Data Handling

Data

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.

The word '**data'** means **collection of information** in the form of numerical figures, or a set of given facts.

Examples:

- The marks obtained by 10 students of a class in a test are:
  76, 83, 95, 100, 56, 32, 80, 67, 75, 46
- The following table gives the data regarding the favourite game of 100 students of a school:

Sports	Cricket	Football	Tennis	Badminton
Number Of Students	40	30	25	5

Sometimes, data is represented **graphically** to give a clear idea of what it represents.

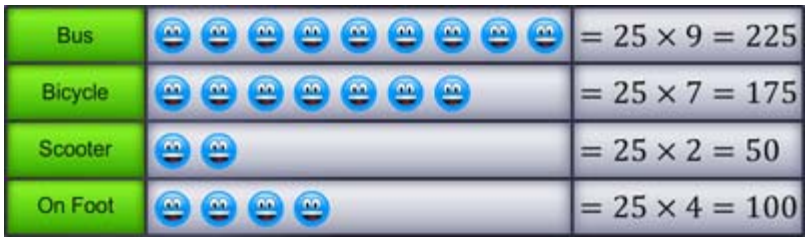
Some commonly used diagrams to **represent numerical data** are:

- A Pictograph
- A bar graph
- Double Bar Graph
- Pie-diagrams or Pie-charts

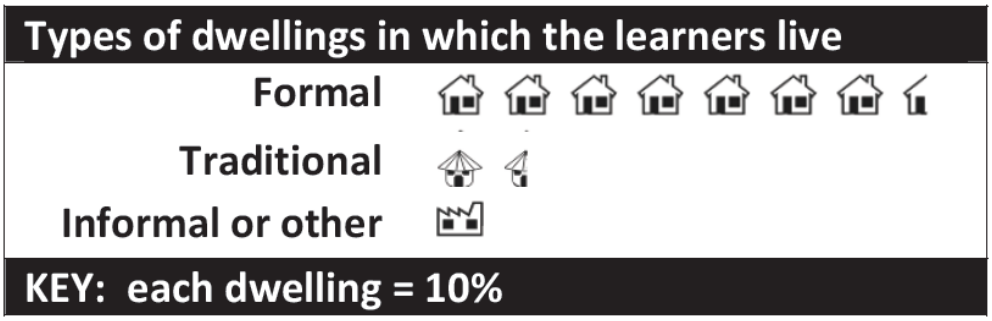
Pictograph

Pictographs represent data through appropriate **pictures**. In pictographs, the same type of **symbol** or **picture** is used to represent the data. Each symbol is used to represent a certain value, and this is mentioned clearly in the graph. For example, one symbol may represent 25 students.

The following pictograph represents the number of students coming to a college by different means of transport:



Example: Anita drew a pictograph to show the number of learners in her school living in each type of dwelling.



(a) What percentage of the learners lives in Formal Dwellings?

(b) What percentage lives in Traditional Dwellings?

Solution:

(a) The pictograph uses 7 full houses plus half a house to represent the percentage of the learners living in Formal Dwellings.

This means that 10% + 10% + 10% + 10% + 10% + 10% + 10% + 5% = 75% of learners live in Formal Dwellings.

(b) It also uses one full house plus half a house to represent the percentage of the learners living in Traditional Dwellings.
This means that 10% + 5% = 15% of the learners live in Traditional Dwellings.

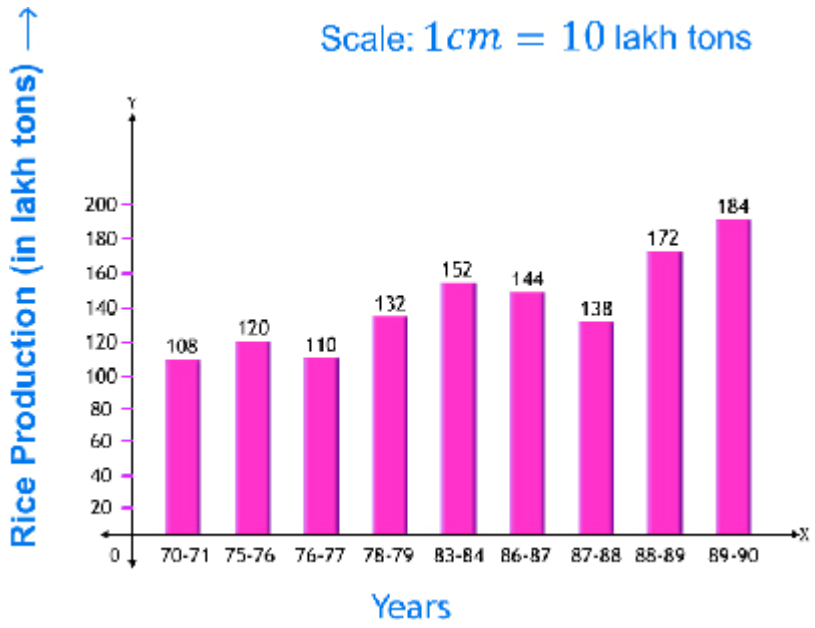
A bar graph

A **representation of data** with the help of **bars or rectangles** in a **diagram** is called a **bar graph** or a **bar diagram**.

Here, each **bar** represents only one value of the data, and hence, there are as many bars as the number of values in the data. The **length** or **height of a bar** indicates the

value of the item. The **width of a bar** and the **gap between the bars** is kept uniform to make the diagram look neat.

The following bar graph represents the production of rice in different years:



Sometimes, organising data becomes a tedious process. In such cases, we group the raw data. We write the groups as intervals. Each group is called a **class interval.** The **class interval** will have a lower class limit and an upper class limit.

The difference between the upper class limit and the lower class limit is called the width or size of the **class interval**. The number of times a particular item appears within a particular class interval is called **frequency**.

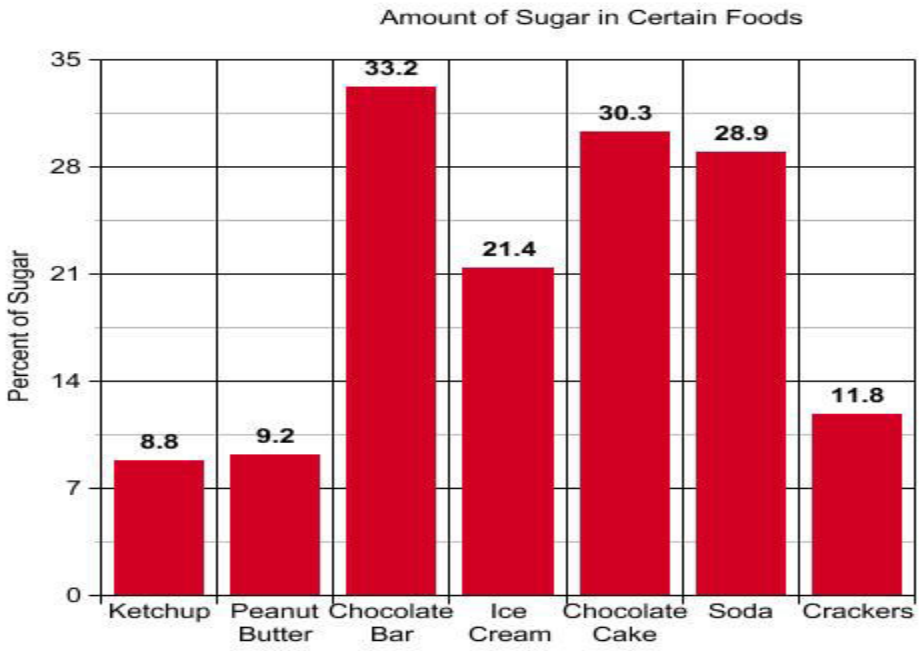
The span of a class interval is called the **width** or **size** of the **class interval**.

We fill up the rows with tally marks and will count the total number of tally marks in each group. The number of tally marks in each group is listed in the frequency column. The completed table is called the **frequency distribution table**. With the data in a table, we can draw a graph.

Class Interval (Marks)	Tally marks	Frequency (No.of students)
0-20	||||	4
20-40	𝍸 ||||	9
40-60	𝍸 𝍸 𝍸 |	16
60-80	𝍸 𝍸 ||||	14
80-100	𝍸 ||	7

FREQUENCY DISTRIBUTION TABLE

Example: The amount of sugar in 7 different foods was measured as a percent. The data is summarized in the bar graph below.



1. What is the title of this bar graph?

Solution: Amount of Sugar in Certain Foods

2. What is the range of values on the (vertical) scale?

Solution: 0 to 35

3. How many categories are in the graph?

Solution: 7

4. Which food had the highest percentage of sugar?

Solution: Chocolate Bar

5. Which food had the lowest percentage of sugar?

Solution: Ketchup

6. What percentage of sugar is in soda?

Solution: 28.9%

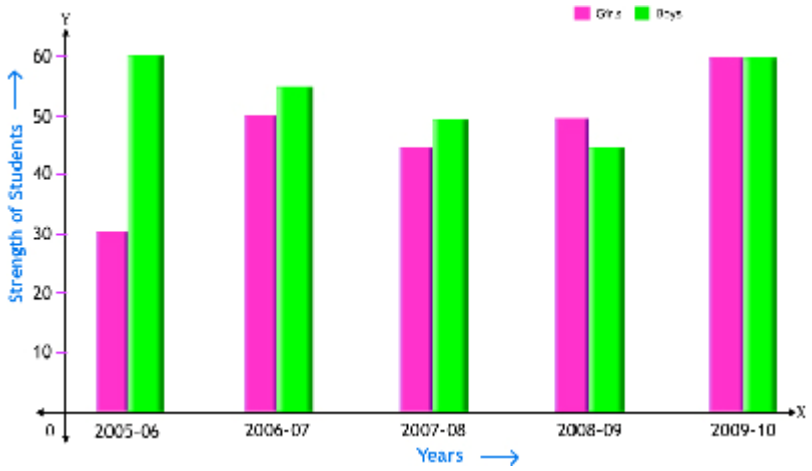
7. What is the difference in percentage of sugar between ice cream and crackers?

Solution: 21.4 - 11.8 = 9.6%

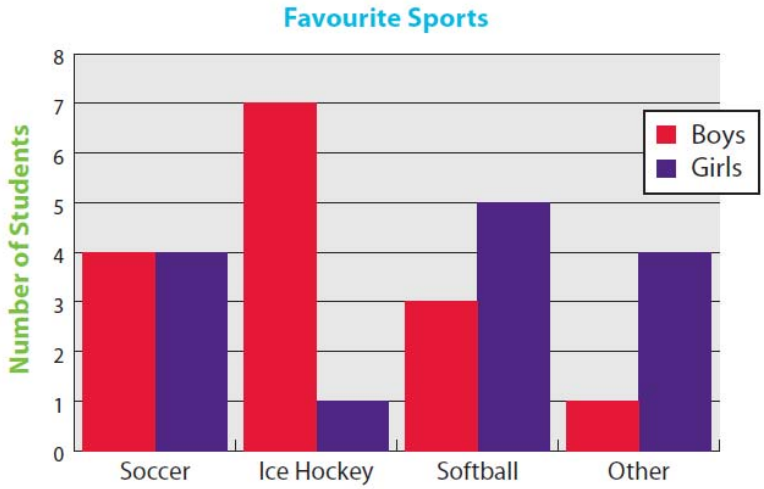
Double Bar Graph

A graph showing two sets of data simultaneously is called a **double bar graph**. It is useful for comparing two sets of data.

The following graph shows the strength of boys and girls in a school in different years:



Example: Here is Daksha's data arranged in a double bar graph.



You can see that this graph shows how many boys favour each sport and how many girls favour each sport.

Here is how you would interpret Daksha's data using the double bar graph.

Answer the following questions using the "Favorite Sports" bar graph.

A. How many students were included in the survey?

Solution: Add the number of students in each bar of the graph: 4 + 4 + 7 + 1 + 3 + 5 + 1 + 4 = 29 students

B. How many students said that softball is their favorite sport?

Solution: 8 students favour softball.

C. How many more boys like ice hockey than girls?

Solution: The number of boys that like ice hockey is 7 and the number of girls is 1. Subtract: 7 – 1 = 6

There are 6 more boys than girls that like ice hockey.

Grouping Data

Example: The marks obtained by 40 students of class VIII in an examination are given below:

16, 17, 18, 3, 7, 23, 18, 13, 10, 21, 7, 1, 13, 21, 13, 15, 19, 24, 16, 2, 23, 5, 12, 18, 8, 12, 6, 8, 16, 5, 3, 5, 0, 7, 9, 12, 20, 10, 2, 23

Divide the data into five groups, namely, 0-5, 5-10, 10-15, 15-20 and 20-25, where 0-5 means marks greater than or equal to 0 but less than 5 and similarly 5-10 means marks greater than or equal to 5 but less than 10, and so on. Prepare a frequency table for the grouped data.

Solution:

Arranging the given observations in ascending order, we get them as

0, 1, 2, 2, 3, 3, 5, 5, 5, 6, 7, 7, 7, 8, 8, 9, 10, 10, 12, 12, 12, 13, 13, 13, 15, 16, 16, 16, 17, 18, 18, 18, 19, 20, 21, 21, 23, 23, 23, 24

Thus, the frequency distribution may be given as under:

Marks	Tally Marks	Frequency
0 - 5	𝍸 I	6
5 - 10	𝍸 𝍸	10
10 - 15	𝍸 III	8
15 - 20	𝍸 IIII	9
20 - 25	𝍸 II	7
	Total	40

Note: Here, each of the groups 0-5, 5-10, 10-15, 15-20 and 20-25 is called a class interval.
In class interval 10-15, the number 10 is called the lower limit and 15 is called the upper limit of the class interval.

The difference between the upper limit and the lower limit of any class interval is called the **class size**.

Thus, the class size in the above frequency distribution is 5.

The mid value of a class is called its class mark and is obtained by adding it's upper and lower class limits and dividing the sum by 2.

Thus, the class mark of 0-5 is (0 + 5)/2 = 2.5

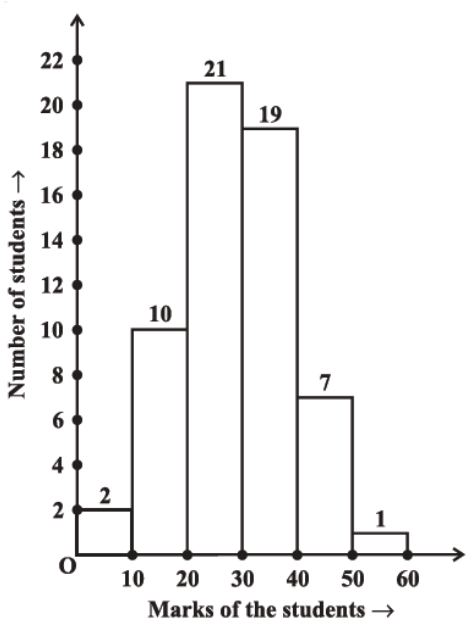
The class mark of 5-10 is (5 + 10)/2 = 7.5, etc.

Bars with a difference

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

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

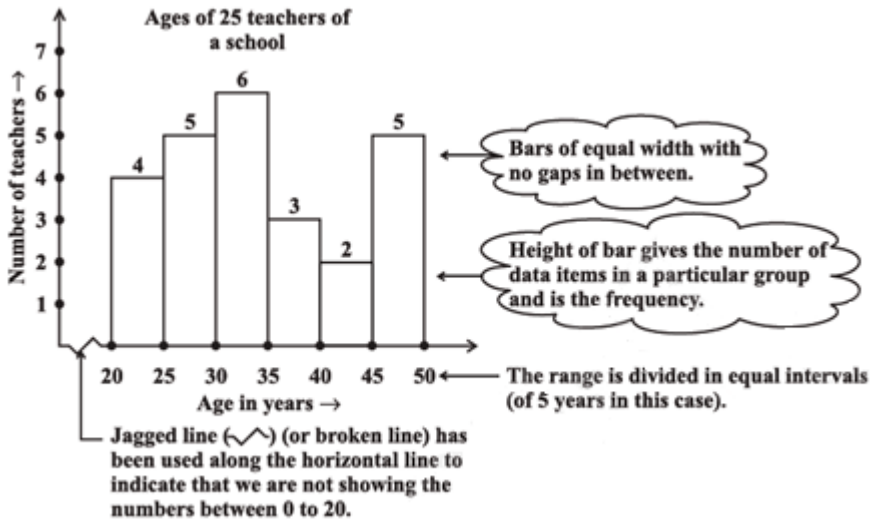


Here we have represented the groups of observations (i.e., class intervals) on the horizontal axis. The **height** of the bars shows the **frequency** of the class-interval.

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



Circle Graph or Pie Chart

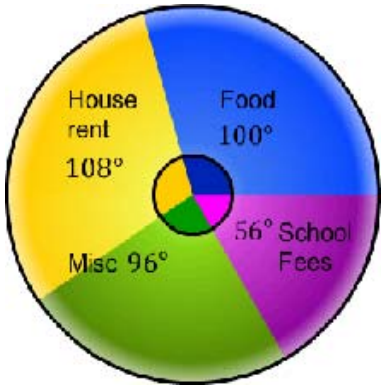
A **pie diagram** or a **pie chart** is a circle divided into several **sectors**.
The **circle** represents the total value of the given data, and the sectors represent the proportion of the components of the total.

It is also called an **angular diagram** or a **circular diagram**.

The monthly expenditure on various items of a family is given below.

Item	Food	House Rent.	Misc.	School Fees
Amount Spent	Rs. 2500	Rs. 2700	Rs. 2400	Rs. 1400

Its representation in a pie diagram is as shown.



Drawing pie charts

Example:

A family's weekly expenditure on its house mortgage, food and fuel is as follows:

Expense	$
Mortgage	300
Food	225
Fuel	75

Draw a pie chart to display the information.

Solution:

$$\text{The total weekly expenditure} = \$300 - \$225 + \$75$$
$$= \$600$$

We can find what percentage of the total expenditure each item equals.

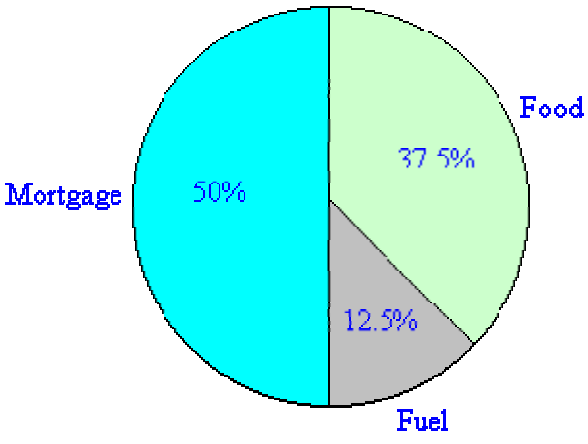
Percentage of weekly expenditure on:

$$\text{Mortgage} = \frac{300}{600} \times 100\% = 50\%$$

$$\text{Food} - \frac{225}{600} \times 100\% - 37.5\%$$

$$\text{Fuel} = \frac{75}{600} \times 100\% = 12.5\%$$

To draw a pie chart, divide the circle into 100 percentage parts. Then allocate the number of percentage parts required for each item.



Note:

- It is simple to read a pie chart. Just look at the required sector representing an item (or category) and read off the value. For example, the weekly expenditure of the family on food is 37.5% of the total expenditure measured.
- A pie chart is used to compare the different parts that make up a whole amount.

Chance and Probability

In our daily life we come across the words like probably, likely, may be, chance and hope etc. All these are synonyms to **probability**.

Probability is defined as the numerical method of measuring uncertainty involved in a situation.

It is widely used in the study of mathematics, statistics, gambling, physical science, biological science, weather forecasting, finance etc. to draw conclusions.

An **experiment** is defined as an action or process that results in well defined outcomes.

An experiment, in which we know all the results, but cannot predict them, is called a **random experiment**.

The possible results of an experiment are called the **outcomes.**

A combination of outcomes is called an **event**.

For example:

When an unbiased die is rolled getting an even number is an **event**.

In this event, the outcomes are {2, 4, and 6}.

When an experiment is performed, outcomes are said to be **equally likely**, if each outcome has the same chance of occurring.

Probability of event E is defined as:

$$\textbf{Probabilityof an event} = \frac{\textbf{No. of favourable outcomes}}{\textbf{Total number of outcomes}}$$

For example, consider the following three experiments:

Experiment: Roll a die two hundred times noting the outcomes.
Event of interest: A six faces upwards.
Trial: Roll the die once.
Number of trials: 200
Outcomes: 1, 2, 3, 4, 5 or 6

Experiment: Toss a coin seventy times noting the outcomes.
Event of interest: A tail faces upwards.
Trial: Toss the coin once.
Number of trials: 70
Outcomes: Head or Tail

Experiment: Spin a spinner one hundred times noting the outcomes.
Event of interest: The spinner stops on the number 3.
Trial: Spin the spinner once.
Number of trials: 100
Outcomes: 1, 2, 3, And 4