Chapter 14: Statistics

Exercise 14.1

Q1. A survey was conducted by a group of students as a part of their environment awareness programme, 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?

Difficulty Level:
Easy

Known:
The number of plants in 20 houses in a locality.
Unknown:
The mean number of plants per house and the method used for finding the mean.

Reasoning:
We can solve this question by any method of finding mean but here we will use direct method to solve this question because the data given is small.

The mean (or average) of observations, as we know, is the sum of the values of all the observations divided by the total number of observations.

We know that if $x_1, x_2, \ldots, x_n$ are observations with respective frequencies $f_1, f_2, \ldots, f_n$, then this means observation x_1 occurs f_1 times, x_2 occurs f_2 times, and so on.

x is the class mark for each interval, you can find the value of x by using

$$\text{class mark, } x_i = \frac{\text{upper limit} + \text{lower limit}}{2}$$

Now, the sum of the values of all the observations $= f_1x_1 + f_2x_2 + \ldots + f_nx_n$, and the number of observations $= f_1 + f_2 + \ldots + f_n$.
So, the mean of the data is given by

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \cdots\cdots + f_nx_n}{f_1 + f_2 + \cdots\cdots + f_n}$$

$$\bar{x} = \frac{\sum f_ix_i}{\sum f_i} \text{ where } i \text{ varies from 1 to } n.$$

Solution:

Number of plants	Number of houses f_i	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
	$\sum f_i = 20$		$\sum f_i x_i = 162$

From the table it can be observed that,

$$\sum f_i = 20$$
$$\sum f_i x_i = 162$$

Mean, $\bar{x} = \dfrac{\sum f_i x_i}{\sum f_i}$

$= \dfrac{162}{20}$

$= 8.1$

Thus, the mean number of plants each house has 8.1.
Here, we have used the direct method because the value of x_i and f_i are small.

Q2. 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.

Difficulty Level:
Moderate

Known:
Distribution of daily wages of 50 workers of a factory is given-

Unknown:
The mean daily wages of the workers of the factory.

Reasoning:

We will use Assumed Mean Method to solve this question because the data given is large.

Sometimes when the numerical values of x_i and f_i are large, finding the product of x_i and f_i becomes tedious. We can do nothing with the f_i, but we can change each x_i to a smaller number so that our calculations become easy. Now we have to subtract a fixed number from each of these x_i.

The first step is to choose one among the x_i as the **assumed mean** and denote it by '***a***'. Also, to further reduce our calculation work, we may take '***a***' to be that x_i which lies in the centre of $x_1, x_2, \ldots, x_n$. So, we can choose ***a***.

The next step is to find the difference 'd_i' between ***a*** and each of the x_i, that is, the deviation of '***a***' from each of the x_i. i.e., $d_i = x_i - a$

The third step is to find the product of d_i with the corresponding f_i, and take the sum of all the $f_i d_i$

Now put the values in the below formula

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i d_i}{\sum f_i} \right)$$

Solution:

We know that,

$$\text{Class mark, } x_i = \frac{\text{Upper class limit} + \text{Lower class limit}}{2}$$

Taking assumed mean, $a = 550$

Daily wages (in ₹)	No of workers f_i	x_i	$d_i = x_i - a$	$f_i d_i$
500 – 520	12	510	– 40	– 480
520 – 540	14	530	– 20	– 280
540 – 560	8	550 (***a***)	0	0
560 – 580	6	570	20	120
580 – 600	10	590	40	400
	$\sum f_i = 50$			$\sum f_i d_i = -240$

It can be observed from the table,

$$\sum f_i = 50$$

$$\sum f_i u_i = -240$$

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i d_i}{\sum f_i} \right)$$

$$= 550 + \left(\frac{-240}{50}\right)$$

$$= 550 - \frac{24}{5}$$

$$= 550 - 4.8$$

$$= 545.2$$

Thus, the mean daily wages of the workers of the factory is ₹ 545.20

Q3. The following distribution shows the daily pocket allowance of children of a locality. The mean pocket allowance is ₹ 18. Find the missing frequency *f*.

Daily pocket allowance (in ₹)	11 – 13	13 – 15	15 – 17	17 – 19	19 – 21	21 – 23	23 – 25
Number of children	7	6	9	13	*f*	5	4

Difficulty Level:
Moderate

Known:
The mean pocket allowance is ₹ 18.

Unknown:
The missing frequency *f*.

Reasoning:
We will use assumed mean method to solve this question.

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i d_i}{\sum f_i}\right)$$

Solution:
We know that,

$$\text{Class mark, } x_i = \frac{\text{Upper class limit} + \text{Lower class limit}}{2}$$

Taking assumed mean, $a = 18$

Daily pocket allowance (in ₹)	**No of children** f_i	x_i	$d_i = x_i - a$	$f_i d_i$
11 – 13	7	12	–6	–42
13 – 15	6	14	–4	–24
15 – 17	9	16	–2	–18
17 – 19	13	18 (***a***)	0	0
19 – 21	*f*	20	2	2*f*
21 – 23	5	22	4	20
23 – 25	4	24	6	24
	$\sum f_i = 40 + f$			$\sum f_i d_i = 2f - 40$

From the table, we obtain

$$\sum f_i = 40 + f$$

$$\sum f_i d_i = 2f - 40$$

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i d_i}{\sum f_i} \right)$$

$$18 = 18 + \left(\frac{2f - 40}{40 + f} \right)$$

$$18 - 18 = \frac{2f - 40}{40 + f}$$

$$2f - 40 = 0$$

$$f = 20$$

Hence, the missing frequency *f* is 20.

Q4. Thirty women were examined in a hospital by a doctor and the number of heart beats per minute were recorded and summarised as follows. Find the mean heart beats 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

Difficulty Level:

Moderate

Known:

The heart beats per minute of 30 women.

Unknown:

The mean heart beats per minute for these women.

Reasoning:

We will use Step-deviation Method to solve this question because the data given is large and will be convenient to apply if all the d_i have a common factor.

Sometimes when the numerical values of x_i and f_i are large, finding the product of x_i and f_i becomes tedious. We can do nothing with the f_i, but we can change each x_i to a smaller number so that our calculations become easy. Now we have to subtract a fixed number from each of these x_i.

The first step is to choose one among the x_i as the **assumed mean** and denote it by '***a***'. Also, to further reduce our calculation work, we may take '***a***' to be that x_i which lies in the centre of $x_1, x_2, \ldots, x_n$. So, we can choose ***a***.

The next step is to find the difference 'd_i' between a and each of the x_i, that is, the deviation of 'a' from each of the x_i. i.e., $d_i = x_i - a$

The third step is to find 'u_i' by dividing d_i and class size h for each of the x_i. i.e., $u_i = \frac{d_i}{h}$

The next step is to find the product of u_i with the corresponding f_i, and take the sum of all the $f_i u_i$

The step-deviation method will be convenient to apply if all the d_i have a common factor.

Now put the values in the below formula

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i}\right) \times h$$

Solution:

We know that,

$$\text{Class mark, } x_i = \frac{\text{Upper class limit} + \text{Lower class limit}}{2}$$

Class size, $h = 3$

Taking assumed mean, $a = 75.5$

Number of heart beats per minute	No. of women f_i	x_i	$d_i = x_i - a$	$u_i = \frac{d_i}{h}$	$f_i u_i$
65 – 68	2	66.5	–9	–3	–6
68 – 71	4	69.5	–6	–2	–8
71 – 74	3	72.5	–3	–1	–3
74 – 77	8	75.5 (***a***)	0	0	0
77 – 80	7	78.5	3	1	7
80 – 83	4	81.5	6	2	8
83 – 86	2	84.5	9	3	6
	$\sum f_i = 30$				$\sum f_i u_i = 4$

From the table, we obtain

$$\sum f_i = 30$$

$$\sum f_i u_i = 4$$

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i}\right) \times h$$

$$= 75.5 + \left(\frac{4}{30}\right) \times 3$$

$$= 75.5 - \frac{2}{5}$$

$$= 75.5 - 0.4$$

$$= 75.9$$

Hence, the mean heartbeat per minute for these women is 75.9

Q5. 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?

Difficulty Level:

Moderate

Known:

The distribution of mangoes according to the number of boxes.

Unknown:

The mean number of mangoes kept in a packing box.

Reasoning:

We will solve this question by step-deviation method.
Hence, the given class interval is not continuous. First, we have to make it continuous.
There is a gap of 1 between two class intervals. Therefore, half of the gap i.e., 0.5 has to be added to the upper-class limit and 0.5 has to be subtracted from the lower-class limit of each interval.

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i} \right) \times h$$

Solution:

We know that,

$$\text{Class mark, } x_i = \frac{\text{Upper class limit } + \text{ Lower class limit}}{2}$$

Class size, $h = 3$

Taking assumed mean, $a = 57$

Number of mangoes	**Number of boxes** f_i	x_i	$d_i = x_i - a$	$u_i = \frac{d_i}{h}$	$f_i u_i$
49.5 – 52.5	15	51	–6	–2	–30
52.5 – 55.5	110	54	–3	–1	–110
55.5 – 58.5	135	57 (***a***)	0	0	0
58.5 – 61.5	115	60	3	1	115
61.5 – 64.5	25	63	6	2	50
	$\sum f_i = 400$				$\sum f_i u_i = 25$

From the table, we obtain

$$\sum f_i = 400$$
$$\sum f_i u_i = 25$$

Mean, $\bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i} \right) \times h$

$$= 57 + \left(\frac{25}{400} \right) \times 3$$
$$= 57 + \frac{1}{16} \times 3$$
$$= 57 + \frac{3}{16}$$
$$= 57 + 0.19$$
$$= 57.19$$

The mean number of mangoes kept in a packing box are 57.19.

Q6. The table below shows the daily expenditure on food of 25 households in a locality.

Daily expenditure (in ₹)	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.

Difficulty Level:
Moderate

Known:
The daily expenditure on food of 25 households in a locality.

Unknown:
The mean daily expenditure on food.

Reasoning:
We will solve this question by step deviation method.

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i} \right) \times h$$

Solution:
We know that,

Class mark, $x_i = \frac{\text{Upper class limit } + \text{ Lower class limit}}{2}$

Class size, $h = 50$

Taking assumed mean, $\boldsymbol{a} = 225$

Daily expenditure (in ₹)	**Number of households** f_i	x_i	$d_i = x_i - a$	$u_i = \frac{d_i}{h}$	$f_i u_i$
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
	$\sum f_i = 25$				$\sum f_i u_i = -7$

From the table, we obtain

$$\sum f_i = 25$$

$$\sum f_i u_i = -7$$

Mean, $\bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i} \right) \times h$

$$= 225 + \left(\frac{-7}{25} \right) \times 50$$

$$= 225 - 14$$

$$= 211$$

Thus, the mean daily expenditure on food is ₹ 211.

Q7. 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.

Difficulty Level:
Moderate

Known:
The concentration of SO_2 in the air (in parts per million, i.e., ppm), for 30 localities in a certain city.

Unknown:
The mean concentration of SO_2 in the air.

Reasoning:
We will solve this question by step-deviation method.

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i} \right) \times h$$

Solution:
We know that,

$$\text{Class mark, } x_i = \frac{\text{Upper class limit } + \text{ Lower class limit}}{2}$$

Class size, $h = 0.04$
Taking assumed mean, $a = 0.14$

Concentration of SO_2 (in ppm)	**Frequency** f_i	x_i	$d_i = x_i - a$	$u_i = \frac{d_i}{h}$	$f_i u_i$
0.00 – 0.04	4	0.02	– 0.12	– 3	– 12
0.04 – 0.08	9	0.06	– 0.08	– 2	– 18
0.08 – 0.12	9	0.10	– 0.04	– 1	– 9
0.12 – 0.16	2	0.14 (***a***)	0	0	0
0.16 – 0.20	4	0.18	0.04	1	4
0.20 – 0.24	2	0.22	0.08	2	4
	$\sum f_i = 30$				$\sum f_i u_i = -31$

From the table, we obtain

$$\sum f_i = 30$$
$$\sum f_i u_i = -31$$

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i} \right) \times h$$
$$= 0.14 + \left(\frac{-31}{30} \right) \times 0.04$$
$$= 0.14 - 0.041$$
$$= 0.099$$

The mean concentration of SO_2 in the air is 0.099.

Q8. 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

Difficulty Level:
Moderate

Known:
The absentee record of 40 students of a class for the whole term.

Unknown:
The mean number of days a student was absent.

Reasoning:
We will solve this question by assumed mean method.

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i d_i}{\sum f_i} \right)$$

Solution:
We know that,

$$\text{Class mark, } x_i = \frac{\text{Upper class limit} + \text{Lower class limit}}{2}$$

Taking assumed mean, $a = 17$

Number of days	**Number of students** f_i	x_i	$d_i = x_i - a$	$f_i d_i$
0 – 6	11	3	–14	–154
6 – 10	10	8	–9	–90
10 – 14	7	12	–5	–35
14 – 20	4	17 (***a***)	0	0
20 – 28	4	24	7	28
28 – 38	3	33	18	48
38 – 40	1	39	22	22
	$\sum f_i = 40$			$\sum f_i d_i = -181$

From the table, we obtain

$$\sum f_i = 40$$

$$\sum f_i d_i = -181$$

Mean, $\bar{x} = a + \left(\frac{\sum f_i d_i}{\sum f_i} \right)$

$= 17 + \left(\frac{-181}{40} \right)$

$= 17 - 4.525$

$= 12.475$

$= 12.48$

Thus, the mean number of days a student was absent is 12.48.

Q9. 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 – 95
Number of cities	3	10	11	8	3

Difficulty Level:

Moderate

Known:

The literacy rate (in percentage) of 35 cities.

Unknown:

The mean literacy rate.

Reasoning:

We will solve this question by assumed mean method.

Mean, $\bar{x} = a + \left(\frac{\sum f_i d_i}{\sum f_i} \right)$

Solution:

We know that,

Class mark, $x_i = \frac{\text{Upper class limit } + \text{ Lower class limit}}{2}$

Taking assumed mean, $a = 70$

Literacy rate	**No of cities** f_i	x_i	$d_i = x_i - a$	$f_i d_i$
45 – 55	3	50	– 20	– 60
55 – 65	10	60	– 10	– 100
65 – 75	11	70 (*a*)	0	0
75 – 85	8	80	10	80
85 – 95	3	90	20	60
	$\sum f_i = 35$			$\sum f_i d_i = -20$

From the table, we obtain

$$\sum f_i = 35$$

$$\sum f_i d_i = -20$$

Mean, $\bar{x} = a + \left(\frac{\sum f_i d_i}{\sum f_i} \right)$

$$= 70 + \left(\frac{-20}{35} \right)$$

$$= 70 - 0.57$$

$$= 69.43$$

Thus, the mean literacy rate is 69.43%.

Chapter 14: Statistics

Exercise 14.2

Q1. The following table shows the ages of the patients admitted in 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.

Difficulty Level:
Moderate

Known:
The ages of the patients admitted in a hospital during a year.

Unknown:
The mode and the mean of the data and their comparison and interpretation.

Reasoning:
We will find the mean by direct method.

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

Modal Class is the class with highest frequency

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

Where,
Class size, h
Lower limit of modal class, l
Frequency of modal class, f_1
Frequency of class preceding modal class, f_0
Frequency of class succeeding the modal class, f_2

Solution:
To find Mean

We know that,

$$\text{Class mark, } x_i = \frac{\text{Upper class limit } + \text{ Lower class limit}}{2}$$

Age (in years)	**Number of patients** f_i	x_i	$f_i x_i$
5 – 15	6	10	6
15 – 25		20	220
25 – 35	21	30	630
35 – 45	23	40	920
45 – 55	14	50	700
55 – 65	5	60	300
	$\sum f_i = 80$		$\sum f_i x_i = 2830$

From the table it can be observed that,

$$\sum f_i = 80$$
$$\sum f_i x_i = 2830$$

Mean, $\bar{x} = \dfrac{\sum f_i x_i}{\sum f_i}$

$$= \frac{2830}{80}$$
$$= 35.37$$

To find mode

We know that, Modal Class is the class with highest frequency

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

From the table, it can be observed that the maximum class frequency is 23, belonging to class interval 35 − 45.

Therefore, Modal class = 35 − 45

Class size, $h = 10$

Lower limit of modal class, $l = 35$

Frequency of modal class, $f_1 = 23$

Frequency of class preceding modal class, $f_0 = 21$

Frequency of class succeeding the modal class, $f_2 = 14$

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

$$= 35 + \left(\frac{23 - 21}{2 \times 23 - 21 - 14} \right) \times 10$$

$$= 35 + \left(\frac{2}{46 - 35} \right) \times 10$$

$$= 35 + \frac{2}{11} \times 10$$

$$= 35 + 1.8$$

$$= 36.8$$

So, the modal age is 36.8 years which means maximum patients admitted to the hospital are of age 36.8 years.

Mean age is 35.37 and average age of the patients admitted is 35.37 years.

Q2. The following data gives information on the observed lifetimes (in hours) of 225 electric 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.

Difficulty Level:

Moderate

Known:

The observed lifetimes (in hours) of 225 electric components.

Unknown:

The modal lifetimes of the components.

Reasoning:

Modal Class is the class with highest frequency

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

Where,

Class size, h

Lower limit of modal class, l

Frequency of modal class, f_1

Frequency of class preceding modal class, f_0

Frequency of class succeeding the modal class, f_2

Solution:

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

From the table, it can be observed that the maximum class frequency is 61, belonging to class interval 60 − 80

Therefore, Modal class = 60 − 80

Class size, $h = 20$

Lower limit of modal class, $l = 60$

Frequency of modal class, $f_1 = 61$

Frequency of class preceding modal class, $f_0 = 52$

Frequency of class succeeding the modal class, $f_2 = 38$

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

$$= 60 + \left(\frac{61 - 52}{2 \times 61 - 52 - 38}\right) \times 20$$

$$= 60 + \left(\frac{9}{122 - 90}\right) \times 20$$

$$= 60 + \frac{9}{32} \times 20$$

$$= 60 + 5.625$$

$$= 65.625$$

Hence, the modal lifetimes of the components are 65.625 hours.

Q3. 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 ₹)	**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

Difficulty Level:

Moderate

Known:

The total monthly household expenditure of 200 families of a village.

Unknown:

The modal and mean monthly expenditure of the families.

Reasoning:

We will find the mean by step-deviation method.

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i}\right) \times h$$

Modal Class is the class with highest frequency

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

Where,

Class size, h

Lower limit of modal class, l

Frequency of modal class, f_1

Frequency of class preceding modal class, f_0

Frequency of class succeeding the modal class, f_2

Solution:

To find mean, we know that,

Class mark, $x_i = \dfrac{\text{Upper class limit } + \text{ Lower class limit}}{2}$

Class size, $h = 500$

Taking assumed mean, $a = 2750$

Expenditure (in ₹)	**Number of families** f_i	x_i	$d_i = x_i - a$	$u_i = \dfrac{d_i}{h}$	$f_i u_i$
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
	$\sum f_i = 200$				$\sum f_i u_i = -35$

From the table, we obtain

$\sum f_i = 200$

$\sum f_i u_i = -35$

Mean, $\bar{x} = a + \left(\dfrac{\sum f_i u_i}{\sum f_i} \right) \times h$

$= 2750 + \left(\dfrac{-35}{200} \right) \times 500$

$= 2750 - \dfrac{175}{2}$

$= 2750 - 87.5$

$= 2662.5$

To find mode

Expenditure (in ₹)	**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

From the table, it can be observed that the maximum class frequency is 40, belonging to class interval 1500 − 2000
Therefore, Modal class = 1500 − 2000

Class size, $h = 500$
Lower limit of modal class, $l = 1500$
Frequency of modal class, $f_1 = 40$
Frequency of class preceding modal class, $f_0 = 24$
Frequency of class succeeding the modal class, $f_2 = 33$

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

$$= 1500 + \left(\frac{40 - 24}{2 \times 40 - 24 - 33}\right) \times 500$$

$$= 1500 + \left(\frac{16}{80 - 57}\right) \times 500$$

$$= 1500 + \frac{16}{23} \times 500$$

$$= 1500 + 347.83$$

$$= 1847.83$$

The modal monthly expenditure of the families is ₹ 1847.83
and the mean monthly expenditure of the families is ₹ 2662.50

Q4. The following data gives the state- wise teacher- student ratio in higher secondary schools of India. Find the mode and mean of the data and interpret the two.

Number 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

Difficulty Level:
Moderate

Known:

The state- wise teacher- student ratio in higher secondary schools of India.

Unknown:

The mode and mean of the data and their interpretation.

Reasoning:

We will find the mean by direct method.

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

Modal Class is the class with highest frequency

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

Where,

Class size, h

Lower limit of modal class, l

Frequency of modal class, f_1

Frequency of class preceding modal class, f_0

Frequency of class succeeding the modal class, f_2

Solution:

To find mean

We know that,

$$\text{Class mark, } x_i = \frac{\text{Upper class limit } + \text{ Lower class limit}}{2}$$

Class size, $h = 500$

Taking assumed mean, $a = 2750$

Number of students per teacher	Number of states / U.T. f_i	x_i	$f_i x_i$
15 – 20	3	17.5	52.5
20 – 25	8	22.5	180
25 – 30	9	27.5	247.5
30 – 35	10	32.5	325
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
	$\sum f_i = 35$		$\sum f_i x_i = 1024$

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

$$= \frac{1024}{35}$$

$$= 29.26$$

To find mode

Number 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

From the table, it can be observed that the maximum class frequency is 10, belonging to class interval 30 − 35

Therefore, Modal class = 30 − 35
Class size, $h = 5$
Lower limit of modal class, $l = 30$

Frequency of modal class, $f_1 = 10$
Frequency of class preceding modal class, $f_0 = 9$
Frequency of class succeeding the modal class, $f_2 = 3$

$$\text{Mode } = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times h$$
$$= 30 + \left(\frac{10-9}{2\times10-9-3}\right)\times 5$$
$$= 30 + \left(\frac{1}{20-12}\right)\times 5$$
$$= 30 + \frac{5}{8}$$
$$= 30 + 0.625$$
$$= 30.625$$
$$= 30.6$$

The modal teacher- student ratio is 30.6 and mean teacher- student ratio is 29.26.
Most states/U.T. have a teacher- student ratio of 30.6 and on an average the ratio is 29.26

Q5. The given distribution shows the number of runs scored by some top batsman of the world in one- day international cricket matches.

Runs scored	Number of batsmen
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.

Difficulty Level:
Moderate

Known:
The number of runs scored by some top batsman of the world in one- day international cricket matches.

Unknown:
The mode of the data.

Reasoning:
Modal Class is the class with highest frequency

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

Where,
Class size, h
Lower limit of modal class, l
Frequency of modal class, f_1
Frequency of class preceding modal class, f_0
Frequency of class succeeding the modal class, f_2

Solution:
From the table, it can be observed that the maximum class frequency is 18, belonging to class interval $4000 - 5000$
Therefore, Modal class = $4000 - 5000$

Class size, $h = 1000$
Lower limit of modal class, $l = 4000$
Frequency of modal class, $f_1 = 18$
Frequency of class preceding modal class, $f_0 = 4$
Frequency of class succeeding the modal class, $f_2 = 9$

$$\begin{aligned}
\text{Mode } &= l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times h \\
&= 4000 + \left(\frac{18-4}{2\times18-4-9}\right)\times1000 \\
&= 4000 + \left(\frac{14}{36-13}\right)\times1000 \\
&= 4000 + \frac{14}{23}\times1000 \\
&= 4000 + 608.695 \\
&= 4608.695 \\
&= 4608.7
\end{aligned}$$

Hence the mode is 4608.7

Q6. A student noted the number of cars passing through a spot on a road for 100 periods each of 3 minutes and summarised it in the table given below. Find mode of the data.

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

Difficulty Level:
Moderate

Known:
The number of cars passing through a spot on a road for 100 periods each of 3 minutes.

Unknown:
The mode of the data.

Reasoning:
Modal Class is the class with highest frequency

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

Where,
Class size, h
Lower limit of modal class, l
Frequency of modal class, f_1

Frequency of class preceding modal class, f_0

Frequency of class succeeding the modal class, f_2

Solution:

From the table, it can be observed that the maximum class frequency is 20, belonging to class interval 40 − 50

Therefore, Modal class = 40 − 50

Class size, $h = 10$

Lower limit of modal class, $l = 40$

Frequency of modal class, $f_1 = 20$

Frequency of class preceding modal class, $f_0 = 12$

Frequency of class succeeding the modal class, $f_2 = 11$

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

$$= 40 + \left(\frac{20 - 12}{2 \times 20 - 12 - 11}\right) \times 10$$

$$= 40 + \left(\frac{8}{40 - 23}\right) \times 10$$

$$= 40 + \frac{8}{17} \times 10$$

$$= 40 + 4.705$$

$$= 40.705$$

$$= 40.7$$

Hence the mode is 40.7

Chapter 14: Statistics

Exercise 14.3

Q1. The following frequency distribution gives the monthly consumption of electricity of 68 consumers of a locality. Find the median, mean and mode of the data and compare them.

Monthly consumption (in units)	Number of consumers
65 – 85	4
85 – 105	5
105 – 125	13
125 – 145	20
145 – 165	14
165 – 185	8
185 – 205	4

Difficulty Level:
Medium

Known:
The frequency distribution of the monthly consumption of electricity of 68 consumers of a locality

Unknown:
The median, mean and mode of the data and the comparison between them.

Reasoning:
We will find the mean by step-deviation method.

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i} \right) \times h$$

Modal Class is the class with highest frequency

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

Where,
Class size, h
Lower limit of modal class, l
Frequency of modal class, f_1

Frequency of class preceding modal class, f_0
Frequency of class succeeding the modal class, f_2

Median Class is the class having Cumulative frequency(cf) just greater than $\frac{n}{2}$

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

Class size, h
Number of observations, n
Lower limit of median class, l
Frequency of median class, f
Cumulative frequency of class preceding median class, cf

Solution:
To find mean, the following relation is used.

$$\text{Class mark, } x_i = \frac{\text{Upper class limit } + \text{ Lower class limit}}{2}$$

Class size, $h = 20$

Taking assumed mean, $a = 135$
d_i, u_i and f_iu_i are calculated according to step-deviation method as follows:

Monthly consumption (in units)	**Number of consumers** f_i	**Class mark** x_i	$d_i = x_i - a$	$u_i = \frac{d_i}{h}$	f_iu_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
Total	**68**				**7**

From the table, we obtain

$$\sum f_i = 68$$

$$\sum f_i u_i = 7$$

Mean, $\overline{\text{X}} = a + \left(\frac{\sum f_i u_i}{\sum f_i}\right) \times h$

$$= 135 + \left(\frac{7}{68}\right) \times 20$$

$$= 135 + \frac{140}{68}$$

$$= 135 + 2.05$$

$$= 137.05$$

To find mode

Monthly consumption (in units)	**Number of consumers**
65 – 85	4
85 – 105	5
105 – 125	13
125 – 145	20
145 – 165	14
165 – 185	8
185 – 205	4

From the table, it can be observed that the maximum class frequency is 20, belonging to class interval 125 − 145.

Class size, h = 20

Modal class = 125 − 145

Lower limit of modal class, l = 125

Frequency of modal class, f_1 = 20

Frequency of class preceding modal class, f_0 = 13

Frequency of class succeeding the modal class, f_2 = 14

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

$$= 125 + \left(\frac{20 - 13}{2 \times 20 - 13 - 14}\right) \times 20$$

$$= 125 + \left(\frac{7}{40 - 27}\right) \times 20$$

$$= 125 + \frac{7}{13} \times 20$$

$$= 125 + \frac{140}{13}$$

$$= 125 + 10.76$$

$$= 135.76$$

To find the median of the given data, cumulative frequency is calculated as follows

Monthly consumption (in units)	**Number of consumers** *f*	**Cumulative frequency** *cf*
65 – 85	4	4
85 – 105	5	4 + 5 = 9
105 – 125	13	9+13 = 22
125 – 145	20	22 + 20 = 42
145 – 165	14	42 + 14 = 56
165 – 185	8	56 + 8 = 64
185 – 205	4	64 + 4 = 68
***n* = 68**		

From the table, we obtain

$$n = 68 \quad \Rightarrow \frac{n}{2} = 34$$

Cumulative frequency(*cf*) just greater than $\frac{n}{2}$ is 42, belonging to class-interval 125 − 145.

Therefore, median class = 125 – 145

Class size, h = 20

Lower limit of median class, l = 125

Frequency of median class, f = 20

Cumulative frequency of class preceding median class, cf = 22

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

$$= 125 + \left(\frac{34 - 22}{20} \right) \times 20$$

$$= 125 + \frac{12}{20} \times 20$$

$$= 125 + 12$$

$$= 137$$

Therefore, median, mode, mean of the given data is 137, 135.76, and 137.05 respectively. The three measures are approximately the same in this case.

Q2. If the median of the distribution given below is 28.5, find the values of *x* and *y*.

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

Difficulty Level:
Medium

Known:
The median of the distribution is 28.5

Unknown:
The values of *x* and *y*

Reasoning:

Median Class is the class having Cumulative frequency(*cf*) just greater than $\frac{n}{2}$

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

Class size, h

Number of observations, n

Lower limit of median class, l

Frequency of median class, f

Cumulative frequency of class preceding median class, cf

Solution:
The cumulative frequency for the given data is calculated as follows.

Class Interval	Frequency	Cumulative frequency
0 – 10	5	5
10 – 20	x	5 + x
20 – 30	20	25 + x
30 – 40	15	40 + x
40 – 50	y	40 + x + y
50 – 60	5	45 + x + y
***n* = 60**		

From the table, it can be observed that

$$n = 60 \quad \Rightarrow \frac{n}{2} = 30$$

$$45 + x + y = 60$$

$$x + y = 15 \qquad \text{(i)}$$

Median of the data is given as 28.5 which lies in interval 20 − 30.

Therefore, median class = 20 – 30

Class size, h = 10

Lower limit of median class, l = 20

Frequency of median class, f = 20

Cumulative frequency of class preceding the median class, cf = 5 + x

Median $= l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$

$$28.5 = 20 + \left(\frac{30 - (5 + x)}{20} \right) \times 10$$
$$28.5 - 20 = \left(\frac{25 - x}{20} \right) \times 10$$
$$8.5 = \frac{25 - x}{2}$$
$$25 - x = 8.5 \times 2$$
$$x = 25 - 17$$
$$x = 8$$

Putting $x = 8$ in equation (i)

$$8 + y = 15$$
$$y = 7$$

Hence, the values of x and y are 8 and 7 respectively.

Q3. A life insurance agent found the following data for distribution of ages of 100 policy holders. Calculate the median age, if policies are given only to persons having age 18 years onwards but less than 60 year.

Age (in years)	Number of policy holders
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

Difficulty Level:
Medium

Known:

The data for distribution of ages of 100 policy holders. The policies are given only to persons having age 18 years onwards but less than 60 years.

Unknown:

The median age.

Reasoning:

Here, class width is not the same. There is no requirement of adjusting the frequencies according to class intervals. The given frequency table is of less than type represented with upper class limits.

Median Class is the class having Cumulative frequency(cf) just greater than $\frac{n}{2}$

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

Class size, h
Number of observations, n
Lower limit of median class, l
Frequency of median class, f
Cumulative frequency of class preceding median class, cf

Solution:

Class intervals with their respective cumulative frequency can be defined as below.

Age (in years)	Cumulative frequency cf	Number of policy holders f
18 – 20	2	2
20 – 25	6	6 – 2 = 4
25 – 30	24	24 – 6 = 18
30 – 35	45	45 – 24 = 21
35 – 40	78	78 – 45 = 33
40 – 45	89	89 – 78 = 11
45 – 50	92	92 – 89 = 3
50 – 55	98	98 – 92 = 6
55 – 60	100	100 – 98 = 2

From the table, it can be observed that

$n = 100 \quad \Rightarrow \frac{n}{2} = 50$

Cumulative frequency (cf) just greater than 50 is 78, belonging to class-interval 35 − 40.

Therefore, median class = 35 – 40
Class size, h = 5
Lower limit of median class, l = 35
Frequency of median class, f = 33
Cumulative frequency of class preceding median class, cf = 45

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

$$= 35 + \left(\frac{50 - 45}{33} \right) \times 5$$

$$= 35 + \frac{5}{33} \times 5$$

$$= 35 + \frac{25}{33}$$

$$= 35 + 0.76$$

$$= 35.76$$

Therefore, median age is 35.76 years.

Q4. The lengths of 40 leaves of a plant are measured correct to the nearest millimetre, and the data obtained is represented 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.)

Difficulty Level:
Medium

Known:

The lengths of 40 leaves of a plant are measured in millimetre.

Unknown:

The median length of the leaves.

Reasoning:

Median Class is the class having Cumulative frequency(*cf*) just greater than $\frac{n}{2}$

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

Class size, h

Number of observations, n

Lower limit of median class, l

Frequency of median class, f

Cumulative frequency of class preceding median class, cf

Solution:

Length (in mm)	Number of leaves *f*	Cumulative frequency *cf*
117.5 – 126.5	3	3
126.5 – 135.5	5	3 + 5 = 8
135.5 – 144.5	9	8 + 9 = 17
144.5 – 153.5	12	17 + 12 = 29
153.5 – 162.5	5	29 + 5 = 34
162.5 – 171.5	4	34 + 4 = 38
171.5 – 180.5	2	38 + 2 = 40
***n* = 40**		

From the table, it can be observed that

$n = 40 \quad \Rightarrow \frac{n}{2} = 20$

Cumulative frequency (*cf*) just greater than 20 is 29, belonging to class 144.5 – 153.5

Therefore, median class = 144.5 – 153.5

Class size, $h = 9$

Lower limit of median class, $l = 144.5$

Frequency of median class, $f = 12$

Cumulative frequency of class preceding median class, cf = 17

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

$$= 144.5 + \left(\frac{20 - 17}{12} \right) \times 9$$

$$= 144.5 + \frac{3}{12} \times 5$$

$$= 144.5 + \frac{5}{4}$$

$$= 144.5 + 1.25$$

$$= 145.75$$

Therefore, median length of leaves is 146.75 mm.

Q5. The following table gives the distribution of the life time 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.

Difficulty Level:
Medium

Known:
The lifetime of 400 neon lamps.

Unknown:
The median lifetime of a lamp.

Reasoning:

Median Class is the class having Cumulative frequency(*cf*) just greater than $\frac{n}{2}$

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

Class size, h

Number of observations, n

Lower limit of median class, l

Frequency of median class, f

Cumulative frequency of class preceding median class, cf

Solution:

Lifetime (in hours)	**Number of lamps** f	**Cumulative frequency** cf
1500 – 2000	14	14
2000 – 2500	56	14 + 56 = 70
2500 – 3000	60	70 + 60 = 130
3000 – 3500	86	130 + 86 = 216
3500 – 4000	74	216 + 74 = 290
4000 – 4500	62	290 + 62 = 352
4500 – 5000	48	352 + 48 = 400
***n* = 400**		

From the table, it can be observed that

$n = 400 \quad \Rightarrow \frac{n}{2} = 200$

Cumulative frequency (*cf*) just greater than 200 is 216, belonging to class 3000 – 3500.
Therefore, median class = 3000 – 3500

Class size, h = 500

Lower limit of median class, l = 3000

Frequency of median class, f = 86

Cumulative frequency of class preceding median class, cf = 130

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

$$= 3000 + \left(\frac{200 - 130}{86} \right) \times 500$$

$$= 3000 + \frac{70}{86} \times 500$$

$$= 3000 + \frac{17500}{43}$$

$$= 3000 + 406.98$$

$$= 3406.98$$

Therefore, median lifetime of lamps is 3406.98 hours.

Q6. 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.

Difficulty Level:
Hard

Known:
The frequency distribution of the number of letters in the English alphabets for 100 surnames.

Unknown:
The median and mean number of letters in the surnames and the modal size of the surnames.

Reasoning:
We will find the mean by step-deviation method.

$$\text{Mean, } \bar{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i} \right) \times h$$

Modal Class is the class with highest frequency

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

Where,
Class size, h

Lower limit of modal class, l

Frequency of modal class, f_1

Frequency of class preceding modal class, f_0

Frequency of class succeeding the modal class, f_2

Median Class is the class having Cumulative frequency(cf) just greater than $\frac{n}{2}$

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

Class size, h

Number of observations, n

Lower limit of median class, l

Frequency of median class, f

Cumulative frequency of class preceding median class, cf

Solution:
To find the median

Number of letters	**Number of surnames** ***f***	**Cumulative frequency** ***cf***
1 – 4	6	6
4 – 7	30	6 + 30 = 36
7 – 10	40	36 + 40 = 76
10 – 13	16	76 + 16 = 92
13 – 16	4	92 + 4 = 96
16 – 19	4	96 + 4 = 100
***n* = 100**		

From the table, it can be observed that

$n = 100 \quad \Rightarrow \frac{n}{2} = 50$

Cumulative frequency (cf) just greater than 50 is 76, belonging to class 7 – 10.

Therefore, median class = 7 – 10

Class size, h = 3

Lower limit of median class, l = 7

Frequency of median class, f = 40

Cumulative frequency of class preceding median class, cf = 36

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

$$= 7 + \left(\frac{50 - 36}{40} \right) \times 3$$

$$= 7 + \frac{14}{40} \times 3$$

$$= 7 + \frac{21}{20}$$

$$= 7 + 1.05$$

$$= 8.05$$

To find the mean

Class mark, $x_i = \frac{\text{Upper class limit } + \text{ Lower class limit}}{2}$

Taking assumed mean, $\boldsymbol{a}$ = 11.5

Number of letters	**Number of surnames** f_i	**Class mark** x_i	$d_i = x_i - a$	$u_i = \frac{d_i}{h}$	$f_i u_i$
1 – 4	6	2.5	– 9	– 3	– 18
4 – 7	30	5.5	– 6	– 2	– 60
7 – 10	40	8.5	– 3	– 1	– 40
10 – 13	16	11.5	0	0	0
13 – 16	4	14.5	3	1	4
16 – 19	4	17.5	6	2	8
Total	**100**				– 106

From the table, we obtain

$$\sum f_i = 100$$
$$\sum f_i u_i = -106$$
Class size, $h = 3$

Mean, $\overline{X} = a + \left(\frac{\sum f_i u_i}{\sum f_i}\right) \times h$

$$= 11.5 + \left(\frac{-106}{100}\right) \times 3$$

$$= 11.5 - \frac{318}{100}$$

$$= 11.5 - 3.18$$

$$= 8.32$$

To find mode

Number of letters	Number of surnames
1 – 4	6
4 – 7	30
7 – 10	40
10 – 13	16
13 – 16	4
16 – 19	4
	***n* = 100**

From the table, it can be observed that the maximum class frequency is 40, belonging to class interval 7 − 10.

Class size, $h = 3$

Modal class = 7 − 10
Lower limit of modal class, $l = 7$
Frequency of modal class, $f_1 = 40$
Frequency of class preceding modal class, $f_0 = 30$
Frequency of class succeeding the modal class, $f_2 = 16$

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

$$= 7 + \left(\frac{40 - 30}{2 \times 40 - 30 - 16}\right) \times 3$$

$$= 7 + \left(\frac{10}{34}\right) \times 3$$

$$= 7 + \frac{15}{17}$$

$$= 7 + 0.88$$

$$= 7.88$$

Therefore, median and mean number of letters in surnames is 8.05 and 8.32 respectively and modal size of surnames is 7.88.

Q7. 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

Difficulty Level:
Medium

Known:
The weights of 30 students of a class.

Unknown:
The median weight of the students.

Reasoning:

Median Class is the class having Cumulative frequency(*cf*) just greater than $\frac{n}{2}$

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

Class size, h
Number of observations, n
Lower limit of median class, l
Frequency of median class, f
Cumulative frequency of class preceding median class, cf

Solution:

Weight (in kg)	**Number of students** f	**Cumulative frequency** cf
40 – 45	2	2
45 – 50	3	2 + 3 = 5
50 – 55	8	5 + 8 = 13
55 – 60	6	13 + 6 = 19
60 – 65	6	19 + 6 = 25
65 – 70	3	25 + 3 = 28
70 – 75	2	28 + 2 = 30
***n* = 30**		

From the table, it can be observed that

$$n = 30 \quad \Rightarrow \frac{n}{2} = 15$$

Cumulative frequency (*cf*) just greater than 15 is 19, belonging to class 55 – 60.

Therefore, median class = 55 – 60

Class size, h = 5

Lower limit of median class, l = 55

Frequency of median class, f = 6

Cumulative frequency of class preceding median class, cf = 13

$$\text{Median } = l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

$$= 55 + \left(\frac{15 - 13}{6} \right) \times 5$$

$$= 55 + \frac{2}{6} \times 5$$

$$= 55 + \frac{5}{3}$$

$$= 55 + 1.67$$

$$= 56.67$$

Therefore, median weight is 56.67 kg.

Chapter 14: Statistics

Exercise 14.4

Q1. The following distribution gives the daily income of 50 workers of a factory.

Daily income (in ₹)	100 − 120	120 − 140	140 − 160	160 – 180	180 – 200
Number of workers	12	14	8	6	10

Convert the distribution above to a less than type cumulative frequency distribution and draw its ogive.

Difficulty Level:
Medium

Known:
The daily income of 50 workers of a factory.

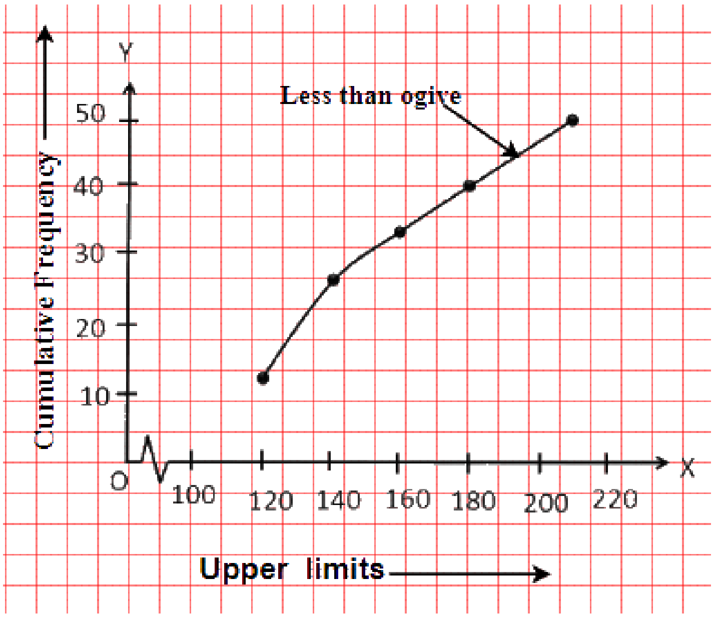
Unknown:
The less than type cumulative frequency distribution and its ogive.

Reasoning:
The representation of cumulative frequency distribution graphically is known as a cumulative frequency curve or an ogive.

Solution:
The frequency distribution table of less than type is as follows:

Daily income (in ₹) (Upper class Limits)	**Cumulative Frequency**
Less than 120	12
Less than 140	$12+14=26$
Less than 160	$26+8=34$
Less than 180	$34+6=40$
Less than 200	$40+10=50$

Taking upper class limits of class intervals on *x*-axis and their respective frequencies on *y*-axis, its ogive can be drawn as follows:



Q2. During the medical check-up of 35 students of a class, their weights were recorded as follows:

Weight in (Kg)	Number of students
Less than 38	0
Less than 40	3
Less than 42	5
Less than 44	9
Less than 46	14
Less than 48	28
Less than 50	32
Less than 52	35

Draw a less than type ogive for the given data. Hence obtain the median weight from the graph and verify the result by using the formula.

Difficulty Level:
Medium

Known:
The weight of 35 students of a class.

Unknown:
The less than type ogive and median weight.

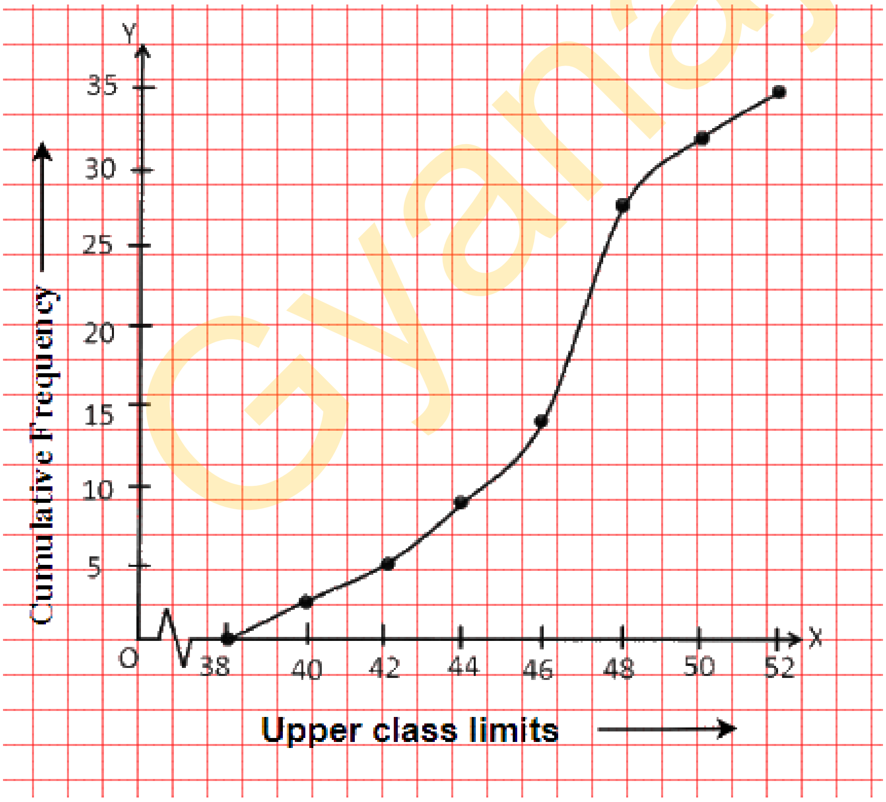
Reasoning:
The representation of cumulative frequency distribution graphically is known as a cumulative frequency curve or an ogive.

Solution:

The given cumulative frequency distributions of less than type are:

Weight in (Kg)	Number of students
Less than 38	0
Less than 40	3
Less than 42	5
Less than 44	9
Less than 46	14
Less than 48	28
Less than 50	32
Less than 52	35

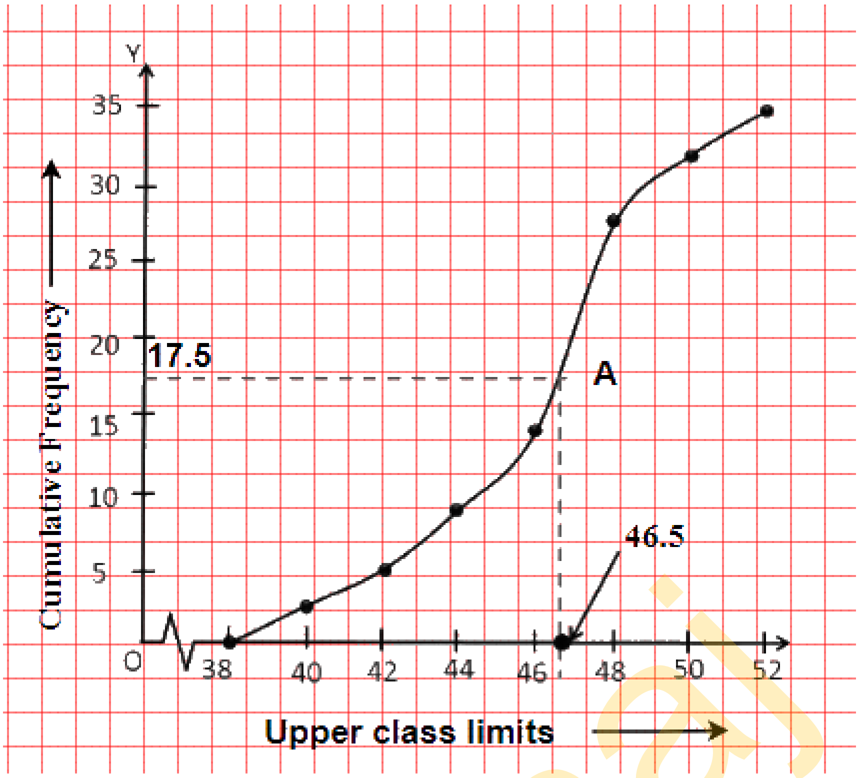
Taking upper class limits on x-axis and their respective cumulative frequencies on y-axis, its ogive can be drawn as follows.



Here, $n = 35 \quad \Rightarrow \frac{n}{2} = 17.5$

Mark the point 'A' whose ordinate is 17.5 and its x-coordinate is 46.5.

Therefore, median of this data is 46.5.



It can be observed that the difference between two consecutive upper-class limits is 2.

The class marks with their respective frequencies are obtained as below

Weight (in Kg)	Frequency	Cumulative Frequency
Less than 38	0	0
38-40	$3-0=3$	3
40-42	$5-3=2$	5
42-44	$9-5=4$	9
44-46	$14-9=5$	14
46-48	$28-14=14$	28
48-50	$32-28=4$	32
50-52	$35-32=3$	35
Total	***n* = 35**	

Cumulative frequency (*cf*) just greater than 17.5 is 28, belonging to class 46 – 48.

Therefore, median class = 46 – 48

Class size, *h* = 2

Lower limit of median class, *l* = 46

Frequency of median class, *f* = 14

Cumulative frequency of class preceding median class, *cf* = 14

$$\text{Median} = l + \left(\frac{\frac{n}{2} - cf}{f}\right) \times h$$

$$= 46 + \left(\frac{17.5 - 14}{14}\right) \times 2$$

$$= 46 + \frac{3.5}{7}$$

$$= 46 + 0.5$$

$$= 46.5$$

Therefore, median of this data is 46.5.
Hence, the value of median is verified.

Q3. The following table gives production yield per hectare of wheat of 100 farms of a village.

Production yield (in kg/ha)	50 – 55	55 – 60	60 – 65	65 – 70	70 – 75	75 – 80
Number of farms	2	8	12	24	38	16

Change the distribution to a more than type distribution and draw ogive.

Difficulty Level:
Medium

Known:
The production yield per hectare of wheat of 100 farms of a village.

Unknown:
The more than type distribution and its ogive.

Reasoning:
The representation of cumulative frequency distribution graphically is known as a cumulative frequency curve or an ogive.

Solution:
The cumulative frequency distribution of more than type can be obtained as follows.

Production Yield (in kg/ha) (Lower class limits)	**Cumulative Frequency**
More than or equal to 50	100
More than or equal to 55	$100 - 2 = 98$
More than or equal to 60	$98 - 8 = 90$
More than or equal to 65	$90 - 12 = 78$
More than or equal to 70	$78 - 24 = 54$
More than or equal to 75	$54 - 38 = 16$

Taking the lower-class limits on *x*-axis and their respective cumulative frequencies on *y*-axis, its ogive can be obtained as follows.