Mode

The **mode** is the value that is repeatedly occurring in a given set. We can also say that the value or number in a data set, which has a high frequency or appears more frequently, is called mode or **modal value**.

Individual series /Mode Formula For Ungrouped Data

96 101 99 100 98 103 97 99 102 95

Mode - 99

Discrete Series / Mode Formula For Grouped Data

There is a possibility that more than one observation has the same frequency, i.e. a data set could have more than one mode. In such a case, the set of data is said to be multimodal.

The frequency table of the given data is as given below:

Scores	5	$\boxed{6}$	7	8	9
Frequency	5	7	6	4	4

We observe that the value 6 has the maximum frequency i.e., it occurs a maximum number of times. Therefore, the mode of the given data is 6.

Continuous Series / Mode Formula For Grouped Data

Height (in cm)	125-130	130-135	135-140	140-145	145-150
Number of students	7	14	10	10	9

$$\text{Mode} = l + \left[\frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)}\right] h$$

Here, the maximum frequency is 14 and the corresponding class is 130-135.

Where
l = lower limit of the modal class

fm = frequency of the modal class

f1= frequency of class preceding the modal class

f2 = frequency of class succeeding the modal class

h = width of the modal class

$$\begin{aligned} \therefore \quad \text{Mode} &= l + \left[\frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)}\right] \times h \\ &= 130 + \left[\frac{14-7}{(14-7)+(14-10)}\right] \times 5 \\ &= 130 + \left[\frac{7}{7+4}\right] \times 5 = 130 + 3.18 \\ &= 133.18 \end{aligned}$$

Studying time (in minutes)	Frequency
0.5 - 10.5	2
10.5 - 20.5	10
20.5 - 30.5	6
30.5 - 40.5	4
40.5 - 50.5	3

Test yourself ?

$$\text{Mode} = l + \left[\frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)}\right] h$$

Measures of Dispersion

Measures of dispersion help to describe the variability in data. Dispersion is a statistical term that can be used to describe the extent to which data is scattered. Thus, measures of dispersion are certain types of measures that are used to quantify the dispersion of data.

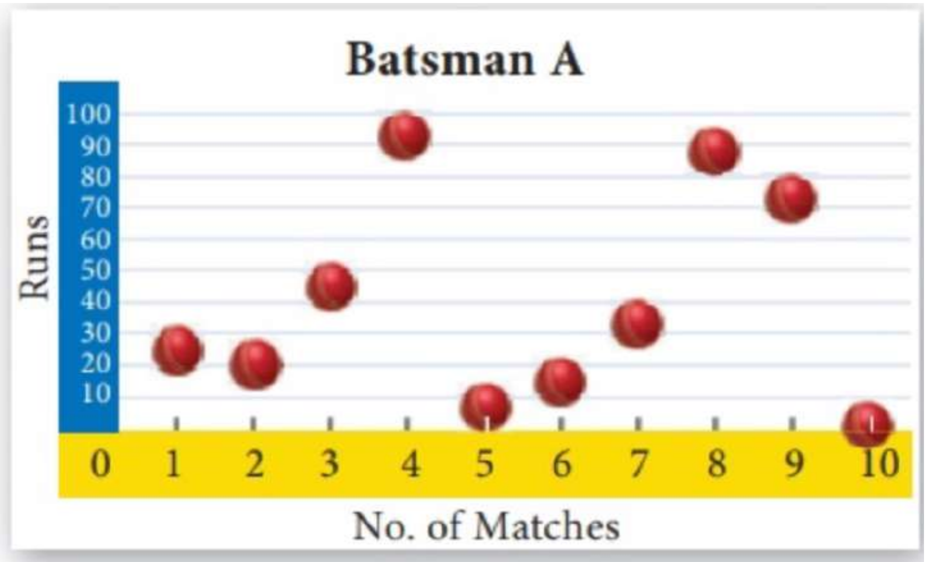
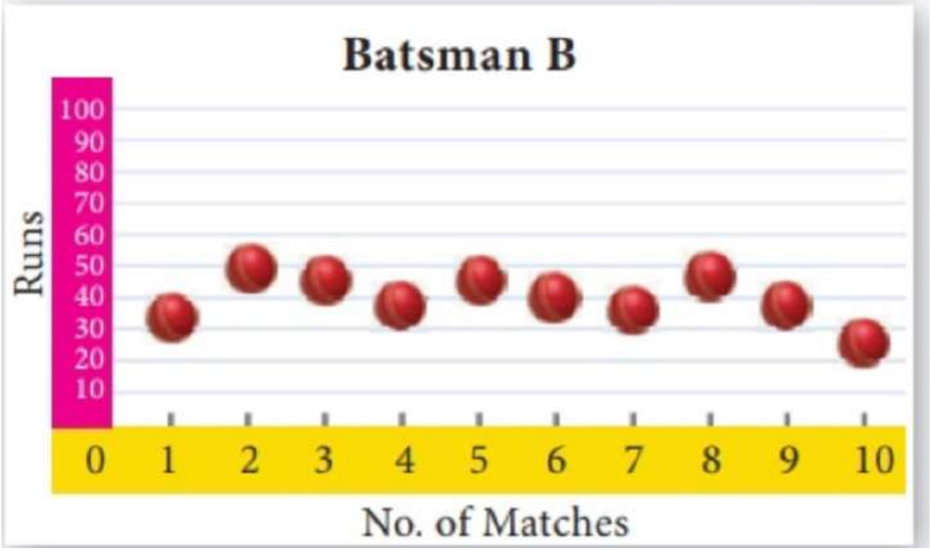


Fig. 8.1(a)



GIVEN:
The mean of both data sets are same (40)

Real Life working fields :

In Weather Forecasting.

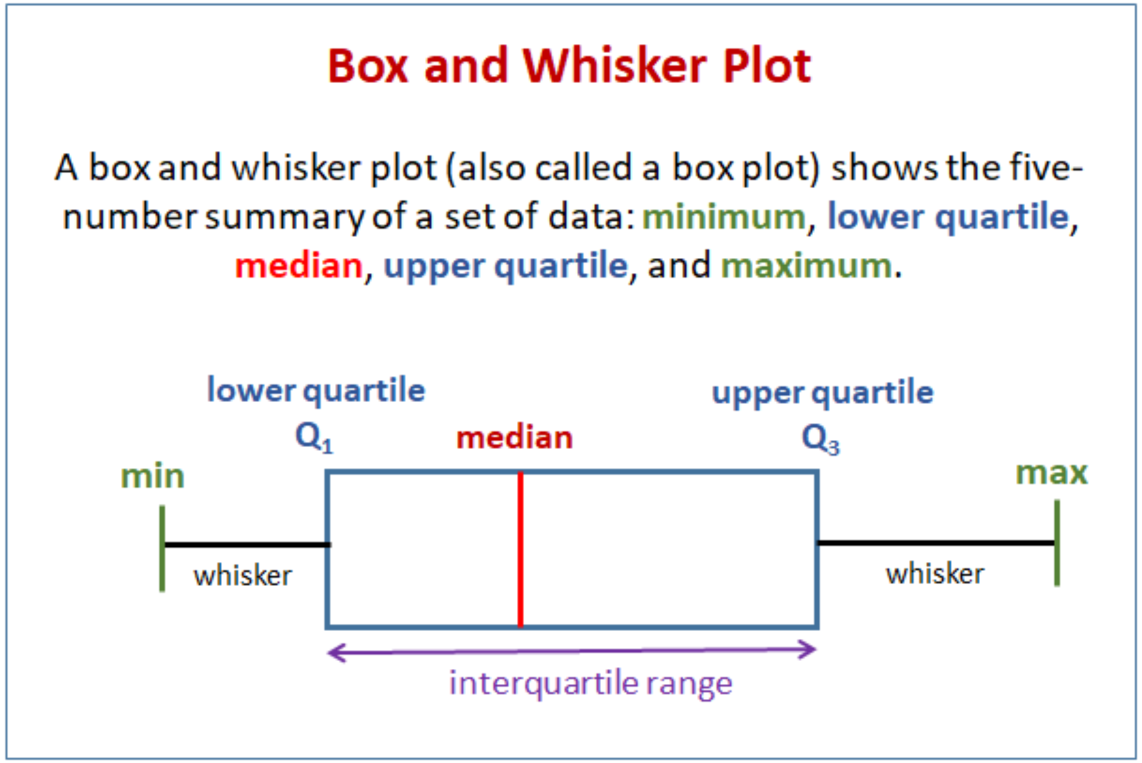
In Healthcare.

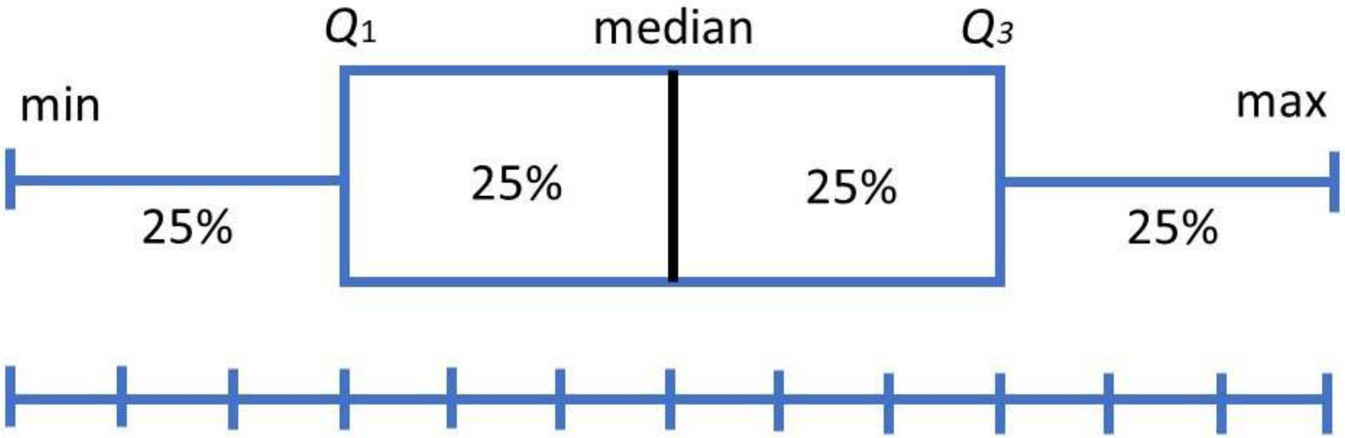
In Real Estate.

In Human Resources.

Different Measures of Dispersion are

1. Range
2. Mean deviation
3. Quartile deviation
4. Standard deviation
5. Variance
6. Coefficient of Variation





Individual series /Range

The difference between the largest value and the smallest value is called Range.

Range $R = L - S$

Coefficient of range = $(L - S) / (L + S)$

where L - Largest value; S - Smallest value

Find the range and coefficient of range of the following data: 25, 67, 48, 53, 18, 39, 44.

Solution Largest value $L = 67$; Smallest value $S = 18$

Range $R = L - S = 67 - 18 = 49$

Coefficient of range $= (L - S) / (L + S)$

Coefficient of range = $(67 - 18) / (67 + 18) = 49/85$

$= 0.576$

Note : Coefficient is scaled between the range, -1 and +1.

Discrete Series/Range

Marks obtained :	$\boxed{4}$	8	12	16	$\boxed{20}$
No. of students :	6	12	18	15	9

Range R = $L - S$ = 20 − 4 = 16

Coefficient of range = $(L - S) / (L + S)$

(20-4)/(20+4) =0.66

Coefficient of range = $(L - S) / (L + S)$

continuous series / Range

Age (in years)	16-18	18-20	20-22	22-24	24-26	26-28
Number of students	0	4	6	8	2	2
Mid value	17	19	21	23	25	27

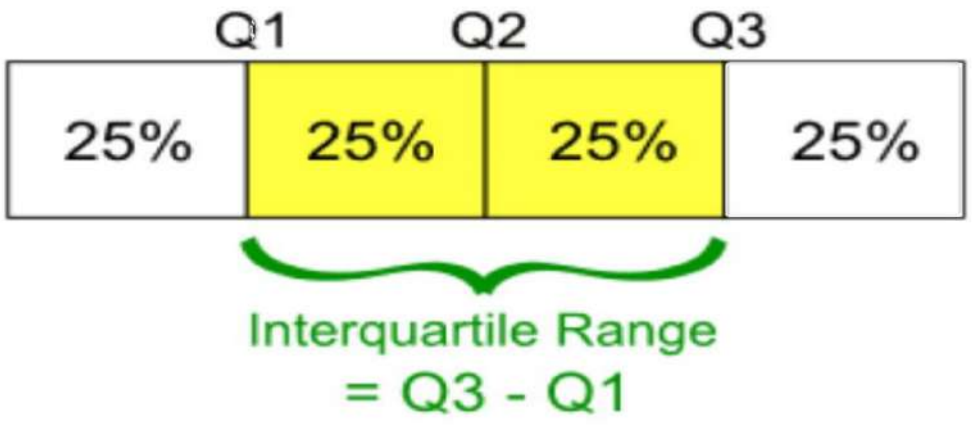
Range = L-S

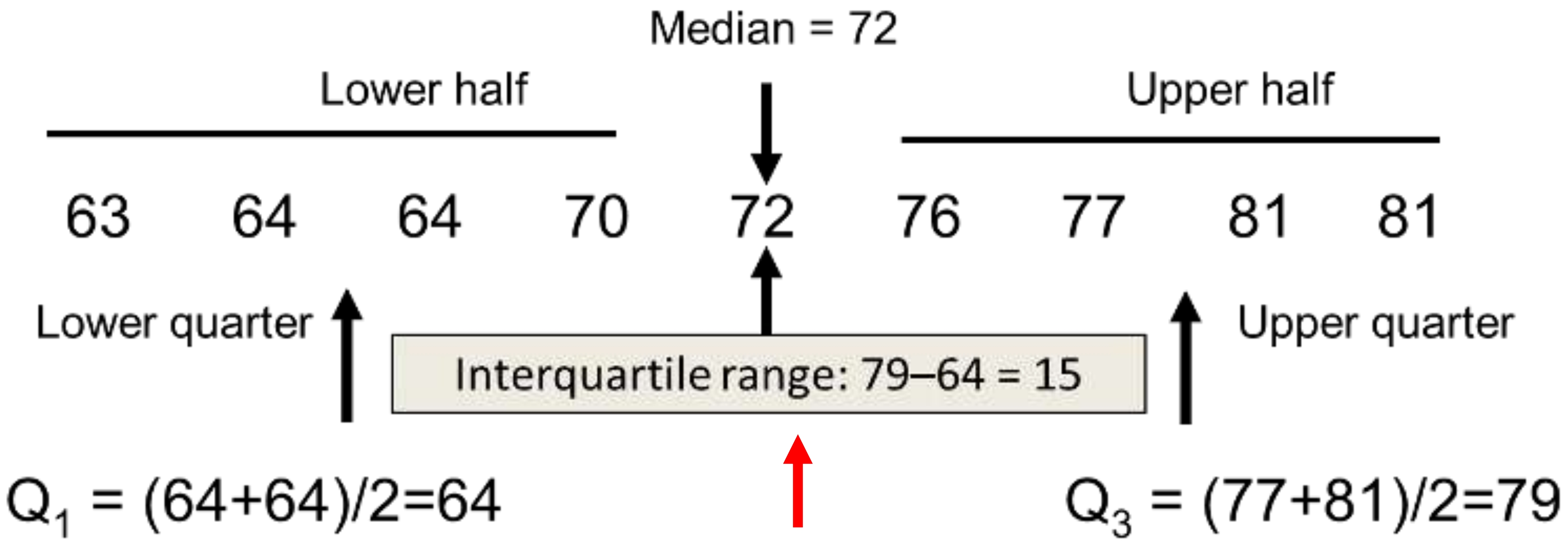
R= 27-17

R= 10 years

Coefficient of range = $(L - S) / (L + S)$

Inter-quartile range





Mean deviation - difference between the mean and each value.

Data : 6, 7, 10, 12, 13, 4, 8, 12.

Mean of the given data= $\frac{Sum of all the terms}{total number of terms}$

$\bar{X}=\frac{6+7+10+12+13+4+8+12}{8}$

$=\frac{72}{8}$

=9

$$\textbf{Mean Deviation} = \frac{\Sigma\,|X-\bar{X}|}{N}$$

where,

X = denotes each value in the data set

X̄ = denotes the mean value of the data set

N = total number of data values

| | = represents absolute value, i.e. it ignores the sign

x_i	$x_i - \bar{x}$	$\lvert x_i - \bar{x}\rvert$
6	6-9=-3	l-3l=3
7	7-9=-2	l-2l=2
10	10-9=1	l1l=1
12	12-9=3	l3l=3
13	13-9=4	l4l=4
4	4-8=-5	l-5l=5
8	8-9=-1	l-1l=1
12	12-9=3	l3l=3
N=8		$\sum_{1}^{8}\lvert x_i - \bar{x}\rvert$=22

Mean deviation about mean= $\frac{\sum\,|\,X_i-\bar{X}|}{8}$

$\frac{22}{8}$

=2.75

variance

Data set					
46	69	32	60	52	41

Step 1: Find the mean

To find the mean, add up all the scores, then divide them by the number of scores.

Mean ($\bar{x}$)
$\bar{x}$ = (46 + 69 + 32 + 60 + 52 + 41) ÷ 6 = **50**

Step 2: Find each score's deviation from the mean

Subtract the mean from each score to get the deviations from the mean.

Since $\bar{x}$ = 50, take away 50 from each score.

Score	Deviation from the mean
46	46 – 50 = **-4**
69	69 – 50 = **19**
32	32 – 50 = **-18**
60	60 – 50 = **10**
52	52 – 50 = **2**
41	41 – 50 = **-9**

Step 3: Square each deviation from the mean

Multiply each deviation from the mean by itself. This will result in positive numbers.

Squared deviations from the mean
$(-4)^2$ = 4 × 4 = **16**
19^2 = 19 × 19 = **361**
$(-18)^2$ = -18 × -18 = **324**
10^2 = 10 × 10 = **100**
2^2 = 2 × 2 = **4**
$(-9)^2$ = -9 × -9 = **81**

Step 4: Find the sum of squares

Add up all of the squared deviations. This is called the sum of squares.

Sum of squares
16 + 361 + 324 + 100 + 4 + 81 = **886**

variance =Sum of squares deviation/n-1 for sample

Step 5: Divide the sum of squares by $n – 1$ or N

Divide the sum of the squares by $n – 1$ (for a sample) or N (for a population).

Since we're working with a sample, we'll use $n – 1$, where n = 6.

Variance
886 ÷ (6 – 1) = 886 ÷ 5 = **177.2**

variance =Sum of squares deviation/n for population

Standard Deviation

DATA SET: 4, 2, 5, 8, 6

Mean:

$\bar{x} = \frac{\sum x}{n}$

=

$\frac{x_1+x_2+x_3+x_4.\ldots.+x_n}{n}$

= (4 + 2 + 5 + 6 + 8) / 5

= 5

$x_n - \bar{x}$

for every value of the sample:

$x_1 - \bar{x} = 4–5 = -1$

$x_2 - \bar{x} = 2–5 = -3$

$x_3 - \bar{x} = 5–5 = 0$

$x_4 - \bar{x} = 8–5 = 3$

$x_5 - \bar{x} = 6–5 = 1$

$\sum (x_n - \bar{x})^2$

$S.D = \sqrt{\frac{\sum (x_n-\bar{x})^2}{n-1}}$

=

$\sqrt{\frac{20}{4}}$

= √5

= 2.236

$x_n - \bar{x}$

for every value of the sample:

$x_1 - \bar{x} = 4–5 = -1$

$x_2 - \bar{x} = 2–5 = -3$

$x_3 - \bar{x} = 5–5 = 0$

$x_4 - \bar{x} = 8–5 = 3$

$x_5 - \bar{x} = 6–5 = 1$

$\sum (x_n - \bar{x})^2$

$= (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_5 - \bar{x})^2$

$= (-1)^2 + (-3)^2 + 0^2 + 3^2 + 1^2$

= 20

Coefficient of Variation

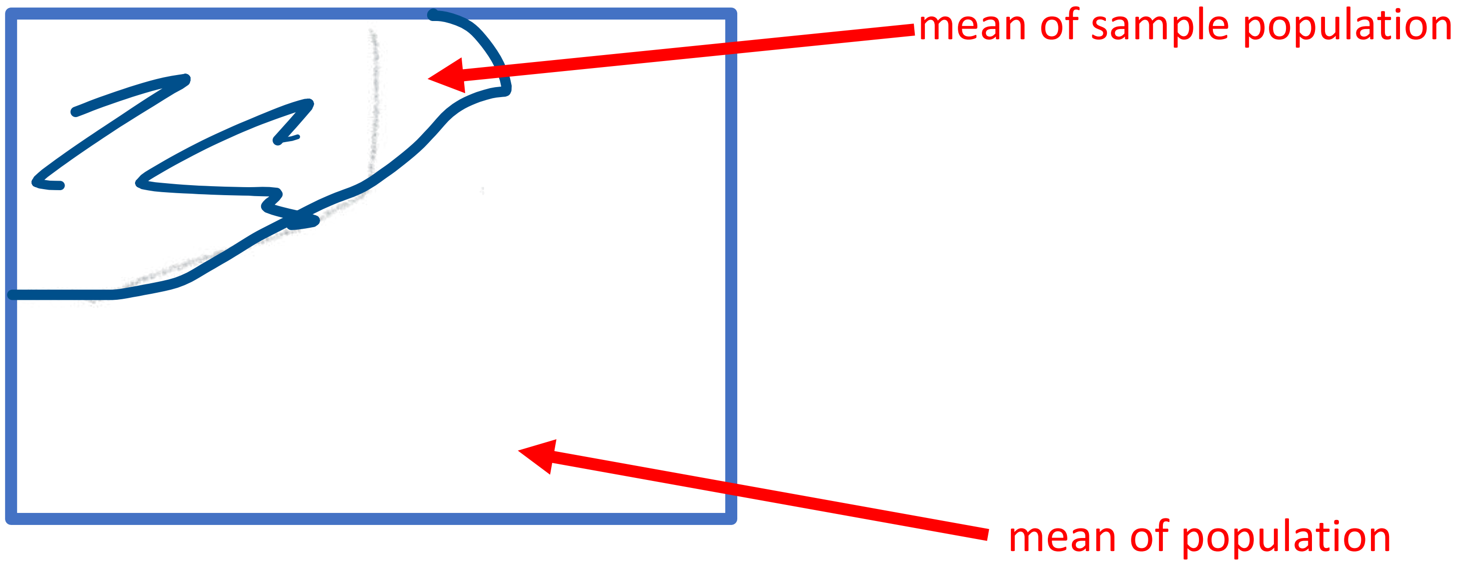
	Coefficient of Variation	Standard Deviation
Population	$\frac{\sigma}{\mu} \times 100$	$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N}}$
Sample	$\frac{s}{\mu} \times 100$	$s = \sqrt{\frac{\sum(x_i - \mu)^2}{N - 1}}$

Coefficient of Variation = (Standard Deviation / Mean) * 100

Standard Error

The standard error (SE) of a statistic is the approximate standard deviation of a statistical sample population. The standard error is a statistical term that measures the accuracy with which a sample distribution represents a population by using standard deviation.

SE = Difference between mean of sample population from mean of population



Age (in years)	16-18	18-20	20-22	22-24	24-26	26-28
Number of students	0	4	6	8	2	2

How to sort unequal intervals !

Age (in years)	16-18	19-20	21-23	24-26	27-29	30-32
No. of students	0	4	6	8	2	2

Lower class – 0.5 and upper class +0.5 class

Age (in years)	15.5-18.5	18.5-20.5	20.5-23.5	23.5-26.5	26.5-29.5	29.5-32.5
No. of students	0	4	6	8	2	2