

Introduction to Statistics: Statistics is a branch of mathematics that deals with the data collection, data analysis and interpretation of collected data. From collection, analysis and interpretation, we reach at meaningful conclusions.

Data is the group of information that represent the qualitative or quantitative attributes of any variable. For example, marks obtained by students of Diploma students in Mathematics.

Measures of Central Tendency: A measure of central tendency is a summary statistic that represents the central position of the given data. In Statistics, the three most common measures of central tendency are: (1) Mean, (2) Median, (3) Mode. All the three measures calculate the location of central position using different methods. Also choosing the best measures of central tendency depends on the type of given data.

(1) Mean: To calculate mean, we add the values of all terms and then divide the sum by number of terms.

i.e., if $x_1, x_2, x_3, \dots, x_N$ is the given data, then mean of this data is $\frac{x_1+x_2+x_3+\dots+x_N}{N}$ and it is represented by \bar{x} . Here N is the number of terms.

e.g.

Let the given data is 9, 12, 7, 15, 17

Mean of the above data is $\bar{x} = \frac{9+12+7+15+17}{5} = \frac{60}{5} = 12$.

Note: (i) Mean is also known as “Arithmetic Mean” or Average.

(ii) The observations $x_1, x_2, x_3, \dots, x_N$ is known as individual series.

(iii) The method written above is known as “Direct Method”.

For Frequency Distribution: If the following is given frequency distribution

x	x_1	x_2	x_3	...	x_n
f	f_1	f_2	f_3	...	f_n

then, mean of this data is given by

$$\bar{x} = \frac{f_1x_1+f_2x_2+f_3x_3+\dots+f_nx_n}{f_1+f_2+f_3+\dots+f_n} = \frac{\Sigma f_i x_i}{\Sigma f_i}$$

or $\bar{x} = \frac{\Sigma f_i x_i}{N}$, where $N = \Sigma f_i$.

Q.1 Find the mean of the following data:

14, 25, 18, 23, 20, 30, 15, 24, 11, 17

Ans. Given data is 14, 25, 18, 23, 20, 30, 15, 24, 11, 17

Here, $N = \text{no. of terms} = 10$

Therefore, Mean is

$$\bar{x} = \frac{14+25+18+23+20+30+15+24+11+17}{10} = \frac{197}{10} = 19.7$$

Q.2 Find the Arithmetic Mean(A.M.) of the following data:

9, 7, 6, 7, 5, 4, 6, 4

Ans. Given data is 9, 7, 6, 7, 5, 4, 6, 4

Here, $N = \text{no. of terms} = 8$

Therefore, Arithmetic Mean(A.M.) is

$$\bar{x} = \frac{9+7+6+7+5+4+6+4}{8} = \frac{48}{8} = 6$$

Q.3 Calculate the Arithmetic Mean(A.M.) for the data 150, 130, 120, 160, 140 .

Ans. Given data is 150, 130, 120, 160, 140

Here, $N = \text{no. of terms} = 5$

Therefore, Arithmetic Mean(A.M.) is

$$\bar{x} = \frac{150+130+120+160+140}{5} = \frac{700}{5} = 140$$

Q.4 Calculate the Mean for the following frequency distribution

x_i	2	4	6	8	10
f_i	3	1	2	1	3

Ans. We know that, mean for frequency distribution is given by $\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$.

x_i	f_i	$f_i x_i$
2	3	6
4	1	4
6	2	12
8	1	8
10	3	30

	$\Sigma f_i = 10$	$\Sigma f_i x_i = 60$
--	-------------------	-----------------------

Therefore, Mean = $\frac{60}{10} = 6$

Q.5 Calculate the Arithmetic Mean(A.M.) for the following frequency distribution

x_i	1	2	3	4	5
f_i	6	4	3	4	3

Ans. We know that, Arithmetic mean for frequency distribution is given by $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i}$.

x_i	f_i	$f_i x_i$
1	6	6
2	4	8
3	3	9
4	4	16
5	3	15
	$\Sigma f_i = 20$	$\Sigma f_i x_i = 54$

Therefore, Arithmetic Mean = $\frac{54}{20} = 2.7$

Q.6 Calculate the Arithmetic Mean(A.M.) for the following frequency distribution

x_i	1	3	5	7	9	11	13	15
f_i	1	2	3	4	4	3	2	1

Ans.

We know that, Arithmetic mean for frequency distribution is given by $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i}$.

x_i	f_i	$f_i x_i$
1	1	1
3	2	6
5	3	15
7	4	28
9	4	36
11	3	33
13	2	26
15	1	15
	$\Sigma f_i = 20$	$\Sigma f_i x_i = 160$

Therefore, Arithmetic Mean = $\frac{160}{20} = 8$

Exercise:

Q.1 Calculate the Arithmetic Mean for the following data:

- (i) 8, 12, 5, 19, 13
- (ii) 25, 75, 90, 60, 50, 35, 95, 20, 80, 40
- (iii) 120, 160, 220, 210, 190
- (iv) 23, 34, 25, 28, 39, 36, 24, 35
- (v) 62, 75, 79, 68, 81, 93, 89
- (vi)

x_i	13	15	17	19	21
f_i	1	3	1	2	3

(vii)

x_i	10	12	14	16	18	20	22
f_i	3	2	4	5	3	1	2

(viii)

x_i	1	2	3	4	5	6	7	8
f_i	6	4	4	3	3	3	4	3

(ix)

x_i	1	2	3	4	5	6
f_i	2	4	6	6	4	2

(x)

x_i	3	6	9	12	15	18	21	24
f_i	1	2	3	4	4	3	2	1

Mean for Continuous Frequency Distribution:

If the following is given continuous frequency distribution

Class	<i>class limit 1</i>	<i>class limit 2</i>	<i>class limit 3</i>	...	<i>class limit n</i>
f	f_1	f_2	f_3	...	f_n

then, we have to do following steps:

Step (i) : Find x_i by taking midpoint of the class.

Step (ii) : Then, $\bar{x} = \frac{f_1x_1+f_2x_2+f_3x_3+\dots+f_nx_n}{f_1+f_2+f_3+\dots+f_n} = \frac{\sum f_i x_i}{\sum f_i}$

or $\bar{x} = \frac{\sum f_i x_i}{N}$, where $N = \sum f_i$.

Q.1 Find the Mean for the following continuous frequency distribution:

class	0-10	10-20	20-30	30-40	40-50
f_i	8	4	3	3	2

Ans.

class	x_i	f_i	$f_i x_i$
0-10	5	8	40
10-20	15	4	60
20-30	25	3	75
30-40	35	3	105
40-50	45	2	90
		$\sum f_i = 20$	$\sum f_i x_i = 370$

Here $N = \sum f_i = 20$

Therefore, Mean = $\frac{\sum f_i x_i}{\sum f_i} = \frac{370}{20} = 18.5$

Q.2 Find the Mean for the following frequency distribution:

class	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
f_i	5	4	1	2	3	2	2	3

Ans.

class	x_i	f_i	$f_i x_i$
5-10	7.5	5	37.5
10-15	12.5	4	50
15-20	17.5	1	17.5
20-25	22.5	2	45
25-30	27.5	3	82.5
30-35	32.5	2	65
35-40	37.5	2	75
40-45	42.5	3	127.5
		$\sum f_i = 22$	$\sum f_i x_i = 500$

Here $N = \sum f_i = 22$

$$\text{Therefore, Mean} = \frac{\sum f_i x_i}{\sum f_i} = \frac{500}{22} = 22.727$$

Exercise:

Q.1 Calculate the Mean for the following data:

(i)

Class	0-2	2-4	4-6	6-8	8-10	10-12
f_i	2	4	6	6	4	2

(ii)

Class	5-10	10-15	15-20	20-25	25-30	30-35	35-40
f_i	5	2	7	4	5	6	1

(iii)

Class	10-20	20-30	30-40	40-50	50-60	60-70
f_i	8	5	4	6	7	9

(iv)

Class	0-3	3-6	6-9	9-12	12-15	15-18	18-21	21-24
f_i	6	4	5	8	3	4	6	2

(2) Median: The median value of a set of data is the middle value of the ordered data. If number of ordered data is odd, then median is the mid value of the data and if number of ordered data is even, then median is the mid value of middle two values.

i.e., if $x_1, x_2, x_3, \dots, x_N$ is the ordered data,

then median of this data is $\left(\frac{N+1}{2}\right)^{th}$ term, if N is odd number and

median is $\frac{\left(\frac{N}{2}\right)^{th} \text{ term} + \left(\frac{N}{2} + 1\right)^{th} \text{ term}}{2}$, if N is even number.

e.g.

Let the given data is 7, 9, 12, 15, 17

This data is in ordered form.

Here $N = 5$

Therefore, Median is $\left(\frac{5+1}{2}\right)^{rd} term = 3^{rd} term = 12$.

Note: (i) The word “ordered” means that we have to arrange the data in order (ascending or descending).

(ii) The above method is applicable on individual observations, i.e., $x_1, x_2, x_3, \dots, x_N$.

For Discrete Frequency Distribution: If the following is given frequency distribution

x	x_1	x_2	x_3	...	x_n
f	f_1	f_2	f_3	...	f_n

then, we have to do following steps:

Step (i) : Find the cumulative frequency.

Step (ii) : Find $\frac{N}{2}$, where $N = \Sigma f_i$.

Step (iii) : Note the cumulative frequency just greater than or equal to $\frac{N}{2}$ and the corresponding value of x . This value of x is the required median.

For Grouped or Continuous Frequency Distribution: If the following is given continuous frequency distribution

Class	<i>class limit 1</i>	<i>class limit 2</i>	<i>class limit 3</i>	...	<i>class limit n</i>
f	f_1	f_2	f_3	...	f_n

then, we have to do following steps:

Step (i) : Find the cumulative frequency.

Step (ii) : Find $\frac{N}{2}$, where $N = \Sigma f_i$.

Step (iii) : Note the cumulative frequency just greater than or equal to $\frac{N}{2}$ and the corresponding class. This class is called the median class.

Step (iv) : Median = $l + \left(\frac{\frac{N}{2} - C}{f}\right) \times h$

where l = lower limit of median class, h = width of the median class,

f = frequency of the median class,

C = cumulative frequency of the class preceding the median class

RAVI BANSAL

Q.1 Calculate the median for the following data:

35, 27, 39, 51, 60

Ans. Given data is 35, 27, 39, 51, 60

Rearranging the data in ascending order, we have

27, 35, 39, 51, 60

Here, $N = \text{no. of terms} = 5$, which is odd.

Therefore, median = $\left(\frac{5+1}{2}\right)^{\text{rd}} \text{ term} = 3^{\text{rd}} \text{ term} = 39$

Q.2 Calculate the median for the following data:

5, 7, 11, 15, 19, 23

Ans. Given data is 5, 7, 11, 15, 19, 23, which is in arranged form.

Here $N = \text{no. of terms} = 6$, which is even.

Therefore, median = $\frac{\left(\frac{6}{2}\right)^{\text{rd}} \text{ term} + \left(\frac{6}{2}+1\right)^{\text{th}} \text{ term}}{2} = \frac{3^{\text{rd}} + 4^{\text{th}}}{2} = \frac{11+15}{2} = \frac{26}{2} = 13.$

Q.3 Find the Median for the following frequency distribution:

x_i	1	2	3	4	5	6
f_i	5	3	2	4	2	3

Ans.

x_i	f_i	c.f.
1	5	5
2	3	8
3	2	10
4	4	14
5	2	16
6	3	19
	$\Sigma f_i = 19$	

Here $N = \Sigma f_i = 19$

$$\frac{N}{2} = \frac{19}{2} = 9.5$$

Therefore, Median = 3.

Note: In cumulative frequency, 10 is just greater than 9.5. So, we selected that value according to the rule.

Q.4 Find the Median for the following frequency distribution:

x_i	2	4	6	8	10
f_i	3	5	8	4	2

Ans.

x_i	f_i	$c.f.$
2	3	3
4	5	8
6	8	16
8	4	20
10	2	22
	$\Sigma f_i = 22$	

Here $N = \Sigma f_i = 22$

$$\frac{N}{2} = \frac{22}{2} = 11$$

Therefore, Median = 6.

Note: In cumulative frequency, 16 is just greater than 11. So, we selected that value according to the rule.

Q.5 Find the Median for the following frequency distribution:

x_i	1	2	3	4	5
f_i	6	4	3	4	3

Ans.

x_i	f_i	$c.f.$
1	6	6
2	4	10
3	3	13
4	4	17
5	3	20
	$\Sigma f_i = 20$	

Here $N = \sum f_i = 20$

$$\frac{N}{2} = \frac{20}{2} = 10$$

Therefore, Median = 2.

Q.6 Find the Median for the following continuous frequency distribution:

class	0-10	10-20	20-30	30-40	40-50
f_i	13	14	26	22	15

Ans.

class	f_i	c.f.
0-10	13	13
10-20	14	27
20-30	26	53
30-40	22	75
40-50	15	90
	$\sum f_i = 90$	

Here $N = \sum f_i = 90$

$$\frac{N}{2} = \frac{90}{2} = 45$$

l = lower limit of median class = 20,

h = width of the median class = 10,

f = frequency of the median class = 26,

C = cumulative frequency of the class preceding the median class = 27

$$\begin{aligned} \text{Therefore, Median} &= l + \left(\frac{\frac{N}{2} - C}{f} \right) \times h = 20 + \left(\frac{45 - 27}{26} \right) \times 10 \\ &= 20 + \left(\frac{18}{26} \right) \times 10 = 20 + 6.92 = 26.92 \end{aligned}$$

Note: In cumulative frequency, 53 is just greater than 45. So, we selected the corresponding class as median class.

Q.7 Find the Median for the following frequency distribution:

class	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
--------------	------	-------	-------	-------	-------	-------	-------	-------

f_i	5	4	6	10	8	15	7	3
-------	---	---	---	----	---	----	---	---

Ans.

<i>class</i>	f_i	<i>c. f.</i>
5-10	5	5
10-15	4	9
15-20	6	15
20-25	10	25
25-30	8	33
30-35	15	48
35-40	7	55
40-45	3	58
	$\Sigma f_i = 58$	

Here $N = \Sigma f_i = 58$

$$\frac{N}{2} = \frac{58}{2} = 29$$

$l =$ lower limit of median class = 25,

$h =$ width of the median class = 5,

$f =$ frequency of the median class = 8,

$C =$ cumulative frequency of the class preceding the median class = 25

$$\begin{aligned} \text{Therefore, Median} &= l + \left(\frac{\frac{N}{2} - C}{f} \right) \times h = 25 + \left(\frac{29 - 25}{8} \right) \times 5 \\ &= 25 + \left(\frac{4}{8} \right) \times 5 = 25 + 2.5 = 27.5 \end{aligned}$$

Note: In cumulative frequency, 33 is just greater than 29. So, we selected the corresponding class as median class.

Exercise:

Q.1 Calculate the Median for the following data:

- (i) 8, 12, 5, 19, 13
- (ii) 25, 75, 90, 60, 50, 35, 95, 20, 80, 40
- (iii) 120, 160, 190, 210, 220
- (iv) 23, 34, 25, 28, 39, 36, 24, 35

(v) 62, 75, 79, 68, 81, 93, 89, 70

(vi)

x_i	13	15	17	19	21
f_i	7	4	8	3	5

(vii)

x_i	10	12	14	16	18	20	22
f_i	3	2	4	5	3	1	2

(viii)

x_i	1	2	3	4	5	6	7	8
f_i	6	4	4	3	3	3	4	3

(ix)

Class	0-2	2-4	4-6	6-8	8-10	10-12
f_i	2	4	6	6	4	2

(x)

Class	5-10	10-15	15-20	20-25	25-30	30-35	35-40
f_i	5	2	7	4	5	6	1

(xi)

Class	10-20	20-30	30-40	40-50	50-60	60-70
f_i	8	5	4	6	7	9

(xii)

Class	0-3	3-6	6-9	9-12	12-15	15-18	18-21	21-24
f_i	6	4	5	8	3	4	6	2

(3) Mode: The mode of a set of data is the value that occurs most frequently, i.e. the data entry which has greatest frequency.

e.g.

Let the given data is 7, 9, 7, 9, 12, 9, 15, 9, 17

Here 9 has the greatest frequency. So mode the given data is 9.

Note: (i) A data set may have one mode, more than one mode or no mode. If all the entries in a data set are different, then there is no mode.

(ii) The above method is applicable on both: "Individual Observations" and "Discrete Frequency Distribution".

For Grouped or Continuous Frequency Distribution: To find mode for continuous frequency distribution, we have to do the following steps:

Step (i) : Determine the class of maximum frequency. This class is called modal class.

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

where l = lower limit of the modal class, h = width of the modal class,

f = frequency of the modal class,

f_1 = frequency of the class preceding the modal class,

f_2 = frequency of the class following the modal class.

Q.1 Find the mode for the following data:

22, 24, 21, 23, 25, 24, 21, 25, 23, 21, 30.

Ans. Given data is 22, 24, 21, 23, 25, 24, 21, 25, 23, 21, 30

Therefore, mode = 21 as 21 has maximum frequency.

Q.2 Find the mode for the following data:

30, 40, 34, 42, 56, 34, 44, 55, 42, 34.

Ans. Given data is 30, 40, 34, 42, 56, 34, 44, 55, 42, 34

Therefore, mode = 34 as 34 has maximum frequency.

Q.3 Find the Mode for the following data:

x_i	1	2	3	4	5	6
f_i	5	3	2	7	4	1

Ans. Given data is

x_i	f_i
-------	-------

1	5
2	3
3	2
4	7
5	4
6	1

Here the maximum frequency is 7 and 4 is its corresponding data entry.
Therefore, mode of the given data is 4.

Q.4 Find the Mode for the following frequency distribution:

x_i	3	6	9	12	15	18
f_i	2	3	6	4	5	3

Ans. Given data is

x_i	f_i
3	2
6	3
9	6
12	4
15	5
18	3

Here the maximum frequency is 6 and 9 is its corresponding data entry.
Therefore, mode of the given data is 9.

Q.5 Find the Mode for the following frequency distribution:

x_i	2	4	6	8	10
f_i	3	5	8	4	2

Ans. Given data is

x_i	f_i
2	3
4	5
6	8
8	4
10	2

Here the maximum frequency is 8 and 6 is its corresponding data entry.
Therefore, mode of the given data is 6.

Q.6 Find the Mode for the following continuous frequency distribution:

<i>class</i>	0-10	10-20	20-30	30-40	40-50
<i>f_i</i>	13	14	26	22	15

Ans. Given data is

<i>class</i>	<i>f_i</i>
0-10	13
10-20	14
20-30	26
30-40	22
40-50	15

Here l = lower limit of the modal class = 20,

h = width of the modal class = 10,

f = frequency of the modal class = 26,

f_1 = frequency of the class preceding the modal class = 14,

f_2 = frequency of the class following the modal class = 22

$$\begin{aligned} \text{Therefore, Mode} &= l + \left(\frac{f - f_1}{2f - f_1 - f_2} \right) \times h = 20 + \left(\frac{26 - 14}{2 \times 26 - 14 - 22} \right) \times 10 \\ &= 20 + \left(\frac{12}{52 - 36} \right) \times 10 = 20 + \left(\frac{12}{16} \right) \times 10 \\ &= 20 + 7.5 = 27.5 \end{aligned}$$

Note: Here maximum frequency is 26. So, we selected the corresponding class as modal class.

Q.7 Find the Mode for the following frequency distribution:

<i>class</i>	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
<i>f_i</i>	5	4	6	10	8	15	7	3

Ans.

Given data is

<i>class</i>	<i>f_i</i>
5-10	5
10-15	4
15-20	6
20-25	10
25-30	8

30-35	15
35-40	7
40-45	3

Here $l = \text{lower limit of the modal class} = 30$,

$h = \text{width of the modal class} = 5$,

$f = \text{frequency of the modal class} = 15$,

$f_1 = \text{frequency of the class preceding the modal class} = 8$,

$f_2 = \text{frequency of the class following the modal class} = 7$

$$\begin{aligned} \text{Therefore, Mode} &= l + \left(\frac{f - f_1}{2f - f_1 - f_2} \right) \times h = 30 + \left(\frac{15 - 8}{2 \times 15 - 8 - 7} \right) \times 10 \\ &= 30 + \left(\frac{7}{30 - 15} \right) \times 10 = 30 + \left(\frac{7}{15} \right) \times 10 \\ &= 30 + 4.67 = 34.67 \end{aligned}$$

Note: Here maximum frequency is 15. So, we selected the corresponding class as modal class.

Exercise:

Q.1 Calculate the Mode for the following data:

- (i) 8, 12, 8, 19, 13, 12, 10, 12, 9, 12
- (ii) 15, 19, 17, 19, 20, 18, 20, 18, 16, 18, 12, 18
- (iii) 32, 35, 39, 35, 36, 37, 35, 40, 35, 32
- (iv) 7, 6, 5, 2, 7, 4, 1, 7, 2, 7, 1
- (v) 62, 75, 65, 78, 65, 62, 80, 62, 65, 72, 65, 75, 65

(vi)

x_i	13	15	17	19	21
f_i	7	4	8	3	5

(vii)

x_i	10	12	14	16	18	20	22
f_i	3	2	4	5	3	1	2

(viii)

x_i	1	2	3	4	5	6	7	8
f_i	2	4	5	3	4	3	4	3

(ix)

Class	0-2	2-4	4-6	6-8	8-10	10-12
f_i	2	4	5	6	4	2

(x)

Class	5-10	10-15	15-20	20-25	25-30	30-35	35-40
f_i	5	2	7	4	5	6	1

(xi)

Class	10-20	20-30	30-40	40-50	50-60	60-70
f_i	3	5	8	6	7	5

(xii)

Class	0-3	3-6	6-9	9-12	12-15	15-18	18-21	21-24
f_i	6	4	5	7	3	4	6	2

Dispersion: The degree, to which numerical data tends to spread about an average value, is called Dispersion.

We know that an average gives us idea of central tendency of the given data but it is necessary to know how the variants are scattered away from the average. We will try to understand this with an example. Consider the works of two typists who typed the following number of pages in 5 working days of a week:

	Mon.	Tue.	Wed.	Thur.	Fri.	Total Pages
1 st typist :	20	20	25	30	35	130
2 nd typist :	15	25	30	25	35	130

We see that each of the typist 1st and 2nd typed 130 pages in 5 working days and so the average in both the cases is 26. Thus there is no difference in the average, but we know that in the first case the number of pages varies from 20 to 35 while in the second case the number of pages varies from 15 to 35. This denotes that the greatest deviation from the mean in the first case is 9 and in the second case it is 11 i.e., there is a

difference between the two series. The variation of this type is termed dispersion. Various measures of dispersion are available. Among all, we will discuss the following: (1) Mean Deviation from Mean, (2) Standard Deviation.

(1) Mean Deviation from Mean:

(i) If $x_1, x_2, x_3, \dots, x_N$ is the given data, then mean deviation from mean is given by

M.D. (\bar{x}) = $\frac{\sum |x_i - \bar{x}|}{N}$, where \bar{x} is the mean of the given data and N is the number of terms.

(ii) For Frequency Distribution (Grouped or Ungrouped):

M.D. (\bar{x}) = $\frac{\sum f_i |x_i - \bar{x}|}{\sum f_i}$, where \bar{x} is the mean of the given data and f_i is the frequency corresponding to x_i .

Q.1 Find the mean deviation about mean for the following data:

14, 25, 18, 23, 20, 30, 15, 24, 12, 19

Ans. Given data is 14, 25, 18, 23, 20, 30, 15, 24, 12, 19

Here, $N = \text{no. of terms} = 10$

So, Mean is

$$\bar{x} = \frac{14+25+18+23+20+30+15+24+12+19}{10} = \frac{200}{10} = 20$$

Now,

x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $
14	-6	6
25	5	5
18	-2	2
23	3	3
20	0	0
30	10	10
15	-5	5
24	4	4
12	-8	8
19	-1	1

$$\sum |x_i - \bar{x}| = 44$$

Hence, mean deviation from mean is given by

$$\begin{aligned} \text{M.D.}(\bar{x}) &= \frac{\sum |x_i - \bar{x}|}{N} \\ &= \frac{44}{10} = 4.4 \end{aligned}$$

Q.2 Find the mean deviation about mean for the following data:

9, 7, 6, 7, 5, 4, 6, 4

Ans. Given data is 9, 7, 6, 7, 5, 4, 6, 4

Here, $N = \text{no. of terms} = 8$

So, Mean is

$$\bar{x} = \frac{9+7+6+7+5+4+6+4}{8} = \frac{48}{8} = 6$$

Now,

x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $
9	3	3
7	1	1
6	0	0
7	1	1
5	-1	1
4	-2	2
6	0	0
4	-2	2
		$\sum x_i - \bar{x} = 10$

Hence, mean deviation from mean is given by

$$\text{M.D.}(\bar{x}) = \frac{\sum |x_i - \bar{x}|}{N} = \frac{10}{8} = 1.25$$

Q.3 Find the mean deviation about mean for the following frequency distribution:

x_i	2	4	6	8	10
f_i	3	1	2	1	3

Ans.

x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$ x_i - \bar{x} $	$f_i x_i - \bar{x} $
-------	-------	-----------	-----------------	-------------------	-----------------------

2	3	6	-4	4	12
4	1	4	-2	2	2
6	2	12	0	0	0
8	1	8	2	2	2
10	3	30	4	4	12
	$\Sigma f_i = 10$	$\Sigma f_i x_i = 60$			$\Sigma f_i x_i - \bar{x} = 28$

We know that, Mean = $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{60}{10} = 6$

Therefore, Mean deviation from mean is given by

$$\text{M.D.}(\bar{x}) = \frac{\Sigma f_i |x_i - \bar{x}|}{\Sigma f_i} = \frac{28}{10} = 2.8$$

Q.4 Find the mean deviation about mean for the following frequency distribution:

x_i	1	2	3	4	5
f_i	6	4	3	4	3

Ans.

x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$ x_i - \bar{x} $	$f_i x_i - \bar{x} $
1	6	6	-1.7	1.7	10.2
2	4	8	-0.7	0.7	2.8
3	3	9	0.3	0.3	0.9
4	4	16	1.3	1.3	5.2
5	3	15	2.3	2.3	6.9
	$\Sigma f_i = 20$	$\Sigma f_i x_i = 54$			$\Sigma f_i x_i - \bar{x} = 26$

We know that, Mean = $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{54}{20} = 2.7$

Therefore, Mean deviation from mean is given by

$$\text{M.D.}(\bar{x}) = \frac{\Sigma f_i |x_i - \bar{x}|}{\Sigma f_i} = \frac{26}{20} = 1.3$$

Q.5 Find the mean deviation about mean for the following continuous frequency distribution:

class	0-10	10-20	20-30	30-40	40-50
f_i	8	4	3	3	2

Ans.

<i>class</i>	x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$ x_i - \bar{x} $	$f_i x_i - \bar{x} $
0-10	5	8	40	-13.5	13.5	108
10-20	15	4	60	-3.5	3.5	14
20-30	25	3	75	6.5	6.5	19.5
30-40	35	3	105	16.5	16.5	49.5
40-50	45	2	90	26.5	26.5	53
		$\Sigma f_i = 20$	$\Sigma f_i x_i = 370$			$\Sigma f_i x_i - \bar{x} = 244$

Here $N = \Sigma f_i = 20$

We know that, Mean = $\frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{370}{20} = 18.5$

Therefore, Mean deviation from mean is given by

$$\text{M.D.}(\bar{x}) = \frac{\Sigma f_i |x_i - \bar{x}|}{\Sigma f_i} = \frac{244}{20} = 12.2$$

Q.6 Find the mean deviation about mean for the following continuous frequency distribution:

Ans.

<i>class</i>	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
f_i	5	4	1	2	3	2	2	3

<i>class</i>	x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$ x_i - \bar{x} $	$f_i x_i - \bar{x} $
5-10	7.5	5	37.5	-15.2	15.2	76
10-15	12.5	4	50	-10.2	10.2	40.8
15-20	17.5	1	17.5	-5.2	5.2	5.2
20-25	22.5	2	45	-0.2	0.2	0.4
25-30	27.5	3	82.5	4.8	4.8	14.4
30-35	32.5	2	65	9.8	9.8	19.6
35-40	37.5	2	75	14.8	14.8	29.6
40-45	42.5	3	127.5	19.8	19.8	59.4
		$\Sigma f_i = 22$	$\Sigma f_i x_i = 500$			$\Sigma f_i x_i - \bar{x} = 245.4$

Here $N = \Sigma f_i = 22$

We know that, Mean = $\frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{500}{22} = 22.7$

Therefore, Mean deviation from mean is given by

$$\text{M.D.}(\bar{x}) = \frac{\Sigma f_i |x_i - \bar{x}|}{\Sigma f_i} = \frac{245.4}{22} = 11.15$$

Exercise:

Q.1 Calculate the Mean Deviation from mean for the following data:

(i) 8, 12, 5, 19, 13

(ii) 120, 160, 220, 210, 190

(iii) 23, 34, 25, 28, 39, 36, 24, 35

(iv)

x_i	13	15	17	19	21
f_i	1	3	1	2	3

(v)

x_i	10	12	14	16	18	20	22
f_i	3	2	4	5	3	1	2

(vi)

x_i	1	2	3	4	5	6	7	8
f_i	6	4	4	3	3	3	4	3

(vii)

x_i	1	2	3	4	5	6
f_i	2	4	6	6	4	2

(viii)

Class	0-2	2-4	4-6	6-8	8-10	10-12
f_i	2	4	6	6	4	2

(ix)

Class	5-10	10-15	15-20	20-25	25-30	30-35	35-40
f_i	5	2	7	4	5	6	1

(x)

Class	10-20	20-30	30-40	40-50	50-60	60-70
f_i	8	5	4	6	7	9

(2) Standard Deviation: It is defined as the positive square root of the mean of the squares of the deviations from the mean and it is denoted by s .

(i) If $x_1, x_2, x_3, \dots, x_N$ is the given data, then standard deviation is given by

$$\text{S.D.} = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N}}, \text{ where } \bar{x} \text{ is the mean of the given data and } N \text{ is the number of terms.}$$

(ii) For Frequency Distribution (Grouped or Ungrouped):

$$\text{S.D.} = \sqrt{\frac{\sum f_i(x_i - \bar{x})^2}{\sum f_i}}, \text{ where } \bar{x} \text{ is the mean of the given data and } f_i \text{ is the frequency corresponding to } x_i.$$

Q.1 Find the standard deviation for the following data:

14, 25, 18, 23, 20, 30, 15, 24, 12, 19

Ans. Given data is 14, 25, 18, 23, 20, 30, 15, 24, 12, 19

Here, $N = \text{no. of terms} = 10$

So, Mean is

$$\bar{x} = \frac{14+25+18+23+20+30+15+24+12+19}{10} = \frac{200}{10} = 20$$

Now,

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
14	-6	36
25	5	25
18	-2	4
23	3	9
20	0	0
30	10	100
15	-5	25
24	4	16
12	-8	64
19	-1	1
		$\sum (x_i - \bar{x})^2 = 280$

Hence, Standard Deviation is given by

$$\begin{aligned} \text{S.D.} &= \sqrt{\frac{\sum(x_i - \bar{x})^2}{N}} \\ &= \sqrt{\frac{280}{10}} = \sqrt{28} = 5.29 \end{aligned}$$

Q.2 Find the standard deviation for the following data:

9, 7, 6, 7, 5, 4, 6, 4

Ans. Given data is 9, 7, 6, 7, 5, 4, 6, 4

Here, $N = \text{no. of terms} = 8$

So, Mean is

$$\bar{x} = \frac{9+7+6+7+5+4+6+4}{8} = \frac{48}{8} = 6$$

Now,

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
9	3	9
7	1	1
6	0	0
7	1	1
5	-1	1
4	-2	4
6	0	0
4	-2	4
		$\sum (x_i - \bar{x})^2 = 20$

Hence, Standard Deviation is given by

$$\text{S.D.} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}} = \sqrt{\frac{20}{8}} = \sqrt{2.5} = 1.58$$

Q.3 Find the standard deviation for the following frequency distribution:

x_i	2	4	6	8	10
f_i	3	1	2	1	3

Ans.

x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
2	3	6	-4	16	48
4	1	4	-2	4	4
6	2	12	0	0	0
8	1	8	2	4	4
10	3	30	4	16	48

$\Sigma f_i = 10$	$\Sigma f_i x_i = 60$			$\Sigma f_i (x_i - \bar{x})^2 = 104$
-------------------	-----------------------	--	--	--------------------------------------

We know that, Mean = $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{60}{10} = 6$

Therefore, Standard deviation given by

$$\begin{aligned} \text{S.D.} &= \sqrt{\frac{\Sigma f_i (x_i - \bar{x})^2}{\Sigma f_i}} = \sqrt{\frac{104}{10}} \\ &= \sqrt{10.4} = 3.22 \end{aligned}$$

Q.4 Find the standard deviation for the following frequency distribution:

x_i	1	2	3	4	5
f_i	6	4	3	4	3

Ans.

x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
1	6	6	-1.7	2.89	17.34
2	4	8	-0.7	0.49	1.96
3	3	9	0.3	0.09	0.27
4	4	16	1.3	1.69	6.76
5	3	15	2.3	5.29	15.87
	$\Sigma f_i = 20$	$\Sigma f_i x_i = 54$			$\Sigma f_i (x_i - \bar{x})^2 = 42.2$

We know that, Mean = $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{54}{20} = 2.7$

Therefore, Standard Deviation is given by

$$\text{S.D.} = \sqrt{\frac{\Sigma f_i (x_i - \bar{x})^2}{\Sigma f_i}} = \sqrt{\frac{42.2}{20}} = \sqrt{2.11} = 1.45$$

Q.5 Find the standard deviation for the following continuous frequency distribution:

class	0-10	10-20	20-30	30-40	40-50
f_i	8	4	3	3	2

Ans.

class	x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
0-10	5	8	40	-13.5	182.25	1458

10-20	15	4	60	-3.5	12.25	49
20-30	25	3	75	6.5	42.25	126.75
30-40	35	3	105	16.5	272.25	816.75
40-50	45	2	90	26.5	702.25	1404.5
		$\Sigma f_i = 20$	$\Sigma f_i x_i = 370$			$\Sigma f_i (x_i - \bar{x})^2 = 3855$

Here $N = \Sigma f_i = 20$

We know that, Mean = $\frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{370}{20} = 18.5$

Therefore, Standard Deviation is given by

$$\text{S.D.} = \sqrt{\frac{\Sigma f_i (x_i - \bar{x})^2}{\Sigma f_i}} = \sqrt{\frac{3855}{20}} = \sqrt{192.75} = 13.88$$

Q.6 Find the standard deviation for the following continuous frequency distribution:

Ans.

class	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
f_i	5	4	1	2	3	2	2	3

class	x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
5-10	7.5	5	37.5	-15.2	231.04	1155.2
10-15	12.5	4	50	-10.2	104.04	416.16
15-20	17.5	1	17.5	-5.2	27.04	27.04
20-25	22.5	2	45	-0.2	0.04	0.08
25-30	27.5	3	82.5	4.8	23.04	69.12
30-35	32.5	2	65	9.8	96.04	192.08
35-40	37.5	2	75	14.8	219.04	438.08
40-45	42.5	3	127.5	19.8	392.04	1176.12
		$\Sigma f_i = 22$	$\Sigma f_i x_i = 500$			$\Sigma f_i (x_i - \bar{x})^2 = 3473.88$

Here $N = \Sigma f_i = 22$

We know that, Mean = $\frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{500}{22} = 22.7$

Therefore, Standard Deviation is given by

$$\text{S.D.} = \sqrt{\frac{\Sigma f_i (x_i - \bar{x})^2}{\Sigma f_i}} = \sqrt{\frac{3473.88}{22}} = \sqrt{157.90} = 12.56$$

Exercise:

Q.1 Calculate the Standard Deviation for the following data:

(i) 8, 12, 5, 19, 13

(ii) 120, 160, 220, 210, 190

(iii) 23, 34, 25, 28, 39, 36, 24, 35

(iv)

x_i	13	15	17	19	21
f_i	1	3	1	2	3

(v)

x_i	10	12	14	16	18	20	22
f_i	3	2	4	5	3	1	2

(vi)

x_i	1	2	3	4	5	6	7	8
f_i	6	4	4	3	3	3	4	3

(vii)

x_i	1	2	3	4	5	6
f_i	2	4	6	6	4	2

(viii)

Class	0-2	2-4	4-6	6-8	8-10	10-12
f_i	2	4	6	6	4	2

(ix)

Class	5-10	10-15	15-20	20-25	25-30	30-35	35-40
f_i	5	2	7	4	5	6	1

(x)

Class	10-20	20-30	30-40	40-50	50-60	60-70
f_i	8	5	4	6	7	9

Correlation Coefficient: The correlation coefficient is a statistical measure of the strength of the relationship between the relative movements of two variables. It ranges between -1 and 1 . If the two variables are in perfect linear relationship, the correlation coefficient will be either 1 or -1 . The sign depends on whether the variables are positively or negatively related. The correlation coefficient is 0 if there is no linear relationship between the variables.