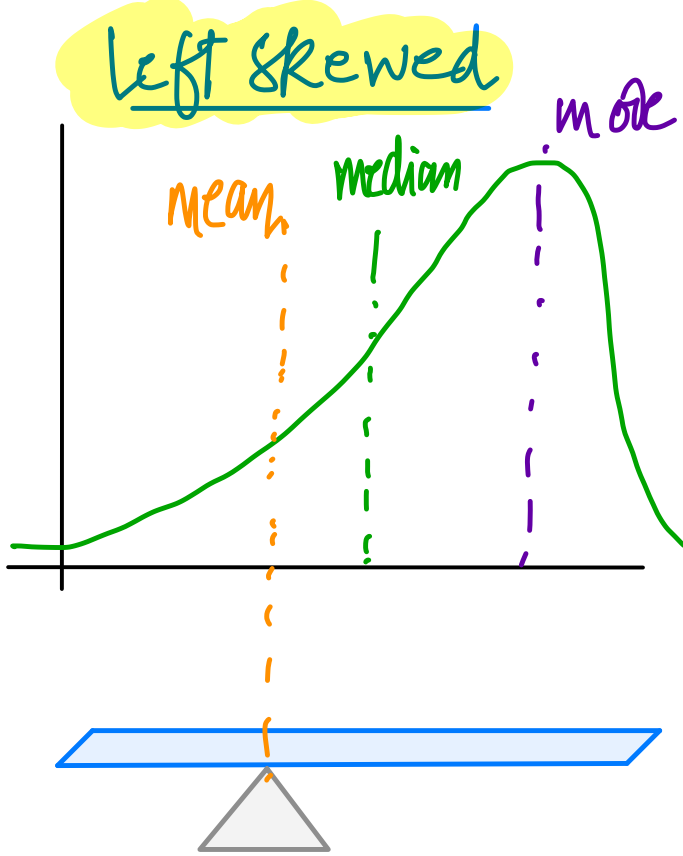


Right skewed



+ve skewness

Left skewed



−ve skewness