

Chapter-13

STATISTICS

2 MARKS QUESTIONS

1. The following data represents the number of hours students spend on homework in a week in a class: 2, 3, 4, 5, 2, 3, 4, 3, 2, 5. Calculate the mean.

Solution:

$$\text{Mean} = (2 + 3 + 4 + 5 + 2 + 3 + 4 + 3 + 2 + 5) / 10$$

$$= 33 / 10$$

$$= 3.3$$

2. In a class of 50 students, the heights of the students were recorded. The mean height was found to be 160 cm. If the height of the teacher is included, whose height is 180 cm, find the new mean.

Solution:

$$\text{The total height of the students} = 50 * 160 \text{ cm}$$

$$\text{New total height including the teacher} = 50 * 160 \text{ cm} + 180 \text{ cm}$$

$$\text{New mean} = (\text{Total height including teacher}) / (\text{Number of students} + 1)$$

$$= (50 * 160 + 180) / 51$$

$$= (8000 + 180) / 51$$

$$= 8180 / 51$$

$$= 160 \text{ cm (approximately)}$$

3. Define the term 'median' and explain how to find it for a given set of data.

Solution:

The median of a set of data is the middle value when the data is arranged in ascending or descending order. To find the median, arrange the data in order and choose the middle value. If there are even data points, the median is the average of the two middle values.

4.If the mean of a data set is 20 and there are 5 observations, find the sum of all the observations.

Solution:

Sum of observations = Mean * Number of observations

$$= 20 * 5$$

$$= 100$$

5.Calculate the mean of the following data set: 10, 15, 20, 25, 30.

Solution:

$$\text{Mean} = (10 + 15 + 20 + 25 + 30) / 5$$

$$= 100 / 5$$

$$= 20$$

6.If the range of a data set is 12 and the minimum value is 5, what is the maximum value?

Solution:

$$\begin{aligned}\text{Maximum value} &= \text{Minimum value} + \text{Range} \\ &= 5 + 12 \\ &= 17\end{aligned}$$

7.Calculate the mean of the following data set: 12, 15, 18, 21, 24, 27.

Solution:

$$\begin{aligned}\text{Mean} &= (12 + 15 + 18 + 21 + 24 + 27) / 6 \\ &= 117 / 6 \\ &= 19.5\end{aligned}$$

8.If the mean of a data set is 30 and the sum of the data points is 150, how many data points are there in the set?

Solution:

$$\begin{aligned}\text{Number of data points} &= \text{Sum of data points} / \text{Mean} \\ &= 150 / 30 \\ &= 5\end{aligned}$$

4 MARKS QUESTIONS

1. A survey was conducted by a group of students as a part of their environment awareness program, in which they collected the following data regarding the number of plants in 20 houses in a locality. Find the mean number of plants per house.

Number of Plants	0-2	2-4	4-6	6-8	8-10	10-12	12-14
Number of Houses	1	2	1	5	6	2	3

Which method did you use for finding the mean, and why?

Solution:

To find the mean value, we will use the direct method because the numerical value of f_i and x_i are small.

Find the midpoint of the given interval using the formula.

Midpoint (x_i) = (upper limit + lower limit)/2

No. of plants (Class interval)	No. of houses Frequency (f_i)	Mid-point (x_i)	$f_i x_i$
0-2	1	1	1
2-4	2	3	6
4-6	1	5	5
6-8	5	7	35
8-10	6	9	54
10-12	2	11	22
12-14	3	13	39
	Sum $f_i = 20$		Sum $f_i x_i = 162$

The formula to find the mean is:

$$\text{Mean} = \bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

$$= 162/20$$

$$= 8.1$$

Therefore, the mean number of plants per house is 8.1.

2. Consider the following distribution of daily wages of 50 workers of a factory.

Daily wages (in Rs.)	500-520	520-540	540-560	560-580	580-600
Number of workers	12	14	8	6	10

Find the mean daily wages of the workers of the factory by using an appropriate method.

Solution:

Find the midpoint of the given interval using the formula.

$$\text{Midpoint } (x_i) = (\text{upper limit} + \text{lower limit})/2$$

In this case, the value of mid-point (x_i) is very large, so let us assume the mean value, $a = 550$.

$$\text{Class interval } (h) = 20$$

$$\text{So, } u_i = (x_i - a)/h$$

$$u_i = (x_i - 550)/20$$

Substitute and find the values as follows:

Daily wages (Class interval)	Number of workers frequency (f_i)	Mid-point (x_i)	$u_i = (x_i - 550)/20$	$f_i u_i$
500-520	12	510	-2	-24
520-540	14	530	-1	-14
540-560	8	550 = a	0	0
560-580	6	570	1	6
580-600	10	590	2	20
Total	Sum $f_i = 50$			Sum $f_i u_i = -12$

So, the formula to find out the mean is:

$$\text{Mean} = \bar{x} = a + h(\sum f_i u_i / \sum f_i) = 550 + [20 \times (-12/50)] = 550 - 4.8 = 545.20$$

Thus, mean daily wage of the workers = Rs. 545.20

3. In a retail market, fruit vendors were selling mangoes kept in packing boxes. These boxes contained varying number of mangoes. The following was the distribution of mangoes according to the number of boxes.

Number of mangoes	50-52	53-55	56-58	59-61	62-64
Number of boxes	15	110	135	115	25

Find the mean number of mangoes kept in a packing box. Which method of finding the mean did you choose?

Solution:

The given data is not continuous, so we add 0.5 to the upper limit and subtract 0.5 from the lower limit as the gap between two intervals is 1.

Here, assumed mean (a) = 57

Class size (h) = 3

Here, the step deviation is used because the frequency values are big.

Class Interval	Number of boxes (f_i)	Mid-point (x_i)	$u_i = (x_i - 57)/h$	$f_i u_i$
49.5-52.5	15	51	-2	-30
52.5-55.5	110	54	-1	-110
55.5-58.5	135	57 = a	0	0
58.5-61.5	115	60	1	115
61.5-64.5	25	63	2	50
	Sum $f_i = 400$			Sum $f_i u_i = 25$

The formula to find out the Mean is:

$$\text{Mean} = \bar{x} = a + h(\sum f_i u_i / \sum f_i)$$

$$= 57 + 3(25/400)$$

$$= 57 + 0.1875$$

$$= 57.19$$

Therefore, the mean number of mangoes kept in a packing box is 57.19

4. The table below shows the daily expenditure on food of 25 households in a locality.

Daily expenditure(in c)	100-150	150-200	200-250	250-300	300-350
Number of households	4	5	12	2	2

Find the mean daily expenditure on food by a suitable method.

Solution:

Find the midpoint of the given interval using the formula.

Midpoint (x_i) = (upper limit + lower limit)/2

Let us assume the mean (a) = 225

Class size (h) = 50

Class Interval	Number of households (f_i)	Mid-point (x_i)	$d_i = x_i - A$	$u_i = d_i/50$	$f_i u_i$
100-150	4	125	-100	-2	-8
150-200	5	175	-50	-1	-5
200-250	12	225 = a	0	0	0
250-300	2	275	50	1	2
300-350	2	325	100	2	4
	Sum $f_i = 25$				Sum $f_i u_i = -7$

Mean = $\bar{x} = a + h(\sum f_i u_i / \sum f_i)$

Mathematics

$$= 225 + 50(-7/25)$$

$$= 225 - 14$$

$$= 211$$

Therefore, the mean daily expenditure on food is 211.

5. To find out the concentration of SO₂ in the air (in parts per million, i.e., ppm), the data was collected for 30 localities in a certain city and is presented below:

Concentration of SO ₂ (in ppm)	Frequency
0.00 – 0.04	4
0.04 – 0.08	9
0.08 – 0.12	9
0.12 – 0.16	2
0.16 – 0.20	4
0.20 – 0.24	2

Find the mean concentration of SO₂ in the air.

Solution:

To find out the mean, first find the midpoint of the given frequencies as follows:

Concentration of SO ₂ (in ppm)	Frequency (f _i)	Mid-point (x _i)	f _i x _i
0.00-0.04	4	0.02	0.08
0.04-0.08	9	0.06	0.54
0.08-0.12	9	0.10	0.90
0.12-0.16	2	0.14	0.28
0.16-0.20	4	0.18	0.72
0.20-0.24	2	0.22	0.44
Total	Sum f _i = 30		Sum (f _i x _i) = 2.96

The formula to find out the mean is

$$\text{Mean} = \bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

$$= 2.96/30$$

$$= 0.099 \text{ ppm}$$

Therefore, the mean concentration of SO₂ in the air is 0.099 ppm.

6. A class teacher has the following absentee record of 40 students of a class for the whole term. Find the mean number of days a student was absent.

Number of days	0-6	6-10	10-14	14-20	20-28	28-38	38-40
Number of students	11	10	7	4	4	3	1

Solution:

Find the midpoint of the given interval using the formula.

Midpoint (x_i) = (upper limit + lower limit)/2

Class interval	Frequency (f_i)	Mid-point (x_i)	$f_i x_i$
0-6	11	3	33
6-10	10	8	80
10-14	7	12	84
14-20	4	17	68
20-28	4	24	96
28-38	3	33	99
38-40	1	39	39
	Sum $f_i = 40$		Sum $f_i x_i = 499$

The mean formula is,

$$\text{Mean} = \bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

$$= 499/40$$

$$= 12.48 \text{ days}$$

Therefore, the mean number of days a student was absent = 12.48.

7. The following table gives the literacy rate (in percentage) of 35 cities. Find the mean

literacy rate.

Literacy rate (in %)	45-55	55-65	65-75	75-85	85-98
Number of cities	3	10	11	8	3

Solution:

Find the midpoint of the given interval using the formula.

$$\text{Midpoint } (x_i) = (\text{upper limit} + \text{lower limit})/2$$

In this case, the value of mid-point (x_i) is very large, so let us assume the mean value, $a = 70$.

$$\text{Class interval } (h) = 10$$

$$\text{So, } u_i = (x_i - a)/h$$

$$u_i = (x_i - 70)/10$$

Substitute and find the values as follows:

Class Interval	Frequency (f_i)	(x_i)	$u_i = (x_i - 70)/10$	$f_i u_i$
45-55	3	50	-2	-6
55-65	10	60	-1	-10
65-75	11	70 = a	0	0
75-85	8	80	1	8
85-95	3	90	2	6
	Sum $f_i = 35$			Sum $f_i u_i = -2$

So, Mean = $\bar{x} = a + (\sum f_i u_i / \sum f_i) \times h$

$$= 70 + (-2/35) \times 10$$

$$= 69.43$$

Therefore, the mean literacy part = 69.43%

8. The given distribution shows the number of runs scored by some top batsmen of the world in one- day international cricket matches.

Run Scored	Number of Batsman
3000-4000	4
4000-5000	18
5000-6000	9
6000-7000	7

7000-8000	6
8000-9000	3
9000-10000	1
10000-11000	1

Find the mode of the data.

Solution:

Given data:

Modal class = 4000 – 5000,

$l = 4000$,

class width (h) = 1000,

$f_m = 18$, $f_1 = 4$ and $f_2 = 9$

Mode Formula:

$$\text{Mode} = l + \left[\frac{(f_m - f_1)}{(2f_m - f_1 - f_2)} \right] \times h$$

Substitute the values

$$\text{Mode} = 4000 + \left[\frac{(18 - 4)}{(36 - 4 - 9)} \right] \times 1000$$

$$= 4000 + (14000/23)$$

$$= 4000 + 608.695$$

$$= 4608.695$$

$$= 4608.7 \text{ (approximately)}$$

Thus, the mode of the given data is 4608.7 runs.

9. A student noted the number of cars passing through a spot on a road for 100 periods each of 3 minutes and summarized it in the table given below. Find the mode of the data:

Number of cars	Frequency
0-10	7
10-20	14
20-30	13
30-40	12
40-50	20
50-60	11
60-70	15
70-80	8

Solution:

Given Data:

Modal class = 40 – 50, $l = 40$,

Class width (h) = 10, $f_m = 20$, $f_1 = 12$ and $f_2 = 11$

$$\text{Mode} = l + \left[\frac{(f_m - f_1)}{(2f_m - f_1 - f_2)} \right] \times h$$

Substitute the values

$$\text{Mode} = 40 + \left[\frac{(20 - 12)}{(40 - 12 - 11)} \right] \times 10$$

$$= 40 + (80/17)$$

$$= 40 + 4.7$$

$$= 44.7$$

Thus, the mode of the given data is 44.7 cars.

10. The life insurance agent found the following data for the distribution of ages of 100 policy holders. Calculate the median age, if policies are given only to the persons whose age is 18 years onwards but less than the 60 years.

Age (in years)	Number of policy holder
Below 20	2
Below 25	6
Below 30	24
Below 35	45
Below 40	78
Below 45	89
Below 50	92
Below 55	98
Below 60	100

Solution:

Class interval	Frequency	Cumulative frequency
15-20	2	2
20-25	4	6
25-30	18	24
30-35	21	45
35-40	33	78
40-45	11	89
45-50	3	92
50-55	6	98
55-60	2	100

Given data: $N = 100$ and $N/2 = 50$

Median class = 35-40

Then, $l = 35$, $cf = 45$, $f = 33$ & $h = 5$

$$\text{Median} = l + \frac{\frac{N}{2} - cf}{f} \times h$$

$$\text{Median} = 35 + [(50 - 45)/33] \times 5$$

$$= 35 + (25/33)$$

$$= 35.76$$

Therefore, the median age = 35.76 years.

11. The lengths of 40 leaves in a plant are measured correctly to the nearest millimeter, and the data obtained is represented as in the following table:

Length (in mm)	Number of leaves
118-126	3
127-135	5
136-144	9
145-153	12
154-162	5
163-171	4
172-180	2

Find the median length of the leaves.

(Hint : The data needs to be converted to continuous classes for finding the median, since the formula assumes continuous classes. The classes then change to 117.5 – 126.5, 126.5 – 135.5, . . . , 171.5 – 180.5.)

Solution:

Since the data are not continuous reduce 0.5 in the lower limit and add 0.5 in the upper limit.

Class Interval	Frequency	Cumulative frequency
117.5-126.5	3	3
126.5-135.5	5	8
135.5-144.5	9	17
144.5-153.5	12	29
153.5-162.5	5	34
162.5-171.5	4	38
171.5-180.5	2	40

So, the data obtained are:

$$N = 40 \text{ and } N/2 = 20$$

$$\text{Median class} = 144.5-153.5$$

$$\text{then, } l = 144.5,$$

$$cf = 17, f = 12 \text{ \& } h = 9$$

$$\text{Median} = l + \frac{\frac{N}{2} - cf}{f} \times h$$

$$\text{Median} = 144.5 + [(20 - 17) / 12] \times 9$$

$$= 144.5 + (9/4)$$

$$= 146.75 \text{ mm}$$

Therefore, the median length of the leaves = 146.75 mm.

12. The following table gives the distribution of a lifetime of 400 neon lamps.

Lifetime (in hours)	Number of lamps
1500-2000	14
2000-2500	56
2500-3000	60
3000-3500	86
3500-4000	74
4000-4500	62
4500-5000	48

Find the median lifetime of a lamp.

Solution:

Class Interval	Frequency	Cumulative
1500-2000	14	14
2000-2500	56	70
2500-3000	60	130
3000-3500	86	216
3500-4000	74	290
4000-4500	62	352

4500-5000	48	400
-----------	----	-----

Data:

$$N = 400 \text{ \& } N/2 = 200$$

$$\text{Median class} = 3000 - 3500$$

$$\text{Therefore, } l = 3000, \text{ cf} = 130,$$

$$f = 86 \text{ \& } h = 500$$

$$\text{Median} = l + \frac{\frac{N}{2} - cf}{f} \times h$$

$$\text{Median} = 3000 + [(200 - 130)/86] \times 500$$

$$= 3000 + (35000/86)$$

$$= 3000 + 406.98$$

$$= 3406.98$$

Therefore, the median lifetime of the lamps = 3406.98 hours

13. The distribution below gives the weights of 30 students of a class. Find the median weight of the students.

Weight(in kg)	40-45	45-50	50-55	55-60	60-65	65-70	70-75
Number of students	2	3	8	6	6	3	2

Solution:

Class Interval	Frequency	Cumulative frequency
40-45	2	2
45-50	3	5
50-55	8	13
55-60	6	19
60-65	6	25
65-70	3	28
70-75	2	30

Given: $N = 30$ and $N/2 = 15$

Median class = 55-60

$l = 55$, $C_f = 13$, $f = 6$ & $h = 5$

$$\text{Median} = l + \frac{\frac{N}{2} - C_f}{f} \times h$$

$$\text{Median} = 55 + [(15 - 13)/6] \times 5$$

$$= 55 + (10/6)$$

$$= 55 + 1.666$$

$$= 56.67$$

Therefore, the median weight of the students = 56.67

7 MARKS QUESTIONS

1. The following distribution shows the daily pocket allowance of children of a locality. The mean pocket allowance is Rs 18. Find the missing frequency f .

Daily Pocket Allowance(in c)	11-13	13-15	15-17	17-19	19-21	21-23	23-35
Number of children	7	6	9	13	f	5	4

Solution:

To find out the missing frequency, use the mean formula.

Given, mean $\bar{x} = 18$

Class interval	Number of children (f_i)	Mid-point (x_i)	$f_i x_i$
11-13	7	12	84
13-15	6	14	84
15-17	9	16	144
17-19	13	18	234
19-21	F	20	$20f$
21-23	5	22	110
23-25	4	24	96

Total	$f_i = 44+f$		Sum $f_i x_i =$ 752+20f
-------	--------------	--	----------------------------

The mean formula is

$$\text{Mean} = \bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{(752 + 20f)}{(44 + f)}$$

Now substitute the values and equate to find the missing frequency (f)

$$\Rightarrow 18 = \frac{(752 + 20f)}{(44 + f)}$$

$$\Rightarrow 18(44 + f) = (752 + 20f)$$

$$\Rightarrow 792 + 18f = 752 + 20f$$

$$\Rightarrow 792 + 18f = 752 + 20f$$

$$\Rightarrow 792 - 752 = 20f - 18f$$

$$\Rightarrow 40 = 2f$$

$$\Rightarrow f = 20$$

So, the missing frequency, $f = 20$.

2. Thirty women were examined in a hospital by a doctor, and the number of heartbeats per minute were recorded and summarised as follows. Find the mean heartbeats per minute for these women, choosing a suitable method.

Number of heart beats per minute	65- 68	68- 71	71- 74	74- 77	77- 80	80- 83	83- 86
Number of women	2	4	3	8	7	4	2

Solution:

From the given data, let us assume the mean as $a = 75.5$

$$x_i = (\text{Upper limit} + \text{Lower limit})/2$$

$$\text{Class size (h)} = 3$$

Now, find the u_i and $f_i u_i$ as follows:

Class Interval	Number of women (f_i)	Mid-point (x_i)	$u_i = (x_i - 75.5)/h$	$f_i u_i$
65-68	2	66.5	-3	-6
68-71	4	69.5	-2	-8
71-74	3	72.5	-1	-3
74-77	8	75.5 = a	0	0
77-80	7	78.5	1	7
80-83	4	81.5	2	8
83-86	2	84.5	3	6
	Sum $f_i = 30$			Sum $f_i u_i = 4$

$$\text{Mean} = \bar{x} = a + h(\sum f_i u_i / \sum f_i)$$

$$= 75.5 + 3 \times (4/30)$$

$$= 75.5 + (4/10)$$

$$= 75.5 + 0.4$$

$$= 75.9$$

Therefore, the mean heart beats per minute for these women is 75.9

3. To find out the concentration of SO_2 in the air (in parts per million, i.e., ppm), the data was collected for 30 localities in a certain city and is presented below:

Concentration of SO_2 (in ppm)	Frequency
0.00 – 0.04	4
0.04 – 0.08	9
0.08 – 0.12	9
0.12 – 0.16	2
0.16 – 0.20	4
0.20 – 0.24	2

Find the mean concentration of SO_2 in the air.

Solution:

To find out the mean, first find the midpoint of the given frequencies as follows:

Concentration of SO_2 (in ppm)	Frequency (f_i)	Mid-point (x_i)	$f_i x_i$
0.00-0.04	4	0.02	0.08
0.04-0.08	9	0.06	0.54
0.08-0.12	9	0.10	0.90

0.12-0.16	2	0.14	0.28
0.16-0.20	4	0.18	0.72
0.20-0.24	2	0.22	0.44
Total	Sum $f_i = 30$		Sum $(f_i x_i) = 2.96$

The formula to find out the mean is

$$\text{Mean} = \bar{x} = \sum f_i x_i / \sum f_i$$

$$= 2.96/30$$

$$= 0.099 \text{ ppm}$$

Therefore, the mean concentration of SO_2 in the air is 0.099 ppm.

4. The following table shows the ages of the patients admitted to a hospital during a year:

Age (in years)	5-15	15-25	25-35	35-45	45-55	55-65
Number of patients	6	11	21	23	14	5

Find the mode and the mean of the data given above. Compare and interpret the two

measures of central tendency.

Solution:

To find out the modal class, let us consider the class interval with high frequency.

Here, the greatest frequency = 23, so the modal class = 35 – 45,

Lower limit of modal class = $l = 35$,

class width (h) = 10,

$f_m = 23$,

$f_1 = 21$ and $f_2 = 14$

The formula to find the mode is

$$\text{Mode} = l + \left[\frac{(f_m - f_1)}{(2f_m - f_1 - f_2)} \right] \times h$$

Substitute the values in the formula, we get

$$\text{Mode} = 35 + \left[\frac{(23 - 21)}{(46 - 21 - 14)} \right] \times 10$$

$$= 35 + (20/11)$$

$$= 35 + 1.8$$

$$= 36.8 \text{ years}$$

So the mode of the given data = 36.8 years

Calculation of Mean:

First find the midpoint using the formula, $x_i = (\text{upper limit} + \text{lower limit})/2$

Class Interval	Frequency (f_i)	Mid-point (x_i)	$f_i x_i$
5-15	6	10	60
15-25	11	20	220
25-35	21	30	630
35-45	23	40	920
45-55	14	50	700

55-65	5	60	300
	Sum $f_i = 80$		Sum $f_i x_i = 2830$

The mean formula is

$$\text{Mean} = \bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

$$= 2830/80$$

$$= 35.375 \text{ years}$$

Therefore, the mean of the given data = 35.375 years

5. The following data gives the information on the observed lifetimes (in hours) of 225

electrical components:

Lifetime (in hours)	0-20	20-40	40-60	60-80	80-100	100-120
Frequency	10	35	52	61	38	29

Determine the modal lifetimes of the components.

Solution:

From the given data the modal class is 60–80.

Lower limit of modal class = $l = 60$,

The frequencies are:

$$f_m = 61, f_1 = 52, f_2 = 38 \text{ and } h = 20$$

The formula to find the mode is

$$\text{Mode} = l + \left[\frac{(f_m - f_1)}{(2f_m - f_1 - f_2)} \right] \times h$$

Substitute the values in the formula, we get

$$\text{Mode} = 60 + [(61 - 52) / (122 - 52 - 38)] \times 20$$

$$\text{Mode} = 60 + [(9 \times 20) / 32]$$

$$\text{Mode} = 60 + (45/8) = 60 + 5.625$$

Therefore, modal lifetime of the components = 65.625 hours.

6. The following data gives the distribution of total monthly household expenditure of 200

families of a village. Find the modal monthly expenditure of the families. Also, find the

mean monthly expenditure:

Expenditure (in Rs.)	Number of families
1000-1500	24
1500-2000	40
2000-2500	33
2500-3000	28
3000-3500	30
3500-4000	22
4000-4500	16
4500-5000	7

Solution:

Given data:

Modal class = 1500-2000,

$l = 1500$,

Frequencies:

$f_m = 40$ $f_1 = 24$, $f_2 = 33$ and

$h = 500$

Mode formula:

$$\text{Mode} = l + \left[\frac{(f_m - f_1)}{(2f_m - f_1 - f_2)} \right] \times h$$

Substitute the values in the formula, we get

$$\text{Mode} = 1500 + \left[\frac{(40 - 24)}{(80 - 24 - 33)} \right] \times 500$$

$$\text{Mode} = 1500 + \left[\frac{(16 \times 500)}{23} \right]$$

$$\text{Mode} = 1500 + (8000/23) = 1500 + 347.83$$

Therefore, modal monthly expenditure of the families = Rupees 1847.83

Calculation for mean:

First find the midpoint using the formula, $x_i = (\text{upper limit} + \text{lower limit})/2$

Let us assume a mean, (a) be 2750.

Class Interval	f_i	x_i	$d_i = x_i - a$	$u_i = d_i/h$	$f_i u_i$
1000-1500	24	1250	-1500	-3	-72
1500-2000	40	1750	-1000	-2	-80

2000-2500	33	2250	-500	-1	-33
2500-3000	28	2750 = a	0	0	0
3000-3500	30	3250	500	1	30
3500-4000	22	3750	1000	2	44
4000-4500	16	4250	1500	3	48
4500-5000	7	4750	2000	4	28
	$f_i = 200$				$f_i u_i = -35$

The formula to calculate the mean,

$$\text{Mean} = \bar{x} = a + (\sum f_i u_i / \sum f_i) \times h$$

Substitute the values in the given formula

$$= 2750 + (-35/200) \times 500$$

$$= 2750 - 87.50$$

$$= 2662.50$$

So, the mean monthly expenditure of the families = Rs. 2662.50

7. The following distribution gives the state-wise teacher-student ratio in higher secondary schools of India. Find the mode and mean of this data. Interpret the two measures

No of students per teacher	Number of states / U.T
15-20	3
20-25	8
25-30	9
30-35	10
35-40	3
40-45	0
45-50	0
50-55	2

Solution:

Given data:

Modal class = 30 – 35,

$l = 30$,

Class width (h) = 5,

$f_m = 10$, $f_1 = 9$ and $f_2 = 3$

Mode Formula:

$$\text{Mode} = l + \frac{[f_m - f_1]}{(2f_m - f_1 - f_2)} \times h$$

Substitute the values in the given formula

$$\text{Mode} = 30 + [(10 - 9) / (20 - 9 - 3)] \times 5$$

$$= 30 + (5/8)$$

$$= 30 + 0.625$$

$$= 30.625$$

Therefore, the mode of the given data = 30.625

Calculation of mean:

Find the midpoint using the formula, $x_i = (\text{upper limit} + \text{lower limit})/2$

Class Interval	Frequency (f_i)	Mid-point (x_i)	$f_i x_i$
15-20	3	17.5	52.5
20-25	8	22.5	180.0
25-30	9	27.5	247.5
30-35	10	32.5	325.0
35-40	3	37.5	112.5
40-45	0	42.5	0
45-50	0	47.5	0
50-55	2	52.5	105.0
	Sum $f_i = 35$		Sum $f_i x_i = 1022.5$

$$\text{Mean} = \bar{x} = \sum f_i x_i / \sum f_i$$

$$= 1022.5/35$$

$$= 29.2 \text{ (approx)}$$

Therefore, mean = 29.2

8. The following frequency distribution gives the monthly consumption of an electricity of 68 consumers in a locality. Find the median, mean and mode of the data and compare them.

Monthly consumption(in units)	No. of customers
65-85	4
85-105	5
105-125	13
125-145	20
145-165	14
165-185	8
185-205	4

Solution:

Find the cumulative frequency of the given data as follows:

Class Interval	Frequency	Cumulative frequency
65-85	4	4
85-105	5	9
105-125	13	22

125-145	20	42
145-165	14	56
165-185	8	64
185-205	4	68
	N = 68	

From the table, it is observed that, $N = 68$ and hence $N/2=34$

Hence, the median class is 125-145 with cumulative frequency = 42

Where, $l = 125$, $N = 68$, $cf = 22$, $f = 20$, $h = 20$

Median is calculated as follows:

$$\begin{aligned}
 \text{Median} &= l + \frac{\frac{N}{2} - cf}{f} \times h \\
 &= 125 + [(34 - 22)/20] \times 20 \\
 &= 125 + 12 \\
 &= 137
 \end{aligned}$$

Therefore, median = 137

To calculate the mode:

Modal class = 125-145,

f_m or $f_1 = 20$, $f_0 = 13$, $f_2 = 14$ & $h = 20$

Mode formula:

$$\begin{aligned}
 \text{Mode} &= l + [(f_1 - f_0)/(2f_1 - f_0 - f_2)] \times h \\
 \text{Mode} &= 125 + [(20 - 13)/(40 - 13 - 14)] \times 20
 \end{aligned}$$

Mathematics

$$= 125 + (140/13)$$

$$= 125 + 10.77$$

$$= 135.77$$

Therefore, mode = 135.77

Calculate the Mean:

Class Interval	f_i	x_i	$d_i = x_i - a$	$u_i = d_i/h$	$f_i u_i$
65-85	4	75	-60	-3	-12
85-105	5	95	-40	-2	-10
105-125	13	115	-20	-1	-13
125-145	20	135 = a	0	0	0
145-165	14	155	20	1	14
165-185	8	175	40	2	16
185-205	4	195	60	3	12
	Sum $f_i = 68$				Sum $f_i u_i = 7$

$$\bar{x} = a + h (\sum f_i u_i / \sum f_i) = 135 + 20 (7/68)$$

$$\text{Mean} = 137.05$$

In this case, mean, median and mode are more/less equal in this distribution.

9. If the median of a distribution given below is 28.5, find the value of x & y .

Class Interval	Frequency
0-10	5
10-20	x
20-30	20
30-40	15
40-50	y
50-60	5
Total	60

Solution:

Given data, $n = 60$

Median of the given data = 28.5

CI	0-10	10-20	20-30	30-40	40-50	50-60
Frequency	5	x	20	15	y	5
Cumulative frequency	5	$5+x$	$25+x$	$40+x$	$40+x+y$	$45+x+y$

Where, $N/2 = 30$

Median class is 20 – 30 with a cumulative frequency = $25 + x$

Lower limit of median class, $l = 20$,

$cf = 5 + x$,

Mathematics

$$f = 20 \text{ \& } h = 10$$

$$\text{Median} = l + \frac{\frac{N}{2} - cf}{f} \times h$$

Substitute the values

$$28.5 = 20 + [(30 - 5 - x)/20] \times 10$$

$$8.5 = (25 - x)/2$$

$$17 = 25 - x$$

Therefore, $x = 8$.

Now, from cumulative frequency, we can identify the value of $x + y$ as follows:

Since,

$$60 = 45 + x + y$$

Now, substitute the value of x , to find y

$$60 = 45 + 8 + y$$

$$y = 60 - 53$$

$$y = 7$$

Therefore, the value of $x = 8$ and $y = 7$.

10.100 surnames were randomly picked up from a local telephone directory and the frequency distribution of the number of letters in the English alphabets in the surnames was obtained as follows:

Number of letters	1-4	4-7	7-10	10-13	13-16	16-19
Number of surnames	6	30	40	16	4	4

Determine the median number of letters in the surnames. Find the mean number of letters in the surnames. Also, find the modal size of the surnames.

Solution:

To calculate median:

Class Interval	Frequency	Cumulative Frequency
1-4	6	6
4-7	30	36
7-10	40	76
10-13	16	92
13-16	4	96
16-19	4	100

Given:

$$N = 100 \text{ \& } N/2 = 50$$

Median class = 7-10

Therefore, $l = 7$, $cf = 36$, $f = 40$ & $h = 3$

$$\text{Median} = l + \frac{\frac{N}{2} - cf}{f} \times h$$

$$\text{Median} = 7 + [(50 - 36)/40] \times 3$$

$$\text{Median} = 7 + (14/40)$$

$$\text{Median} = 8.05$$

Calculate the Mode:

Modal class = 7-10,

Where, $l = 7$, $f_1 = 40$, $f_0 = 30$, $f_2 = 16$ & $h = 3$

$$\text{Mode} = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right) \times h$$

$$\text{Mode} = 7 + [(40 - 30)/(2 \times 40 - 30 - 16)] \times 3$$

$$= 7 + (10/34)$$

$$= 7.88$$

Therefore mode = 7.88

Calculate the Mean:

Class Interval	f_i	x_i	$f_i x_i$
1-4	6	2.5	15
4-7	30	5.5	165
7-10	40	8.5	340
10-13	16	11.5	184
13-16	4	14.5	58

Mathematics

16-19	4	17.5	70
	Sum $f_i = 100$		Sum $f_i x_i = 832$

$$\text{Mean} = \bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

$$\text{Mean} = 832/100 = 8.32$$

Therefore, mean = 8.32