Recall

Arithmetic mean — Is a measure to represent the entire data.
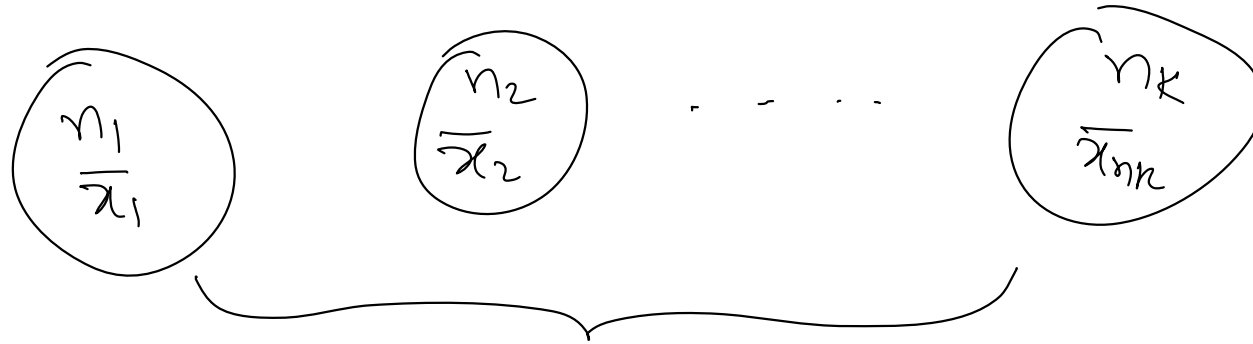
Individual observation: $\bar{x} = \dfrac{\sum_{1}^{n} x_i}{n} \longrightarrow$ no. of observations.

$\longrightarrow$ frequency.

Discrete observation: $\bar{x} = \dfrac{\sum_{1}^{n} f x}{\sum f}$

Continuous data $\bar{x} = A + \dfrac{\sum_{1}^{n} f d}{\sum f} \, x_i$

$d = \dfrac{m - A}{i}$

$\nearrow$ mid pt of CI

$i \longrightarrow$ length of the CI

$\downarrow$ assumed mean.

Combined Arithmetic mean:-

Circles containing:
$$\frac{n_1}{\bar{x}_1} \qquad \frac{n_2}{\bar{x}_2} \qquad \cdots \qquad \frac{n_k}{\bar{x}_{nk}}$$

$$\bar{X} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + \cdots + n_k \bar{x}_k}{n_1 + n_2 + \cdots + n_k}.$$

$\rightarrow$ The average salary of male employees in a firm was Rs.5200 and that of females was Rs.4200. The mean salary of all employees was Rs.5000. Find the Percentage of male and female employees.

Soln

Avg salary of male employees $\overline{x}_1 = 5200$; $\overline{x}_2 = 4200$ $\rightarrow$ Avg salary of female employees

$\overline{x} = 5000.$

$n_1 \rightarrow$ no of male employees

$n_2 \rightarrow$ no of female employees

$$\overline{x} = \frac{n_1 \overline{x}_1 + n_2 \overline{x}_2}{n_1 + n_2}$$

% of male & female employees.

| Male 5200 | female 4200 |

5000.
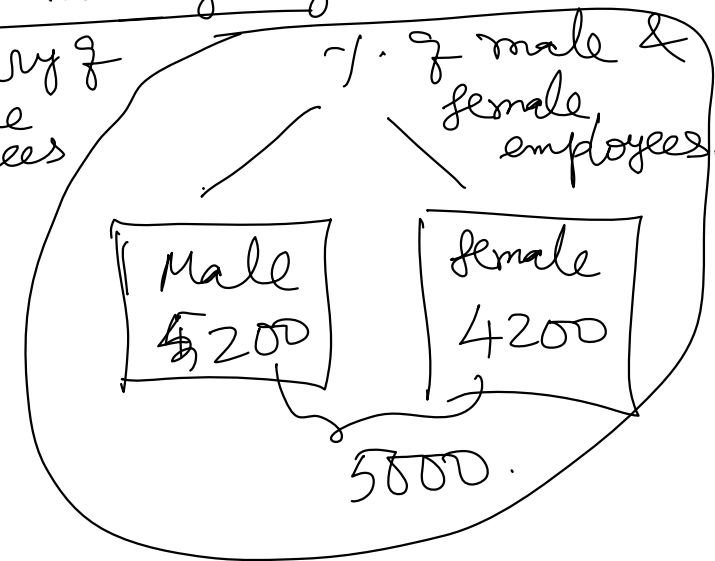
$$5000 = \frac{n_1(5200) + n_2(4200)}{n_1 + n_2}$$

$$5000n_1 + 5000n_2 = 5200n_1 + 4200n_2$$

$$5000n_2 - 4200n_2 = 5200n_1 - 5000n_1$$

$$800n_2 = 200n_1$$

$$\frac{n_1}{n_2} = \frac{800}{200} = \frac{4}{1}.$$
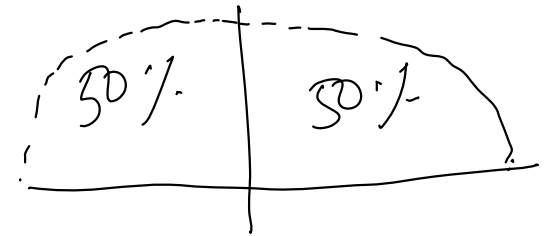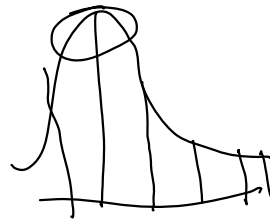
% of male employees is $= \frac{4}{5} \times 100 = 80$ //

% of female employees is $= \frac{1}{5} \times 100 = 20$. //

Disadvantages of A.M.

1) It is affected by the extreme values.

2) It cannot be determined by inspection nor it can be located graphically.

3) A.M cannot be calculated if any observation is missing.

4) In extremely asymmetrical distribution, A.M is not a suitable measure of location.

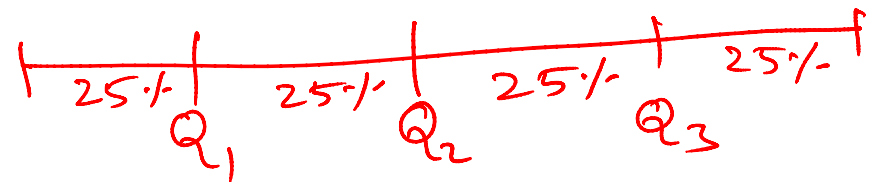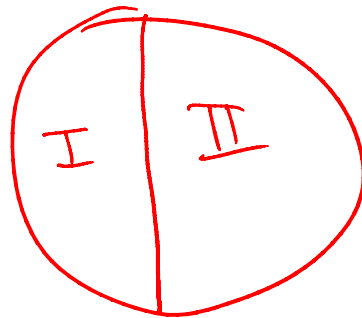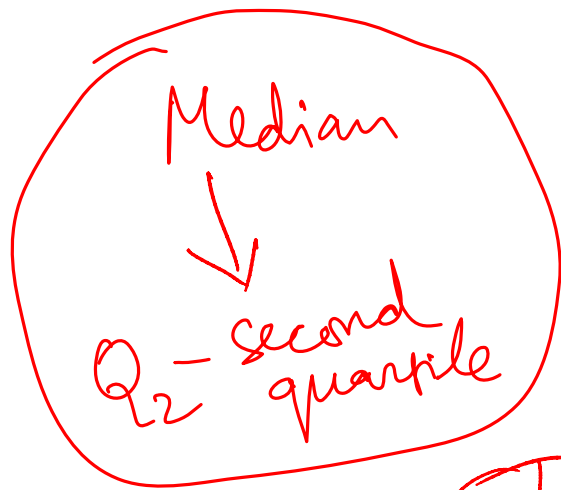50% | 50%

Median $\rightarrow$ Middle value of the data.

Individual observations:- Size of $\left(\dfrac{N+1}{2}\right)^{th}$ item $\rightarrow$ no. of observation

Discrete data :- Size of $\left(\dfrac{N+1}{2}\right)^{th}$ item $\rightarrow N = \sum f$

Continuous data :- Median $= L + \left[\dfrac{\dfrac{N}{2} - cf}{f}\right] \times i$

$\rightarrow$ length of the C.I.

Lower limit Median class

$\sum f$

frequency of Median class

$\{$ Cummulative frequency preceding the median class

Median

$\downarrow$

Q_2 – second quartile

I | II

25% | 25% | 25% | 25%
Q_1 ... Q_2 ... Q_3

Percentiles

Deciles

Quartiles
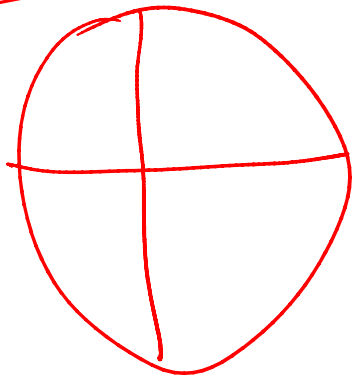
I Quartile $Q_1 = L + \dfrac{\frac{N}{4} - Cf}{f} \times i$

III Quartile $Q_3 = L + \dfrac{\frac{3N}{4} - Cf}{f} \times i$

Find the median of $\underline{20}$.

Find the median of $10, \boxed{20, 30,} 40.$

$\underbrace{\qquad}$ → 25

$\left(\dfrac{1}{2} = 0.5\right)$

$\rightarrow \boxed{\dfrac{N+1}{2}}$

$\dfrac{N}{2}$

Advantage → Unaffected by the extreme values.
(outliers)

$\longrightarrow$ Find the median for
40, 25, 60, 35, 44 & 50.

Soln

Arrange in ascending order.

25, 35, $\boxed{40, \; 44}$, 50, 60.
                    42

$N = 6$.

Median = Size of $\left(\dfrac{N+1}{2}\right)^{th}$ item = Size of 3.5^{th} item

= Avg of 3rd position & 4th position

= $\dfrac{40+44}{2}$ = 42 ✓

Find the median of 30, 60, 20, 15 & 70.

Soln   Arranging in ascending order

15, 20, 30, 60, 70

$N = 5$.

Median is = Size of $\left(\dfrac{N+1}{2}\right)^{th}$ item = Size of 3rd item.

Median = 30

→ obtain median for the following

x: 10 20 30 40 50 60 70 80

f: 3 15 20 6 10 8 7 11

Soln   Table is formed (as follows)

x	f	Cummulative frequency
10	3	3
20	15	18
30	20	38
40	6	44
50	10	54
60	8	62
70	7	69
80	11	80

→ Q_1 class

$N = \sum f = 80$

$$Q_1 = \text{Size of } \left(\frac{N+1}{4}\right)^{th} \text{item}$$

$$= \text{Size of } \left(\frac{81}{4}\right)^{th} \text{item}$$

↓ 20.25

$$\text{Median} = \text{Size} \left(\frac{N+1}{2}\right)^{th} \text{item}$$

$$= \text{Size} \left(\frac{80+1}{2}\right)^{th} \text{item}$$

$$= \text{Size } 40.5^{th} \text{item}$$

$$\boxed{\text{Median} = 40}$$

$$\boxed{Q_1 = 30}$$

→ Find the median

CI:	10-20	20-30	30-40	40-50	50-60
f:	20	90	150	100	70

Soln   form the table

CI	f	Cf
10-20	20	20
20-30	90	110
30-40	150	260
40-50	100	360
50-60	70	430
	430	

Median class → 30-40

$$\frac{N}{2} = \frac{430}{2} = 215$$

$$L = 30 \; ; \; f = 150 \; ; \; Cf = 110.$$

$$\text{Median} = L + \frac{\frac{N}{2} - Cf}{f} \times i$$

$$= 30 + \frac{215 - 110}{150} \times 10$$

$$= 30 + \frac{105}{15} = 37.$$

→ Given median is 46. Determine the missing frequencies.

CI:	10–20	20–30	30–40	40–50	50–60	60–70	70–80	
f:	12	30	?	65	?	25	18	(229)

Soln:-

CI	f	Cf
10–20	12	12
20–30	30	42
30–40	x	42+x
40–50	65	107+x
50–60	y	107+x+y
60–70	25	132+x+y
70–80	18	150+x+y
	229	

$150 + x + y = 229$

$\boxed{x + y = 79}$

$L = 40; \quad \dfrac{N}{2} = \dfrac{229}{2} = 114.5$

$f = 65; \quad Cf = 42+x; \quad i = 10.$

$\text{Median} = L + \dfrac{\frac{N}{2} - Cf}{f} \times i$

$46 = 40 + \dfrac{114.5 - 42 - x}{65} \times 10$

$$6 = \frac{76.5 - x}{65} \times 10$$

$$\frac{6 \times 65}{10} = 76.5 - x$$

$$x = 76.5 - \frac{390}{10}$$

$$x = 76.5 - 39$$

$$x = 37.5$$

$$x + y = 79$$

$$y = 79 - x$$

$$= 79 - 37.5$$

$$y = \underline{41.5}$$