

7

Statistics



Let's study.

- Sub-divided bar-diagram
- Percentage bar-diagram
- Primary and secondary data
- Grouped and ungrouped frequency distribution
- Cumulative frequency distribution
- Mean, Median and Mode for ungrouped data



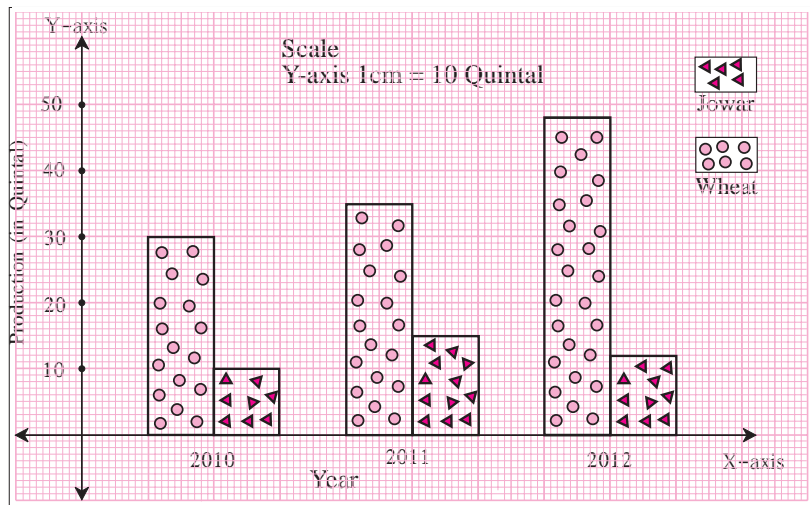
Let's recall.

In earlier standards, we have learnt how to draw a simple bar-diagram and a joint bar-diagram. Also, we have observed various graphs from newspapers, magazines, television etc. and gathered information from them.

It is very important to decide according to the nature of the data, what diagram or graph would be suitable to represent it.

A farmer has produced Wheat and Jowar in his field. The following joint bar diagram shows the production of Wheat and Jowar. From the given diagram answer the following questions :

- Which crop production has increased consistently in 3 years ?
- By how many quintals the production of Jowar has reduced in 2012 as compared to 2011?
- What is the difference between the production of Wheat in 2010 and 2012?
- Complete the following table using this diagram.



Production (in Quintal) Year	Wheat	Jowar	Total
2010			
2011			
2012	48	12	60



Let's learn

Sub-divided bar diagram

To compare the information in the given data, we can also draw another type of bar-diagram

To draw it, we add the numerical values of the entities, decide a scale and show the total by a bar proportional to the scale. Then we divide the bar in parts, proportional to the entities we had added. Hence this type of diagram is called a sub-divided bar diagram.

Now let us show the information in the previous example by a sub-divided bar diagram.

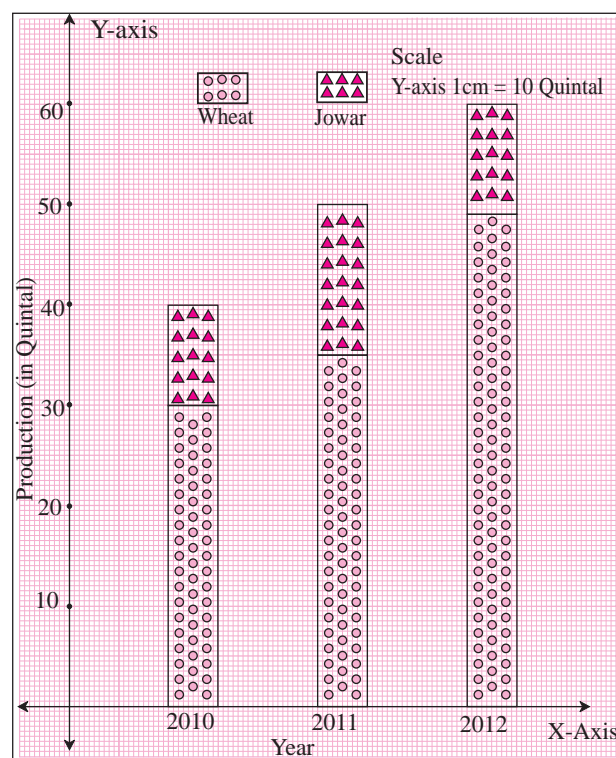
(i) Show the total production of the year 2010 by a bar. The height of the bar should be to the decided scale.

(ii) Show the production of wheat by lower part of the bar, the height of which is to the scale.

(iii) Obviously, remaining part of the bar denotes the production of Jowar for the year.

(iv) Similarly draw divided bars to show productions of the years 2011 and 2012.

When two quantities are compared using percentages, it is more informative. We have studied this before. For example, if there is Rs.600 profit on Rs.2,000 and Rs.510 profit on Rs.1,500; Rs.600 looks greater amount. But if we calculate their percentages they are 30% and 34% respectively. Hence it is clear that Rs.510 profit on Rs.1,500 is a more profitable transaction.



Percentage bar diagram

To compare the given information, in a different way, it is converted into percentages and then a sub-divided bar diagram is drawn. Such diagram is known as 'Percentage bar-diagram'.

The information in the previous example is converted into percentages as shown in the adjacent table.

Year	Production of Wheat (Quintal)	Production of Jowar (Quintal)	Percentage production of Wheat as compared to total production
2010	30	10	$\frac{30}{40} \times 100 = 75\%$
2011	35	15	$\frac{35}{50} \times 100 = 70\%$
2012	48	12	$\frac{48}{60} \times 100 = 80\%$

The information is shown in the percentage bar diagram by following steps

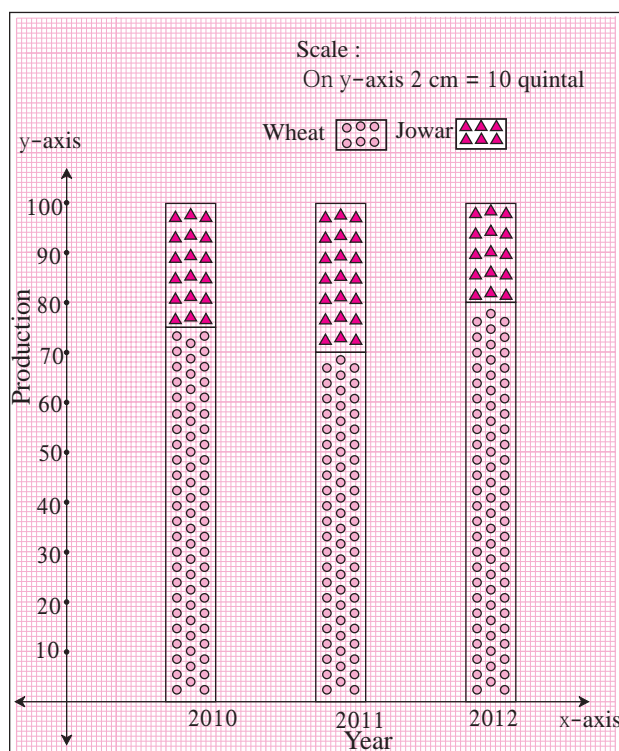
(i) Yearly productions of Wheat and Jowar are converted into percentages.

(ii) The height of each bar to scale is taken as 100.

(iii) The percentage of production of Wheat is shown by the lower part of the bar to the scale.

(iv) The remaining upper part of the bar shows percentage production of Jowar.

Information of more than two entities can be shown by a sub divided bar diagram or by a percentage bar diagram.



Solved examples :

Ex. 1. In the neighbouring figure, percentage bar-diagram is given. Percentage expenses on different items of two families are given. Answer the following questions based on it :

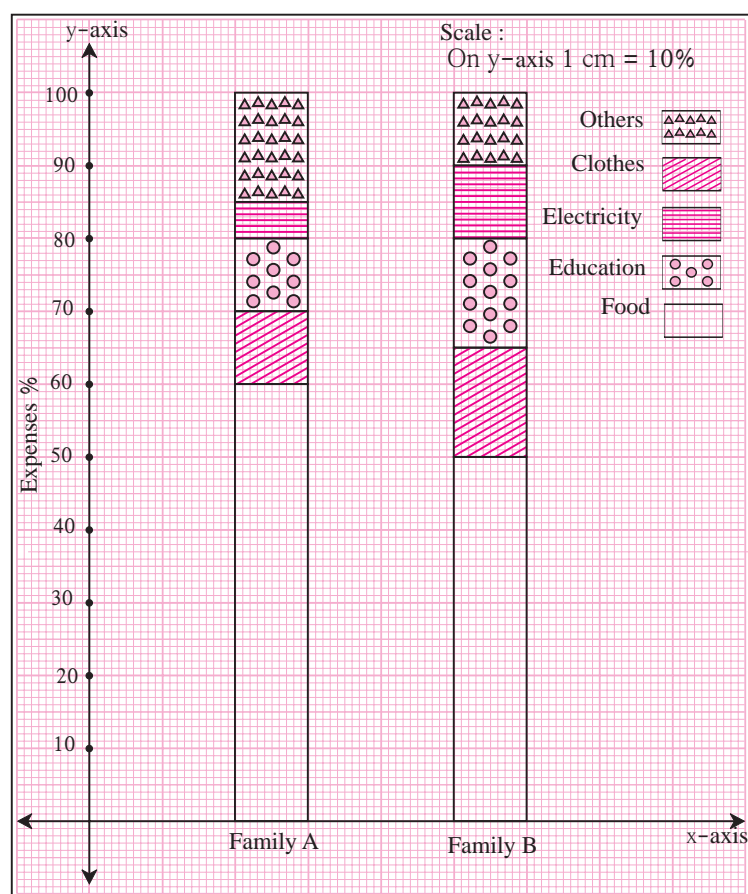
(i) Write the percentage expenses of every component for each family.

(ii) Which family spends more percent of expenses on food as compared to the other and by how much ?

(iii) What are the percentage expenses on other items ?

(iv) Which family shows more percentage expenses on electricity?

(v) Which family's percentage expense is more on education ?



Solution :

(i)

Expenses Family	Food	Clothes	Education	Electricity	Others
A	60%	10%	10%	5%	15%
B	50%	15%	15%	10%	10%

(ii) Food expenses of family A are more by 10% as compared with family B.

(iii) Other expenses of family A are 15% and that of family B are 10%.

(iv) Percentage expenses on electricity of family B is greater.

(v) Percentage expenses on education are more of family B.

Practice set 7.1

(1) The following table shows the number of Buses and Trucks in nearest lakh units. Draw percentage bar-diagram. (Approximate the percentages to the nearest integer)

Year	No. of Trucks	No. of Buses
2005-2006	47	9
2007-2008	56	13
2008-2009	60	16
2009-2010	63	18

(2) In the table given below, the information is given about roads. Using this draw sub-divided and percentage bar-diagram (Approximate the percentages to the nearest integer)

Year	Permanent Roads (Lakh km.)	Temporary Roads (Lakh km.)
2000-2001	14	10
2001-2002	15	11
2003-2004	17	13
2007-2008	20	19

Activity : In the following table, the information of number of girls per 1000 boys is given in different States. Fill in the blanks and complete the table.

States	Boys	Girls	Total	Percentage of boys	Percentage of girls
Assan	1000	960	1960	$\frac{1000}{1960} \times \frac{100}{1} = 51\%$	$100 - 51 = 49\%$
Bihar	1000	840	1840		
Punjab	1000	900			
Keala	1000	1080			
Maha-rashtra	1000	900			

Draw percentage bar-diagram from this information and discuss the findings from the diagram.



Use your brain power!

On page 111, for the given activity, the information of number of girls per 1000 boys is given for five states.

The literacy percentage of these five States is given below.

Assam (73%), Bihar (64%), Punjab (77%), Kerala (94%), Maharashtra (83%).

Think of the number of girls and the literacy percentages in the respective states. Can you draw any conclusions from it ?



Let's discuss.

To show following information diagrammatically, which type of bar-diagram is suitable?

- (1) Literacy percentage of four villages.
- (2) The expenses of a family on various items.
- (3) The numbers of girls and boys in each of five divisions.
- (4) The number of people visiting a science exhibition on each of three days.
- (5) The maximum and minimum temperature of your town during the months from January to June.
- (6) While driving a two-wheeler, number of people wearing helmets and not wearing helmet in 100 families.



Let's learn.

Statistics

Suppose, a large group (population) is to be studied with a particular aspect. (For example, blood pressures of senior citizens in a locality) For the purpose, a sufficiently small part of the group is selected randomly. This small group represents the large group (sample). The necessary information is gathered from the representative group which, in general, is numerical in most of the cases. The analysis of the information enables us to draw conclusions. The study of this type is called 'Statistics'.

The word Statistics is originated from the Latin word 'status', which means situation of a state. This suggests that in ancient times statistics was used for administrative purposes. Today, it is used in many fields of knowledge.

Sir Ronald Aylmer Fisher (17 February 1890 - 29 July 1962) is known as Father of Statistics.

Data collection

Teacher : Suppose, you want to know how much agricultural land is owned by every family in the village. What will you do?

Robert : We will visit each house in the village and record the information about agricultural land owned by them.

Teacher : Correct, my dear students, when we collect information of a group it is called as 'data'. Generally it is numerical. We must know the purpose of collecting it. If some one collects the information personally by asking questions, taking measurements, etc. it is called as the 'Primary Data'.

Afrin : So, the data collected regarding agricultural land, as Robert said, is primary data.

Teacher : Yes, well said Afrin !

Ramesh : But what to do if we want to collect the above data in a short time ?

Teacher : What Ramesh is saying is right. In this situation we have to use another method of data collection. Think what it could be ?

Ketaki : We can go to village Talathi office and can get the information from their records.

Teacher : Correct, in some situations, because of lack of time, lack of resources, we can't collect information personally. In such cases, we have to use the information, already collected in the form of records, information published in journals, case-studies etc. The data collected from such sources is known as 'Secondary data'. So as suggested by Ketaki, the data collected from village Talathi office, regarding agricultural land is secondary data.

See the following examples :

- (i) The chart made from information published in news paper is secondary data.
- (ii) The feedback of customers in a restaurant regarding quality of the food is primary data.
- (iii) The heights of students recorded by actual measurements is primary data.

Primary data	Secondary data
1. It requires more time. 2. It is up to date and detailed information. 3. It is correct and reliable.	1. It is readily available, so needs less time 2. It is taken from already collected data. It is not necessarily up to date. It may lack in details also. 3. It may be less reliable.

Activity : You gather information for several reasons. Take a few examples and discuss whether the data is primary or secondary.

Practice set 7.2

- (1) Classify following information as primary or secondary data.
- (i) Information of attendance of every student collected by visiting every class in a school.
 - (ii) The information of heights of students was gathered from school records and sent to the head office, as it was to be sent urgently.
 - (iii) In the village Nandpur, the information collected from every house regarding students not attending school.
 - (iv) For science project, information of trees gathered by visiting a forest.,



Let's recall.

Classification of data

Ex.(1) The record of marks out of 20 in Mathematics in the first unit test is as follows.

20, 6, 14, 10, 13, 15, 12, 14, 17, 17, 18, 11, 19, 9, 16, 18, 14, 7, 17, 20,
8, 15, 16, 10, 15, 12, 18, 17, 12, 11, 11, 10, 16, 14, 16, 18, 10, 7, 17, 14,
20, 17, 13, 15, 18, 20, 12, 12, 15, 10

What is the above information called ?

-Primary data

What is each of the numbers in the data called ?

-A score

Answer the following questions, from the above information.

- (i) How many students scored 15 marks ?
- (ii) How many students scored more than 15 marks ?
- (iii) How many students scored less than 15 marks ?
- (iv) What is the lowest score of the group ?
- (v) What is the highest score the group ?



Let's discuss.

(1) Was it easy to find out the answers of the above questions ? Did you refer the data frequently ?

(2) What should we do to find answers easily ?

Shamim : We had to refer the data frequently. It was tedious and boring. If we write the data in ascending or descending order the above answers could be found easily.

According to Shamim's suggestion, let us arrange the data in ascending order.

6, 7, 7, 8, 9, 10, 10, 10, 10, 10, 11, 11, 11, 12, 12, 12, 12, 12, 13, 13,
14, 14, 14, 14, 14, 15, 15, 15, 15, 15, 16, 16, 16, 16, 17, 17, 17, 17, 17, 17,
18, 18, 18, 18, 18, 19, 20, 20, 20, 20

Verify that the ascending order of scores helps to find the answers of the questions in Ex. (1) easily.



Let's recall.

Martin : Writing the data in a tabular form can also make the above work easy, We have studied this in previous year. This table is known as 'frequency distribution table'.

Teacher : Correct Martin ! Now let us prepare a table of the information given in example (1).

In example (1), the lowest score is 6 and the highest score is 20. Hence in the table, we write, numbers from 6 to 20 in the column of scores. In second column we record tally marks and in last column, frequency by counting the tally marks. (Complete the table)

Frequency Distribution Table

Score	Tally Marks	Frequency (No.of students)
6		1
7		2
8		
9		
10		5
11		
12		
13		
14		
15		
16		
17		6
18		
19		
20		4
		Total N = 50

Make it sure that the sum of all frequencies, N is 50.



Let's discuss.

- Is the above table is very long ?
- If the number of observations are more, is it difficult to make a table ?

Grouped frequency distribution table

Teacher : From above discussion, we conclude that, when number of observations is large, preparing a table is time consuming. What can be done to condense the data and save time ?

Rohit : In this situation, we can group the scores in the data.

Teacher : Well done Rohit! If we group the scores, that means if we make their classes, then the data will be condensed and time can be saved. Such a table is known as grouped frequency distribution table.

These are two methods of preparing a grouped frequency distribution table.

- (1) Inclusive method, (2) Exclusive method.

(1) Inclusive method (Discrete classes)

6, 7, 7, 8, 9, 10, 10, 10, 10, 10, 11, 11, 11, 12, 12, 12, 12, 12, 13, 13, 14, 14, 14, 14, 14, 15, 15, 15, 15, 15, 16, 16, 16, 16, 17, 17, 17, 17, 17, 18, 18, 18, 18, 18, 19, 20, 20, 20, 20

In the above scores the smallest is \square and the largest is \square . The difference between largest and smallest scores is $20 - 6 = 14$. This difference is called as 'Range of the data'.

By noticing the range, how can we classify the data into convenient classes ?

We can take classes like this.

- (i) 6 to 8, 9 to 11, 12 to 14, 15 to 17, 18 to 20 or
(ii) 6 to 10, 11 to 15, 16 to 20.

By taking classes 6 – 10, 11 – 15 & 16 – 20, let us prepare frequency distribution table.

Grouped Frequency Distribution Table (inclusive method)

Class	Tally marks	Frequency (No.of students)
6 – 10		10
11 – 15
16 – 20	20
		N = 50

While preparing this table, 6, 10 and all scores between them are included in the class 6–10 hence such classes are known as 'Inclusive Classes' of discrete class.



Let's learn.

Basic terms in statistics :

- (1) **Class :** When the observations are divided into suitable groups, each of the groups is called a 'Class'.
- (2) **Class-Limit :** The end values of the classes are called class-limits.
For the class 6-10, the lower class limit is 6 and the upper class limit is 10.
- (3) **Frequency :** The total number of observations in to each class is called the 'frequency' of the that class.
In the above table, there are 20 observations in the class 11 to 15. Hence frequency of the class 11 – 15 is 20

4. **Class width or Class Size or Class-interval :** When continuous classes are given, the difference between upper class limit and lower class limit is known as class-width.

For example, if 5 – 10, 10 – 15, 15–20, ...are given classes,

class width of 5–10 is $10 - 5 = 5$

5. **Class mark :** The average of the lower class limit and the upper class limit for a given class is known as class mark

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

For example, class mark the for class 11 to 15 = $\frac{\boxed{} + \boxed{}}{2} = \frac{26}{2} = 13$

(2) Exclusive method (Continuous classes)

Ex. 6, 10, 10.5, 11, 15.5, 19, 20, 12, 13 are the given observations. By taking classes 6-10, 11-15, 16-20 prepare grouped frequency distribution table

Solution :

Classes	Tally marks	Frequency (f)
6-10		2
11-15		3
16-20		2

In the above table, we could not include observations 10.3 and 15.7.

Because the numbers 10.3 and 15.7 cannot be included in any of the classes 6-10, 11-15, 16-20. Hence in order to include them, we have to change the structure of the classes. Therefore if we take class intervals as 5-10 10-15, 15-20 the problem will be solved. The scores 10.3 and 15.7 can be included in the classes 10-15 and 15-20 respectively. But still a question arises. In which interval the score 10 should be included ? In 5-10 or 10-20 ? To overcome the difficulty, we follow a convention. We will include the score 10 in the class 10-15 instead of 5-10. That is the upper class limit of a class should be excluded from the class. Therefore, this is called the exclusive method of classification.

Now taking classes accordingly and as per the convention of exclusion, we can prepare the table as follows.

Grouped frequency distribution table (Exclusive method)

Class interval Marks	Tally marks	Frequency (No. of students)
5-10		1
10-15		5
15-20		2
20-25		1

**Remember this !****Frequency Distribution Table****Ungrouped**

Ages of 9th std. students	No. of Students
14	12
15	23
16	10

Grouped**Inclusive Method
(Discrete classes)**

Shoe size	No. of Students
2-4	12
5-7	29
8-10	7

**Exclusive Method
(Continuous classes)**

Height (cm)	No. of students
145-150	18
150-155	27
155-160	3

Practice set 7.3

- (1) For class interval 20-25 write the lower class limit and the upper class limit.
- (2) Find the class-mark of the class 35-40.
- (3) If class mark is 10 and class width is 6 then find the class.
- (4) Complete the following table.

Classes (age)	Tally marks	Frequency (No. of students)
12-13		<input type="text"/>
13-14		<input type="text"/>
14-15		<input type="text"/>
15-16		<input type="text"/>
		$N = \sum f = 35$

- (5) In a 'tree plantation' project of a certain school there are 45 students of 'Harit Sena.' The record of trees planted by each student is given below :

3, 5, 7, 6, 4, 3, 5, 4, 3, 5, 4, 7, 5, 3, 6, 6, 5, 3, 4, 5, 7, 3, 5, 6, 4, 4, 3,
5, 6, 6, 4, 3, 5, 7, 3, 4, 5, 7, 6, 4, 3, 5, 4, 4, 7.

Prepare a frequency distribution table of the data.

- (6) The value of π upto 50 decimal places is given below :

3.14159265358979323846264338327950288419716939937510

From this information prepare an ungrouped frequency distribution table of digits appearing after the decimal point.

- (7) In the tables given below, class-mark and frequencies is given. Construct the frequency tables taking inclusive and exclusive classes.

(i)

Class width	Frequency
5	3
15	9
25	15
35	13

(ii)

Class width	Frequency
22	6
24	7
26	13
28	4

- (8) In a school, 46 students of 9th standard, were told to measure the lengths of the pencils in their compass-boxes in centimeters. The data collected was as follows.

16, 15, 7, 4.5, 8.5, 5.5, 5, 6.5, 6, 10, 12,
 13, 4.5, 4.9, 16, 11, 9.2, 7.3, 11.4, 12.7, 13.9, 16,
 5.5, 9.9, 8.4, 11.4, 13.1, 15, 4.8, 10, 7.5, 8.5, 6.5,
 7.2, 4.5, 5.7, 16, 5.7, 6.9, 8.9, 9.2, 10.2, 12.3, 13.7,
 14.5, 10

By taking inclusive classes 0-5, 5-10, 10-15.... prepare a grouped frequency distribution table.

- (9) In a village, the milk was collected from 50 milkmen at a collection center in litres as given below :

27, 75, 5, 99, 70, 12, 15, 20, 30, 35, 45, 80,
 77, 90, 92, 72, 4, 33, 22, 15, 20, 28, 29, 14,
 16, 20, 72, 81, 85, 10, 16, 9, 25, 23, 26, 46,
 55, 56, 66, 67, 51, 57, 44, 43, 6, 65, 42, 36,
 7, 35.

By taking suitable classes, prepare grouped frequency distribution table.

- (10) 38 people donated to an organisation working for differently abled persons. The amount in rupees were as follows :

101, 500, 401, 201, 301, 160, 210, 125, 175, 190, 450, 151,
 101, 351, 251, 451, 151, 260, 360, 410, 150, 125, 161, 195,
 351, 170, 225, 260, 290, 310, 360, 425, 420, 100, 105, 170,
 250, 100.

- (i) By taking classes 100-149, 150-199, 200-249... prepare grouped frequency distribution table.

- (ii) From the table, find the number of people who donated rupees 350 or more.



Let's learn.

Less than Cumulative frequency less than the upper class limit

Ex. : The following information is regarding marks in mathematics, obtain out of 40, scored by 50 students of 9th std. in the first unit test.

Class Interval	Frequency (no.of students)
0-10	02
10-20	12
20-30	20
30-40	16
	Total N = 50

(1) From the table, fill in the blanks in the following statements.

(i) For class interval 10-20 the lower class limit is and upper class limit is

(ii) How many students obtained marks less than 10 ?

(iii) How many students obtained marks less than 20 ? $2 + \text{$ = 14

(iv) How many students obtained marks less than 30 ? $\text{$ + $\text{$ = 34

(v) How many students obtained marks less than 40 ? $\text{$ + $\text{$ = 50



Remember this !

The sum of the frequency of a certain class and all the frequencies of previous classes is called as cumulative frequency less than the upper class limit for that given class. In short, it is also called as 'less than type' cumulative frequency.

The Meaning of less than type cumulative frequency :

Classes marks	Frequency	Less than type cumulative frequency
0-10	2	2
10-20	12	$2 + 12 = \text{$
20-30	20	$\text{$ + 20 = 34
30-40	16	$34 + \text{$ = 50
Total 50		

Class	Cumulative frequency	Meaning of less than type cumulative frequency
0-10	2	2 students got less than 10 marks
10-20	14	14 students got less than 20 marks
20-30	34	34 students got less than 30 marks
30-40	50	50 students got less than 40 marks
Total 50		

(2) Cumulative frequency more than or equal to the lower class limit

Classes	Frequency	Cumulative freq.	Classes	Cum. freq.	Meaning of cumulative frequency more than or equal to the lower class limit
0-10	2	50	0-10	50	50 students got 0 or more than 0 marks
10-20	12	$50 - 2 = 48$	10-20	48	48 students got 10 or more than 10 marks
20-30	20	$48 - 12 = 36$	20-30	36	36 students got 20 or more than 20 marks
30-40	16	$36 - 20 = 16$	30-40	16	16 students got 30 or more than 30 marks
Total 50					

Ex. A sports club has organised a table-tennis tournaments. The following table gives the distribution of players ages. Find the cumulative frequencies equal to or more than the lower class limit and complete the table.

Solution : Equal to lower limit or more than lower limit type of cumulative table.

Age (Year)	Tally marks	Frequency (No. of students)	Equal to lower limit or more than lower limit
10-12		09	50
12 - 14		<input type="text"/>	<input type="text"/> - 9 = 41
14-16		<input type="text"/>	$41 - 23 =$ <input type="text"/>
15 - 16		05	<input type="text"/> - 13 = <input type="text"/>
		Total N = 50	

Practice set 7.4

(1) Complete the following cumulative frequency table :

Class (Height in cm)	Frequency (No. of students)	Less than type frequency
150-153	05	05
153-156	07	$05 +$ <input type="text"/> $=$ <input type="text"/>
156-159	15	<input type="text"/> $+ 15 =$ <input type="text"/>
159-162	10	<input type="text"/> $+ $ <input type="text"/> $= 37$
162-165	05	$37 + 5 = 42$
165-168	03	<input type="text"/> $+ $ <input type="text"/> $= 45$
	Total N = 45	

(2) Complete the following Cumulative Frequency Table :

Class (Monthly income in Rs.)	Frequency (No. of individuals)	More than or equal to type cumulative frequency
1000-5000	45
5000-10000	19
10000-15000	16
15000-20000	02
20000-25000	05
	Total N = 87	

(3) The data is given for 62 students in a certain class regarding their mathematics marks out of 100. Take the classes 0-10, 10-20.. and prepare frequency distribution table and cumulative frequency table more than or equal to type.

55, 60, 81, 90, 45, 65, 45, 52, 30, 85, 20, 10,
 75, 95, 09, 20, 25, 39, 45, 50, 78, 70, 46, 64,
 42, 58, 31, 82, 27, 11, 78, 97, 07, 22, 27, 36,
 35, 40, 75, 80, 47, 69, 48, 59, 32, 83, 23, 17,
 77, 45, 05, 23, 37, 38, 35, 25, 46, 57, 68, 45,
 47, 49.

From the prepared table, answer the following questions :

- How many students obtained marks 40 or above 40 ?
 - How many students obtained marks 90 or above 90 ?
 - How many students obtained marks 60 or above 60 ?
 - What is the cumulative frequency of equal to or more than type of the class 0-10?
- (4) Using the data in example (3) above, prepare less than type cumulative frequency table and answer the following questions.
- How many students obtained less than 40 marks ?
 - How many students obtained less than 10 marks ?
 - How many students obtained less than 60 marks ?
 - Find the cumulative frequency of the class 50-60.



Let's learn.

Measures of central tendency

Central Tendency : If the data collected in a survey of a group is sufficiently large, then it generally shows a peculiar property. The numbers in the data tend to cluster around a certain number. This property is called the **central tendency** of the group.

The number around which the numbers in the data tend to cluster is called **measure** of central tendency. It is supposed that the measure is a representative of the data.

In statistics, the measures of central tendency mainly used are as follows.

The following measures of central tendency are used :

- (1) **Mean** : The arithmetical average of all observations in the given data is known as its 'Arithmetic mean' or simply 'mean'.

$$\text{Mean} = \frac{\text{The sum of all observations in the data}}{\text{Total number of observation}}$$

Ex. (1) Find the mean of numbers 25, 30, 27, 23 and 25.

Solution :
$$\frac{25 + 30 + 27 + 23 + 25}{5} = \frac{130}{5} = 26$$

Ex. (2) The first unit test of 40 marks was conducted for a class of 35 students. The marks obtained by the students were as follows. Find the mean of the marks.

40, 35, 30, 25, 23, 20, 14, 15, 16, 20, 17, 37,
37, 20, 36, 16, 30, 25, 25, 36, 37, 39, 39, 40,
15, 16, 17, 30, 16, 39, 40, 35, 37, 23, 16.

Solution : Here, we can add all observations, but it will be a tedious job. Here 3 students obtained 30 marks each. So their sum is $30 + 30 + 30 = 90$, which is 30×3 . In this way the sum of marks of all students is worked out in the following table.

In statistics, it is convenient to use the Greek letter Σ (sigma) to show the sum of numbers.

In the adjacent table, 956 is the sum of the products $14 \times 1 + 15 \times 6 + \dots + 40 \times 3$. These are the products of the frequencies and the scores; in short of f 's and x 's.

The product of first frequency and first score is 14×1 , which we write as $f_1 \times x_1$.

The product of first frequency and first score is 15×2 , which we write as $f_2 \times x_2$.

In general, the product of i^{th} frequency and i^{th} score is written as $f_i \times x_i$.

So, using the letter Σ , the sum of the products $f_1 \times x_1 + f_2 \times x_2 + \dots + f_i \times x_i$ is in short written as $\Sigma f_i \times x_i$

Marks (x_i)	No. of students (f_i)	$f_i \times x_i$
14	1	$14 \times 1 = 14$
15	2	$15 \times 2 = \dots$
16	5	$16 \times \dots = \dots$
17	2	$17 \times 2 = 34$
20	3	$\dots \times 3 = \dots$
23	2	$23 \times 2 = \dots$
25	3	$25 \times 3 = \dots$
30	3	$\dots \times \dots = \dots$
35	2	$35 \times 2 = 70$
36	2	$\dots \times \dots = \dots$
37	4	$\dots \times \dots = \dots$
39	3	$39 \times 3 = 117$
40	3	$\dots \times \dots = 120$
	N= <input type="text"/>	$\Sigma f_i x_i = 956$

$$\begin{aligned} \bar{x} &= \frac{\Sigma f_i x_i}{N} = \frac{956}{35} \\ &= 27.31 \text{ marks (approximately)} \\ \therefore \text{ mean of the given data is } 27.31. \end{aligned}$$

(2) **Median** : The scores are arranged in ascending or descending order. The number appearing exactly at the middle position in this order is known as 'Median' of the observations.

If the number of observations is even then the median is the average of the middle two numbers.

Ex. (1) Find the median of 54, 63, 66, 72, 98, 87, 92.

Solution : Let us write the given observations in the ascending order.

54, 63, 66, 72, 78, 87, 92

Here the 4th number is at the middle position, which is 72

\therefore Median of the scores = 72

Ex. (2) Find the median of the data. 30, 25, 32, 23, 42, 36, 40, 33, 21, 43

Solution : Let us write the given observations in the ascending order.

21, 23, 25, 30, 32, 33, 36, 40, 42, 43

Here number of observations = 10 which is an even number.

\therefore the 5th and 6th numbers are in the middle position.

Those numbers are 32 and 33

$$\therefore \text{median} = \frac{32 + 33}{2} = \frac{65}{2} = 32.5$$



Let's recall.

If the number of observations is 'n' and

(i) if 'n' is odd, which observation is the median of the data ?

(ii) if 'n' is even, the average of which two numbers is the median ?

(3) **Mode** : The score which is repeated maximum number of times in the given data is known as the 'mode' of the data.

Ex. (1) Find the mode of 90, 55, 67, 55, 75, 75, 40, 35, 55, 95

Solution : If the data is arranged in ascending order, it is easy to find the observation repeating maximum number of times.

Ascending order of given data. 35, 40, 55, 55, 55, 67, 75, 75, 90, 95

The observation repeated maximum number of times = 55.

\therefore mode for the given data is 55.

Ex (2) The ages of workers in a certain factory are given in the following table.

Age (Year)	19	21	25	27	30
Workers	5	15	13	15	7

Find the mode of their ages.

Solution : Here the maximum frequency is 15; but this is the frequency is of two observations.

\therefore Mode = 21 and 29

\therefore mode for ages is 21 years and 29 years.

Practice set 7.5

- (1) Yield of soyabean per acre in quintal in Mukund's field for 7 years was 10, 7, 5, 3, 9, 6, 9. Find the mean of yield per acre.
- (2) Find the median of the observations, 59, 75, 68, 70, 74, 75, 80.
- (3) The marks (out of 100) obtained by 7 students in Mathematics' examination are given below. Find the mode for these marks.
99, 100, 95, 100, 100, 60, 90
- (4) The monthly salaries in rupees of 30 workers in a factory are given below.
5000, 7000, 3000, 4000, 4000, 3000, 3000, 3000, 8000, 4000,
4000, 9000, 3000, 5000, 5000, 4000, 4000, 3000, 5000, 5000,
6000, 8000, 3000, 3000, 6000, 7000, 7000, 6000, 6000, 4000
From the above data find the mean of monthly salary.
- (5) In a basket there are 10 tomatoes. The weight of each of these tomatoes in grams is as follows 60, 70, 90, 95, 50, 65, 70, 80, 85, 95.
Find the median of the weights of tomatoes.
- (6) A hockey player has scored following number of goals in 9 matches.
5, 4, 0, 2, 2, 4, 4, 3, 3. Find the mean, median and mode of the data.
- (7) The calculated mean of 50 observations was 80. It was later discovered that observation 19 was recorded by mistake as 91. What was the correct mean?
- (8) Following 10 observations are arranged in ascending order as follows.
2, 3, 5, 9, $x + 1$, $x + 3$, 14, 16, 19, 20
If the median of the data is 11, find the value of x .
- (9) The mean of 35 observations is 20, out of which mean of first 18 observations is 15 and mean of last 18 observation is 25. Find the 18th observation.
- (10) The mean of 5 observations is 50. One of the observations was removed from the data, hence the mean became 45. Find the observation which was removed.
- (11) There are 40 students in a class, out of them 15 are boys. The mean of marks obtained by boys is 33 and that for girls is 35. Find out the mean of all students in the class.
- (12) The weights of 10 students (in kg) are given below :
40, 35, 42, 43, 37, 35, 37, 37, 42, 37. Find the mode of the data.
- (13) In the following table, the information is given about the number of families and the siblings in the families less than 14 years of age. Find the mode of the data.

No. of siblings	1	2	3	4
Families	15	25	5	5

- (14) Find the mode of the following data.

Marks	35	36	37	38	39	40
No. of students	09	07	09	04	04	02

Which is a suitable measure of Central Tendency ? The answer of this question is related to the purpose of the survey.

For example, the number of runs scored by a player in continuous 11 matches are 41, 58, 35, 80, 23, 12, 63, 48, 107, 9, 73 respectively. To find his overall performance we have to consider the runs he scored in each match. Hence, the suitable measure in this example is mean.

If a company has to decide, the number of shirts to be manufactured 'different sizes'. For this out of 34, 36, 38, 40, 42, 44, which size shirts are used by maximum customers are to be found. By observation it can be found. Here the mode is useful to decide, how many shirts to be manufactured of each size.

Problem Set 7

(1) Write the correct alternative answer for each of the following questions.

- (i) Which of the following data is not primary ?
 - (A) By visiting a certain class, gathering information about attendance of students.
 - (B) By actual visit to homes, to find number of family members.
 - (C) To get information regarding plantation of soyabean done by each farmer from the village Talathi.
 - (D) Review the cleanliness status of canals by actually visiting them.
- (ii) What is the upper class limit for the class 25-35 ?
 - (A) 25 (B) 35 (C) 60 (D) 30
- (iii) What is the class-mark of class 25-35 ?
 - (A) 25 (B) 35 (C) 60 (D) 30
- (iv) If the classes are 0-10, 10-20, 20-30... then in which class should the observation 10 be included ?
 - (A) 0-10 (B) 10-20 (C) 0-10 and 10-20 in these 2 classes (D) 20-30
- (v) If \bar{x} is the mean of x_1, x_2, \dots, x_n and \bar{y} is the mean of y_1, y_2, \dots, y_n and \bar{z} is the mean of $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n$ then $\bar{z} = ?$
 - (A) $\frac{\bar{x} + \bar{y}}{2}$ (B) $\bar{x} + \bar{y}$ (C) $\frac{\bar{x} + \bar{y}}{n}$ (D) $\frac{\bar{x} + \bar{y}}{2n}$
- (vi) The mean of five numbers is 80, out of which mean of 4 numbers is 46, find the 5th number :
 - (A) 4 (B) 20 (C) 434 (D) 66
- (vii) Mean of 100 observations is 40. The 9th observation is 30. If this is replaced by 70 keeping all other observations same, find the new mean.
 - (A) 40.6 (B) 40.4 (C) 40.3 (D) 40.7
- (viii) What is the mode of 19, 19, 15, 20, 25, 15, 20, 15?
 - (A) 15, (B) 20 (C) 19 (D) 25

(ix) What is the median of 7, 10, 7, 5, 9, 10 ?

(A) 7 (B) 9 (C) 8 (D) 10

(x) From following table, what is the cumulative frequency of less than type for the class 30-40 ?

Class	0-10	10-20	20-30	30-40	40-50
Frequency	7	3	12	13	2

(A) 13 (B) 15 (C) 35 (D) 22

(2) The mean salary of 20 workers is Rs.10,250. If the salary of office superintendent is added, the mean will increase by Rs.750. Find the salary of the office superintendent.

(3) The mean of nine numbers is 77. If one more number is added to it then the mean increases by 5. Find the number added in the data.

(4) The monthly maximum temperature of a city is given in degree celcius in the following data. By taking suitable classes, prepare the grouped frequency distribution table

29.2, 29.0, 28.1, 28.5, 32.9, 29.2, 34.2, 36.8, 32.0, 31.0,
30.5, 30.0, 33, 32.5, 35.5, 34.0, 32.9, 31.5, 30.3, 31.4,
30.3, 34.7, 35.0, 32.5, 33.5, 29.0, 29.5, 29.9, 33.2, 30.2

From the table, answer the following questions.

(i) For how many days the maximum temperature was less than 34°C ?

(ii) For how many days the maximum temperature was 34°C or more than 34°C ?

(5) If the mean of the following data is 20.2, then find the value of p .

x_i	10	15	20	25	30
f_i	6	8	p	10	6

(6) There are 68 students of 9th standard from model Highschool, Nandpur. They have scored following marks out of 80, in written exam of mathematics.

70, 50, 60, 66, 45, 46, 38, 30, 40, 47, 56, 68,
80, 79, 39, 43, 57, 61, 51, 32, 42, 43, 75, 43,
36, 37, 61, 71, 32, 40, 45, 32, 36, 42, 43, 55,
56, 62, 66, 72, 73, 78, 36, 46, 47, 52, 68, 78,
80, 49, 59, 69, 65, 35, 46, 56, 57, 60, 36, 37,
45, 42, 70, 37, 45, 66, 56, 47

By taking classes 30-40, 40-50, prepare the less than type cumulative frequency table
Using the table, answer the following questions :

(i) How many students, have scored marks less than 80 ?

(ii) How many students have scored marks less than 40 ?

(iii) How many students have scored marks less than 60 ?

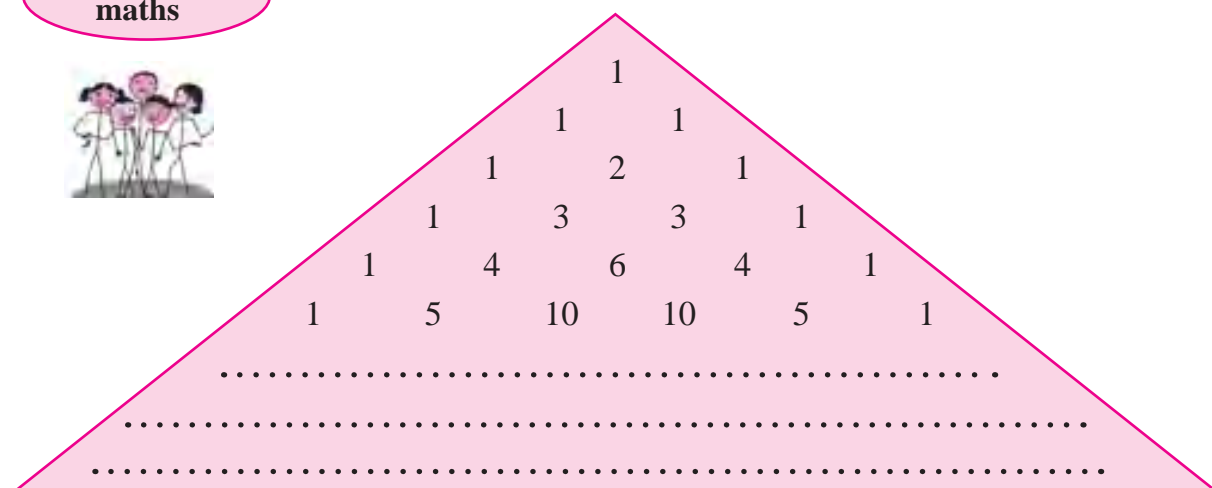
- (7) By using data in example (6), and taking classes 30-40, 40-50... prepare equal to or more than type cumulative frequency table and answer the following questions based on it.
- How many students have scored marks 70 or more than 70 ?
 - How many students have scored marks 30 or more than 30 ?
- (8) There are 10 observations arranged in ascending order as given below.
45, 47, 50, 52, x , $x+2$, 60, 62, 63, 74. The median of these observations is 53.
Find the value of x . Also find the mean and the mode of the data.



**Fun with
maths**



Pascal's Triangle (Meru prastar)



This arrangement is known as Pascal's triangle. Write the remaining 3 lines of above arrangement. The numbers obtained in above arrangement in horizontal lines denote the coefficients of the expansion of $(x + y)^n$. See the following expansions :

$$(x + y)^0 = 1$$

$$(x + y)^1 = 1x + 1y$$

$$(x + y)^2 = 1x^2 + 2xy + 1y^2$$

$$(x + y)^3 = 1x^3 + 3x^2y + 3xy^2 + y^3$$

$$(x + y)^4 = 1x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + 1y^4$$

Observe the degrees of x and y in the above expansions.

Try to write the expansion of $(x + y)^{10}$ from above arrangement.

- (7) Savings of Ramesh is ₹ 48000 ; Savings of Suresh is ₹ 51000 ; Savings of Priti is ₹ 36000
 (8) (i) ₹ 213000 (ii) ₹ 7500 (iii) No tax.

7. Statistics

Practice set 7.2

- (1) Primary data : (i), (iii), (v) Secondary data : (ii), (iv)

Practice set 7.3

- (1) Lower limit of class = 20, Upper limit of class = 25 (2) 37.5 (3) 7–13

Practice set 7.4

- (3) (i) 38 (ii) 3 (iii) 19 (iv) 62 (4) (i) 24 (ii) 3 (iii) 43 (iv) 43

Practice set 7.5

- (1) 7 quintal (2) 74 (3) 100 (4) ₹ 4900 (5) 75 gram
 (6) Mean = 3, Median = 3, Mode = 4 (7) 78.56 (8) $x = 9$ (9) 20 (10) 70
 (11) 34.25 (12) 37 kg (13) 2 (14) 35 and 37

Problem set 7

- (1) (i) C (ii) B (iii) D (iv) B (v) A (vi) D
 (vii) B (viii) A (ix) C (x) C
 (2) ₹ 26000 (3) ₹ 127
 (4) (i) 24 (ii) 06
 (5) $p = 20$
 (6) (i) 66 (ii) 14 (iii) 45
 (7) (i) 11 (ii) 68
 (8) $x = 52$, Mean = 55.9, Mode = 52

