CHAPTER

5

Data Handling

0852CH05

5.1 Looking for Information

In your day-to-day life, you might have come across information, such as:

(a) Runs made by a batsman in the last 10 test matches.

(b) Number of wickets taken by a bowler in the last 10 ODIs.

(c) Marks scored by the students of your class in the Mathematics unit test.

(d) Number of story books read by each of your friends etc.

The information collected in all such cases is called **data**. Data is usually collected in the context of a situation that we want to study. For example, a teacher may like to know the average height of students in her class. To find this, she will write the heights of all the students in her class, organise the data in a systematic manner and then interpret it accordingly.

Sometimes, data is represented **graphically** to give a clear idea of what it represents. Do you remember the different types of graphs which we have learnt in earlier classes?

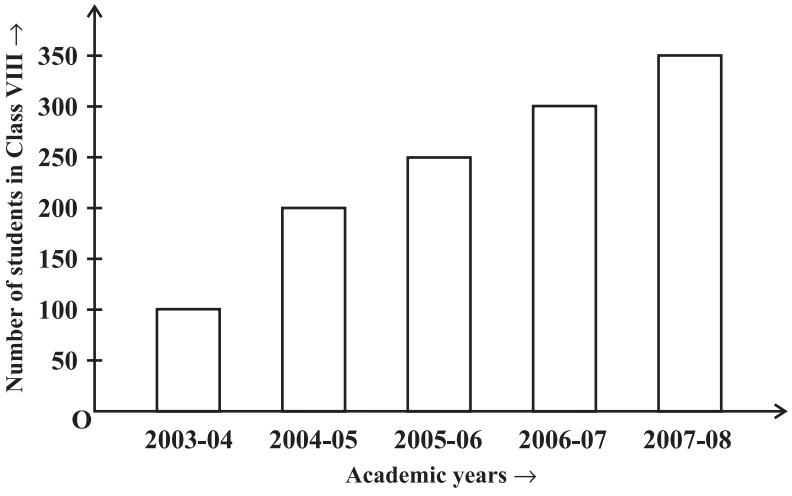
1. A Pictograph: Pictorial representation of data using symbols.

 = 100 cars ← One symbol stands for 100 cars		
July	  	= 250  denotes $\frac{1}{2}$ of 100
August	  	= 300
September	   	= ?

(i) How many cars were produced in the month of July?

(ii) In which month were maximum number of cars produced?

2. **A bar graph:** A display of information using bars of uniform width, their heights being proportional to the respective values.



Bar heights give the quantity for each category.

Bars are of equal width with equal gaps in between.

(i) What is the information given by the bar graph?

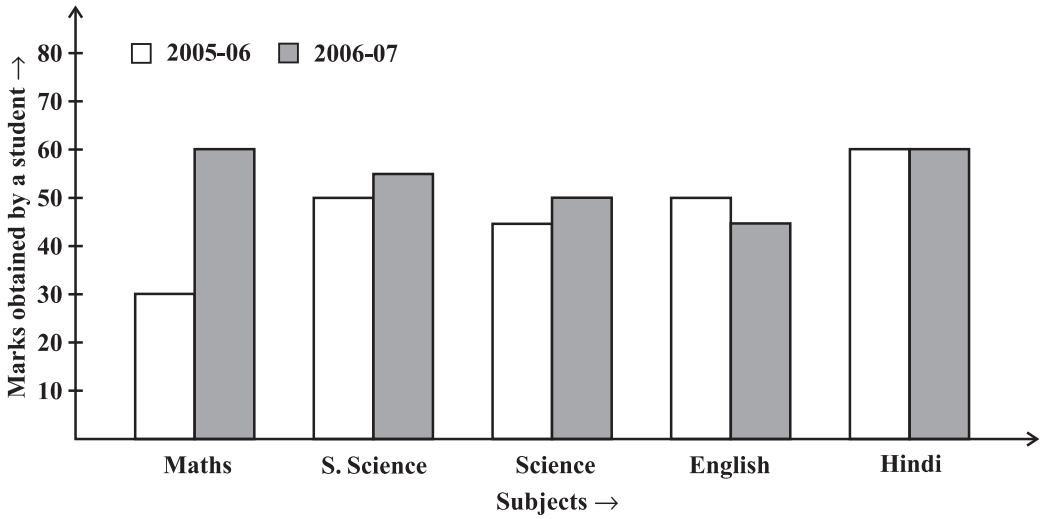
(ii) In which year is the increase in the number of students maximum?

(iii) In which year is the number of students maximum?

(iv) State whether true or false:

'The number of students during 2005-06 is twice that of 2003-04.'

3. **Double Bar Graph:** A bar graph showing two sets of data simultaneously. It is useful for the comparison of the data.



(i) What is the information given by the double bar graph?

(ii) In which subject has the performance improved the most?

(iii) In which subject has the performance deteriorated?

(iv) In which subject is the performance at par?

THINK, DISCUSS AND WRITE

If we change the position of any of the bars of a bar graph, would it change the information being conveyed? Why?

TRY THESE

Draw an appropriate graph to represent the given information.

1.

Month	July	August	September	October	November	December
Number of watches sold	1000	1500	1500	2000	2500	1500

2.

Children who prefer	School A	School B	School C
Walking	40	55	15
Cycling	45	25	35

3. Percentage wins in ODI by 8 top cricket teams.

Teams	From Champions Trophy to World Cup-06	Last 10 ODI in 07
South Africa	75%	78%
Australia	61%	40%
Sri Lanka	54%	38%
New Zealand	47%	50%
England	46%	50%
Pakistan	45%	44%
West Indies	44%	30%
India	43%	56%

5.2 Organising Data

Usually, data available to us is in an unorganised form called **raw data**. To draw meaningful inferences, we need to organise the data systematically. For example, a group of students was asked for their favourite subject. The results were as listed below:

Art, Mathematics, Science, English, Mathematics, Art, English, Mathematics, English, Art, Science, Art, Science, Science, Mathematics, Art, English, Art, Science, Mathematics, Science, Art.

Which is the most liked subject and the one least liked?

It is not easy to answer the question looking at the choices written haphazardly. We arrange the data in Table 5.1 using tally marks.

Table 5.1

Subject	Tally Marks	Number of Students
Art	𝍸 II	7
Mathematics	𝍸	5
Science	𝍸 I	6
English	IIII	4

The number of tallies before each subject gives the number of students who like that particular subject.

This is known as the **frequency** of that subject.

Frequency gives the number of times that a particular entry occurs.

From Table 5.1, Frequency of students who like English is 4

Frequency of students who like Mathematics is 5

The table made is known as **frequency distribution table** as it gives the number of times an entry occurs.

TRY THESE

1. A group of students were asked to say which animal they would like most to have as a pet. The results are given below:
dog, cat, cat, fish, cat, rabbit, dog, cat, rabbit, dog, cat, dog, dog, dog, cat, cow, fish, rabbit, dog, cat, dog, cat, cat, dog, rabbit, cat, fish, dog.
Make a frequency distribution table for the same.

5.3 Grouping Data

The data regarding choice of subjects showed the occurrence of each of the entries several times. For example, Art is liked by 7 students, Mathematics is liked by 5 students and so on (Table 5.1). This information can be displayed graphically using a pictograph or a bargraph. Sometimes, however, we have to deal with a large data. For example, consider the following marks (out of 50) obtained in Mathematics by 60 students of Class VIII:

21, 10, 30, 22, 33, 5, 37, 12, 25, 42, 15, 39, 26, 32, 18, 27, 28, 19, 29, 35, 31, 24, 36, 18, 20, 38, 22, 44, 16, 24, 10, 27, 39, 28, 49, 29, 32, 23, 31, 21, 34, 22, 23, 36, 24, 36, 33, 47, 48, 50, 39, 20, 7, 16, 36, 45, 47, 30, 22, 17.

If we make a frequency distribution table for each observation, then the table would be too long, so, for convenience, we make groups of observations say, 0-10, 10-20 and so on, and obtain a frequency distribution of the number of observations falling in each

group. Thus, the frequency distribution table for the above data can be.

Table 5.2

Groups	Tally Marks	Frequency
0-10	II	2
10-20	𝍸 𝍸	10
20-30	𝍸 𝍸 𝍸 𝍸 I	21
30-40	𝍸 𝍸 𝍸 IIII	19
40-50	𝍸 II	7
50-60	I	1
	Total	**60**

Data presented in this manner is said to be **grouped** and the distribution obtained is called **grouped frequency distribution**. It helps us to draw meaningful inferences like –

(1) Most of the students have scored between 20 and 40.

(2) Eight students have scored more than 40 marks out of 50 and so on.

Each of the groups 0-10, 10-20, 20-30, etc., is called a **Class Interval** (or briefly a class).

Observe that 10 occurs in both the classes, i.e., 0-10 as well as 10-20. Similarly, 20 occurs in classes 10-20 and 20-30. But it is not possible that an observation (say 10 or 20) can belong simultaneously to two classes. To avoid this, we adopt the convention that the common observation will belong to the higher class, i.e., 10 belongs to the class interval 10-20 (and not to 0-10). Similarly, 20 belongs to 20-30 (and not to 10-20). In the class interval, 10-20, 10 is called the **lower class limit** and 20 is called the **upper class limit**. Similarly, in the class interval 20-30, 20 is the lower class limit and 30 is the upper class limit. Observe that the difference between the upper class limit and lower class limit for each of the class intervals 0-10, 10-20, 20-30 etc., is equal, (10 in this case). This difference between the upper class limit and lower class limit is called the **width** or **size** of the class interval.

TRY THESE

1. Study the following frequency distribution table and answer the questions given below.

Frequency Distribution of Daily Income of 550 workers of a factory

Table 5.3

Class Interval (Daily Income in ₹)	Frequency (Number of workers)
100-125	45
125-150	25

150-175	55
175-200	125
200-225	140
225-250	55
250-275	35
275-300	50
300-325	20
Total	**550**

(i) What is the size of the class intervals?
(ii) Which class has the highest frequency?
(iii) Which class has the lowest frequency?
(iv) What is the upper limit of the class interval 250-275?
(v) Which two classes have the same frequency?

2. Construct a frequency distribution table for the data on weights (in kg) of 20 students of a class using intervals 30-35, 35-40 and so on.

40, 38, 33, 48, 60, 53, 31, 46, 34, 36, 49, 41, 55, 49, 65, 42, 44, 47, 38, 39.

5.3.1 Bars with a difference

Let us again consider the grouped frequency distribution of the marks obtained by 60 students in Mathematics test. (Table 5.4)

Table 5.4

Class Interval	Frequency
0-10	2
10-20	10
20-30	21
30-40	19
40-50	7
50-60	1
Total	**60**

This is displayed graphically as in the adjoining graph (Fig 5.1).

Is this graph in any way different from the bar graphs which you have drawn in Class VII? Observe that, here we have represented the groups of observations (i.e., class intervals)

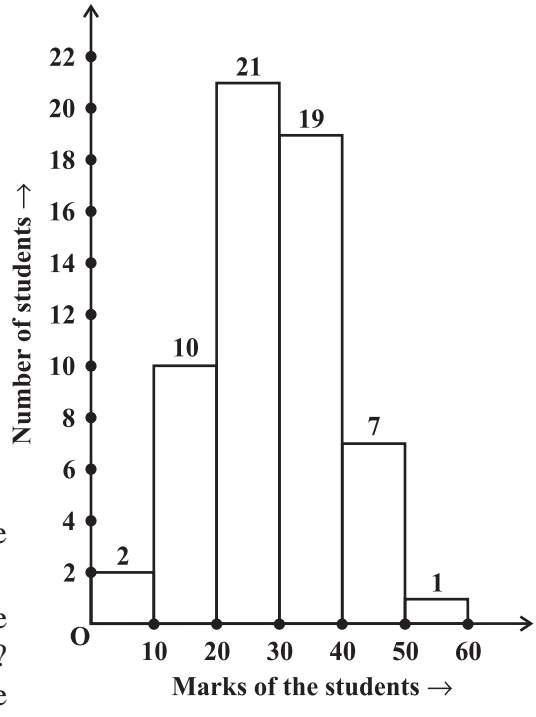


Fig 5.1

on the horizontal axis. The **height** of the bars show the **frequency** of the class-interval. Also, there is no gap between the bars as there is no gap between the class-intervals.

The graphical representation of data in this manner is called a **histogram**.
The following graph is another histogram (Fig 5.2).

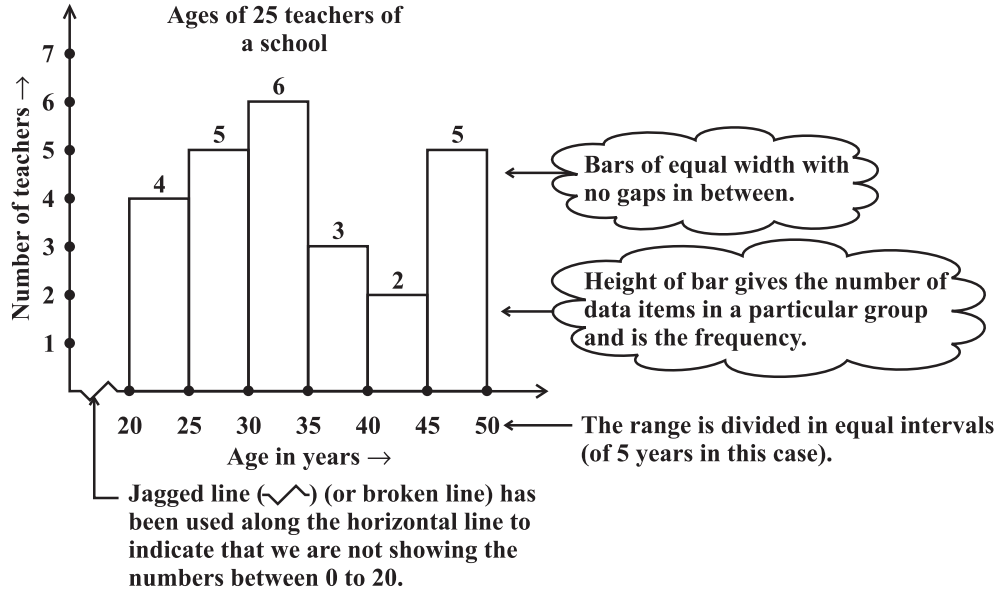


Fig 5.2

From the bars of this histogram, we can answer the following questions:

(i) How many teachers are of age 45 years or more but less than 50 years?

(ii) How many teachers are of age less than 35 years?

TRY THESE

1. Observe the histogram (Fig 5.3) and answer the questions given below.

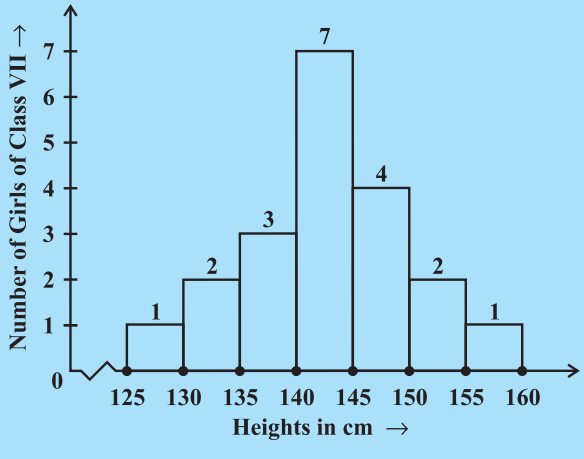


Fig 5.3

(i) What information is being given by the histogram?

(ii) Which group contains maximum girls?

(iii) How many girls have a height of 145 cms and more?
(iv) If we divide the girls into the following three categories, how many would there be in each?

150 cm and more	— Group A
140 cm to less than 150 cm	— Group B
Less than 140 cm	— Group C

EXERCISE 5.1

1. For which of these would you use a histogram to show the data?

(a) The number of letters for different areas in a postman's bag.
(b) The height of competitors in an athletics meet.
(c) The number of cassettes produced by 5 companies.
(d) The number of passengers boarding trains from 7:00 a.m. to 7:00 p.m. at a station.

Give reasons for each.

2. The shoppers who come to a departmental store are marked as: man (M), woman (W), boy (B) or girl (G). The following list gives the shoppers who came during the first hour in the morning:

W W W G B W W M G G M M W W W W G B M W B G G M W W M M W W
W M W B W G M W W W W G W M M W W M W G W M G W M M B G G W

Make a frequency distribution table using tally marks. Draw a bar graph to illustrate it.

3. The weekly wages (in ₹) of 30 workers in a factory are.

830, 835, 890, 810, 835, 836, 869, 845, 898, 890, 820, 860, 832, 833, 855, 845, 804, 808, 812, 840, 885, 835, 835, 836, 878, 840, 868, 890, 806, 840

Using tally marks make a frequency table with intervals as 800–810, 810–820 and so on.

4. Draw a histogram for the frequency table made for the data in Question 3, and answer the following questions.

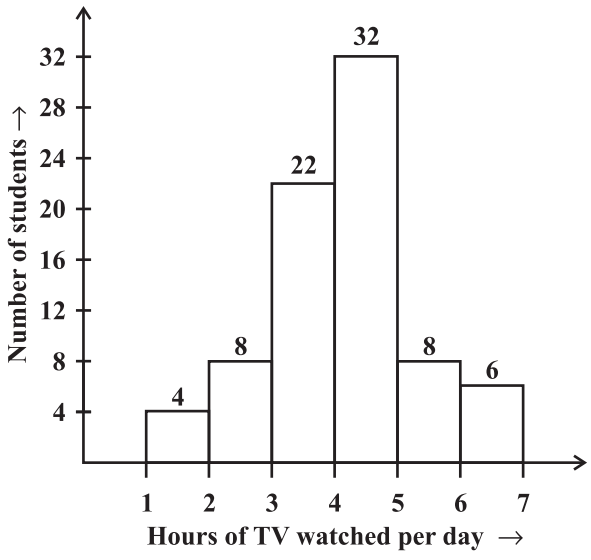
(i) Which group has the maximum number of workers?
(ii) How many workers earn ₹ 850 and more?
(iii) How many workers earn less than ₹ 850?

5. The number of hours for which students of a particular class watched television during holidays is shown through the given graph.

Answer the following.

(i) For how many hours did the maximum number of students watch TV?
(ii) How many students watched TV for less than 4 hours?

(iii) How many students spent more than 5 hours in watching TV?



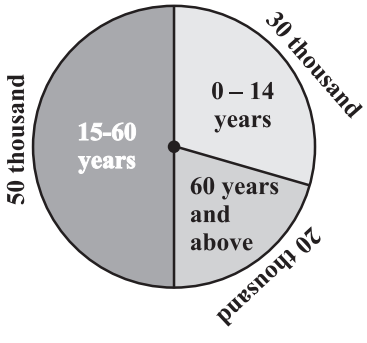
5.4 Circle Graph or Pie Chart

Have you ever come across data represented in circular form as shown (Fig 5.4)?

The time spent by a child during a day

(i)

Age groups of people in a town



(ii)

Fig 5.4

These are called **circle graphs**. A circle graph shows the relationship between a whole and its parts. Here, the whole circle is divided into sectors. The size of each sector is proportional to the activity or information it represents.

For example, in the above graph, the proportion of the sector for hours spent in sleeping

$$= \frac{\text{number of sleeping hours}}{\text{whole day}} = \frac{8 \text{ hours}}{24 \text{ hours}} = \frac{1}{3}$$

So, this sector is drawn as $\frac{1}{3}$rd part of the circle. Similarly, the proportion of the sector for hours spent in school $= \frac{\text{number of school hours}}{\text{whole day}} = \frac{6 \text{ hours}}{24 \text{ hours}} = \frac{1}{4}$

So this sector is drawn $\frac{1}{4}$ th of the circle. Similarly, the size of other sectors can be found.

Add up the fractions for all the activities. Do you get the total as one?

A circle graph is also called a **pie chart**.

TRY THESE

1. Each of the following pie charts (Fig 5.5) gives you a different piece of information about your class. Find the fraction of the circle representing each of these information.

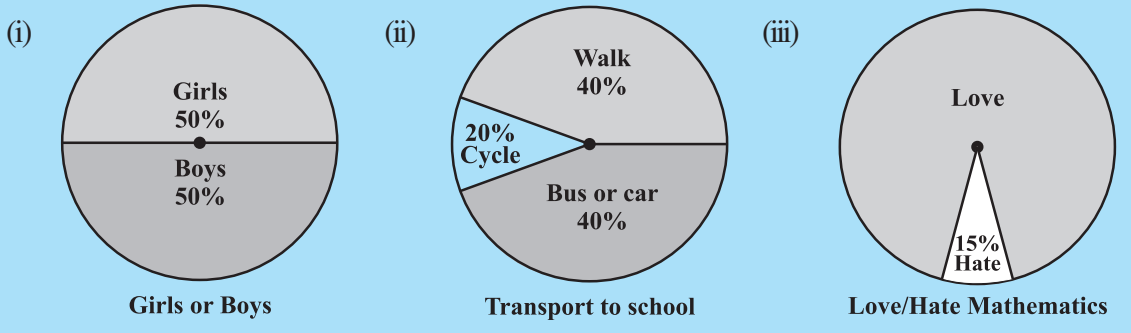
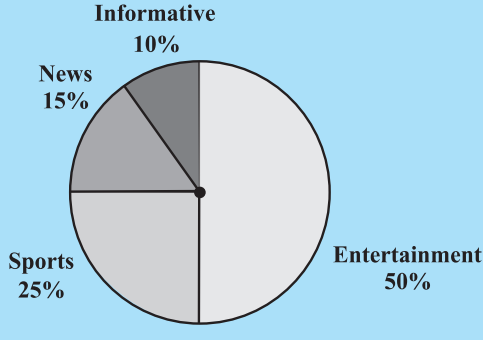


Fig 5.5

2. Answer the following questions based on the pie chart given (Fig 5.6).
   (i) Which type of programmes are viewed the most?
   (ii) Which two types of programmes have number of viewers equal to those watching sports channels?



Viewers watching different types of channels on T.V.

Fig 5.6

5.4.1 Drawing pie charts

The favourite flavours of ice-creams for students of a school is given in percentages as follows.

Flavours	Percentage of students Preferring the flavours
Chocolate	50%
Vanilla	25%
Other flavours	25%

Let us represent this data in a pie chart.

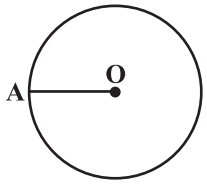
The total angle at the centre of a circle is 360°. The central angle of the sectors will be

a fraction of 360°. We make a table to find the central angle of the sectors (Table 5.5).

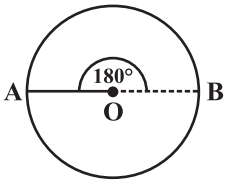
Table 5.5

Flavours	Students in per cent preferring the flavours	In fractions	Fraction of 360°
Chocolate	50%	$\frac{50}{100}=\frac{1}{2}$	$\frac{1}{2}$ of 360° = 180°
Vanilla	25%	$\frac{25}{100}=\frac{1}{4}$	$\frac{1}{4}$ of 360° = 90°
Other flavours	25%	$\frac{25}{100}=\frac{1}{4}$	$\frac{1}{4}$ of 360° = 90°

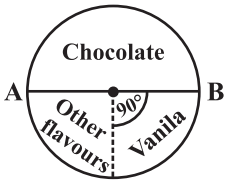
1. Draw a circle with any convenient radius.
   Mark its centre (O) and a radius (OA).



2. The angle of the sector for chocolate is 180°.
   Use the protractor to draw ∠AOB = 180°.



3. Continue marking the remaining sectors.



Example 1: Adjoining pie chart (Fig 5.7) gives the expenditure (in percentage) on various items and savings of a family during a month.

(i) On which item, the expenditure was maximum?
(ii) Expenditure on which item is equal to the total savings of the family?
(iii) If the monthly savings of the family is ₹ 3000, what is the monthly expenditure on clothes?

Solution:

(i) Expenditure is maximum on food.
(ii) Expenditure on Education of children is the same (i.e., 15%) as the savings of the family.

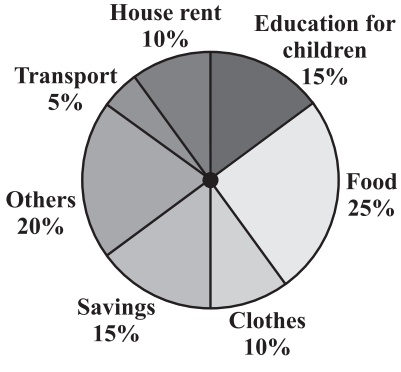


Fig 5.7