Chapter 5

Data Handling

Introduction to Data Handling

What is Data?

The word data means the collection of information in the form of numerical figures or a set of given facts. Data is a collection of numbers gathered to give some information.

Data is a collection of facts, such as numbers, words, measurements, observations or even just descriptions of things.

For example:

(a) Runs made by Virat Kohli in the last 25 one day (ODIs) matches.

(b) The number of wickets taken by Jasprit Bumrah in the last 20 T-20 matches.

(c)Marks obtained by the Rohan in English unit test.

Types of Data

There are two types of data.

i) Raw data

ii) Grouped data

i) Raw data

When some data is collected and presented randomly, then it is called Raw Data. A collection of observations gathered initially is called raw data. or Data obtained from direct observation is called raw data. or Data mostly available to us in an unorganized form is called raw data. The marks obtained by 10 students in a monthly test are an example of raw data also known as ungrouped data.

ii) Grouped data

When the same data are classified into groups, then it is Grouped Data. To present the data in a more meaningful way, we condense the data into a convenient number of classes or groups.

Understanding of Data

Some Terms Related to Data

Tally marks

Tally marks are the symbolic representation of the occurrence of an observation in a particular table.

Frequency

The number of times a particular observation occurs in the given data is called the frequency of the observation.

Class Interval

Large Data is classified into Grouped Data (Class Interval) 1 – 10: 1 → lower limit, 10 → upper limit.

The table below shows the height (in cm) of 12 students of class VIII

Class Interval	Tally Marks	Frequency
120 - 130	||||	4
130 - 140	𝍸	6
140 - 150	||	2
Sum		**12**

Frequency Distribution

A way of presenting data that exhibits the values of the variable and corresponding frequencies is called a frequency distribution.

For example,

A selection test was given to a group of 20 students. The test was completed by them in the following times (in minutes). 38, 40, 42, 41, 39, 27, 28, 26, 30, 42, 41, 43, 45, 46,

37, 37, 43, 44, 49, 36.

Prepare a frequency distribution table taking class interval 25 – 30, 30 – 35, 35 – 40, 40 – 45, 45 – 50.

The table is as follows:

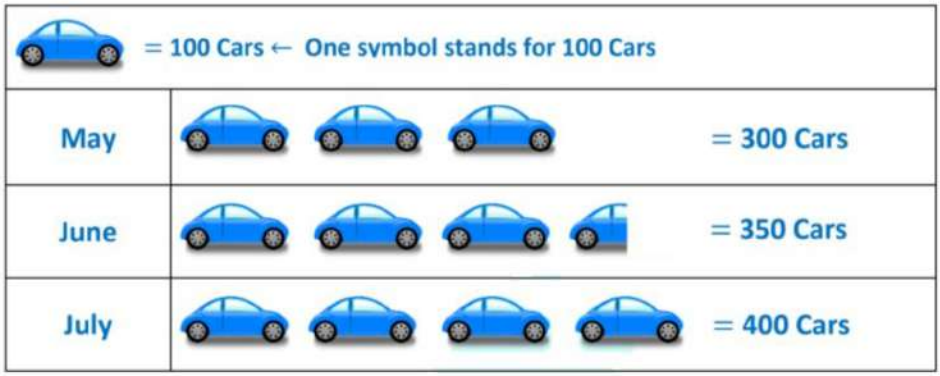
Class interval	Tally Marks	Frequency
25 – 30	|||	3
30 – 35	|	1
35 – 40	𝍸	5
40 – 45	𝍸 |||	8
45 – 50	|||	3

Common observation will belong to the higher class. For example, 30 will belong to 30 – 35 (not 25 – 30). Similarly, 45 will belong to 45 – 50 (not 40 – 45).

Pictograph

It represents data through appropriate pictures. Generally, the same type of symbols or pictures is used to represent data. Each picture and symbols are used to represent a certain value and it is clearly mentioned in the graph.

The given pictograph represents the number of cars produced in May, June, and July.



A Bar Graph

A display of information using bars of uniform widths, their heights being proportional to the respective values.

For example,

Following data gives total marks (out of 500) obtained by six children of a particular class. Represent the data on a bar graph.

Students Name	Aman	Rishabh	Rahim	Gopal	Sony	Tarun
Marks Obtained	350	400	200	260	300	440

To choose an appropriate scale we make equal divisions taking increments of 50. Thus 1 unit will represent 50 marks. Now represent the data on the bar graph.

Double Bar Graph

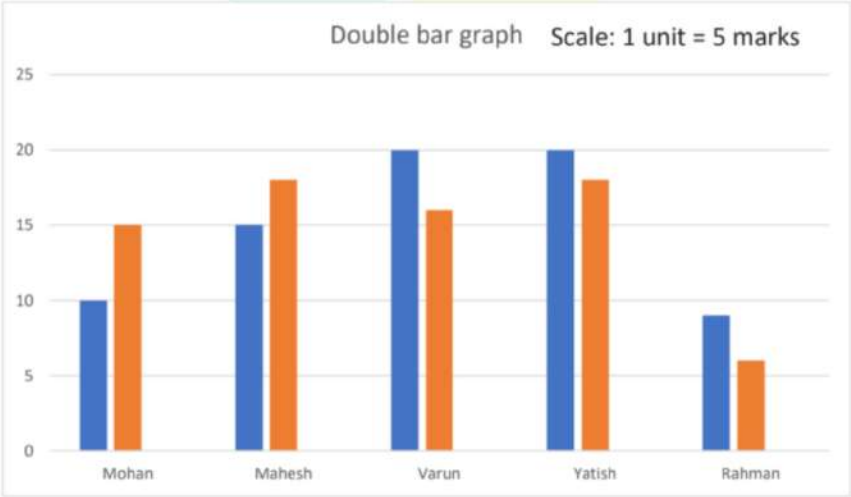
A bar graph showing two sets of data simultaneously. It is useful for the comparison of the data.

For example,

A class teacher prepared a report card of unit test-I (out of 20) and unit test-II (out of 20) of 5 students.

Students	Mohan	Mahesh	Varun	Yatish	Rahman
Unit test-I	10	15	20	20	9
Unit test-II	15	18	16	18	6

She displayed the information in the double bar graph in the parent-teacher meeting to visualize the performance of the students.



Organisation of Data

Data available to us is in an unorganized form called raw data. to draw meaningful inferences, we need to organise the data systematically.

For example,

The following data shows the number of children in 25 families:

1 1 2 3 4 3 2 1 1 4 5 2 4

2 2 1 3 3 2 5 5 3 3 4 5

Make the frequency table for the data.

Number of Children	Tally Marks	Frequency
1	𝍸	5
2	𝍸 |	6
3	𝍸 |	6
4	||||	4
5	||||	4

In the given example, we notice that the data are given in a random manner. We just organise the data on the basis of some observation like, out of 25 families 5 families have only 1 child, 6 families have only 2 children, 6 families have only three children, 4 families have only 4 children and 4 families have 5 children.

Grouping Data

The data regarding the choice of subjects showed the occurrence of each of the entries several times.

For example,

Consider the following marks (out of 50) obtained in Mathematics by 30 students:

21 10 30 22 33 5 37 12 25 42 15 39 26

32 18 27 50 19 29 35 31 24 36 18 20 38 22 44 16 24

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

Groups	Tally Marks	Frequency
0 – 10	|	1
10 – 20	|||| 𝍸	7
20 – 30	|||| 𝍸 \	10
30 – 40	|||| 𝍸	9
40 – 50	||	2
50 – 60	|	1
	Total	30

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

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

b) Three students have scored more than 40 marks out of 50 and so on.

Each of the group 0 – 10, 10 – 20, and so on, is called a Class Interval.

Histogram

What is Histogram?

The graphical representation of data in a manner, such that the height of bars shows the frequency of class interval and also there is no gap between the bars as there is

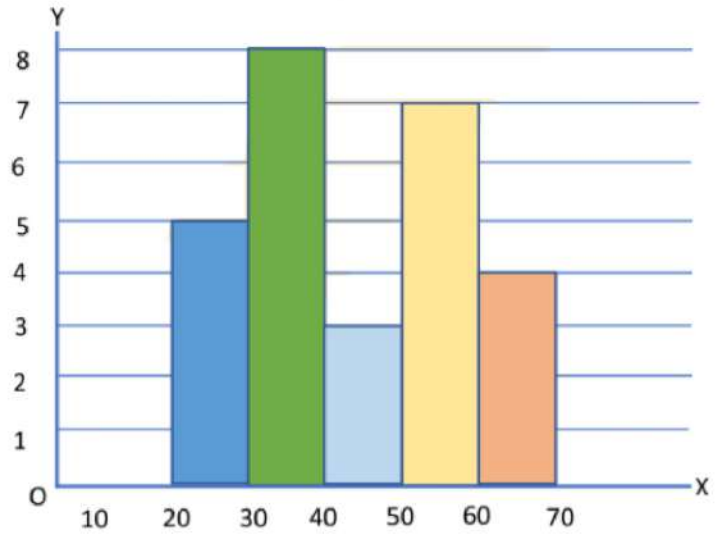
no gap between the class-intervals is called a histogram.

1) X-Axis in the histogram represents the Class Intervals.

2) Y-Axis in the histogram represents the frequencies.

Construct a histogram for the frequency distribution table below:

Class interval	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70
Frequency	5	8	3	7	4

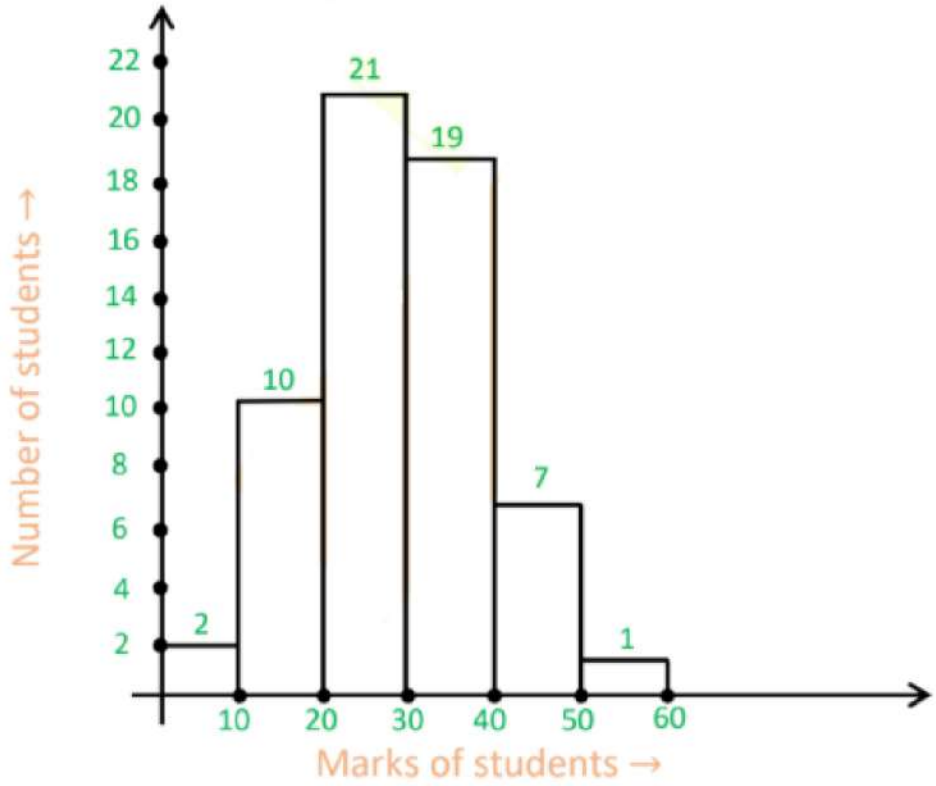


Bars with a difference

For example,

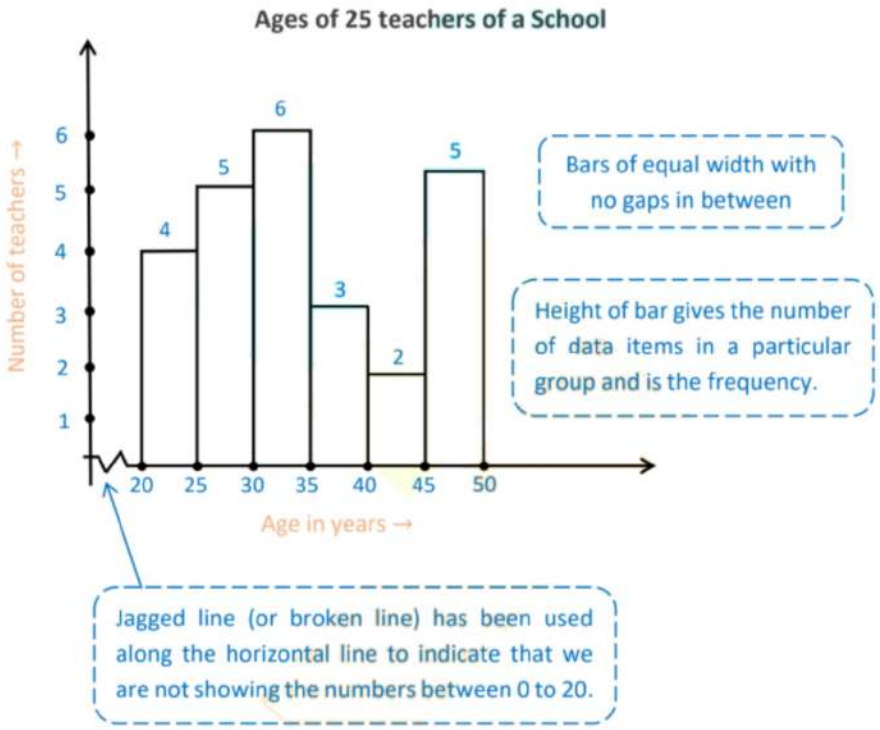
Let us consider the grouped frequency distribution of the marks obtained by 60 students in the English Unit test.

Groups	Frequency
0 – 10	2
10 – 20	10
20 – 30	21
30 – 40	19
40 – 50	7
50 – 60	1
Total	60



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.



Pie Chart

What is Pie Chart?

It is also called a circle graph. A circle graph shows the relationship between a whole and its parts. The whole circle divided into sectors. The size of each sector is proportional to the activity or information it represents.

The pie chart below shows the time spent by a child in a day.

Activity	Time
Sleep	8 hours
School	6 hours
Play	3 hours
Home Work	3 hours
Other	3 hours

In the above graph, the proportion of the sector for hours spent in sleeping.

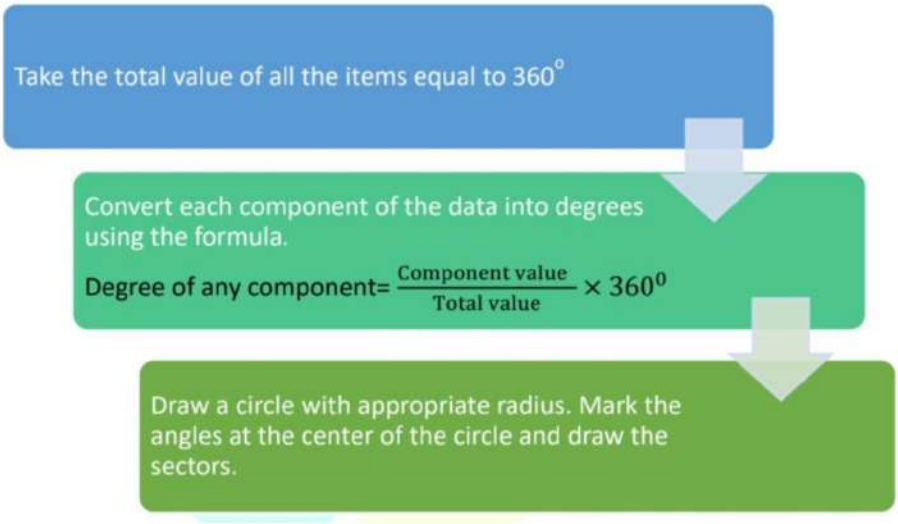
$$= \frac{number\ of\ sleeping\ hours}{Whole\ day} = \frac{8}{24} = \frac{1}{3}$$

So, this sector is drawn as 1/3rd part of the circle. Similarly, the proportion of the sector for hours spent in School

$$= \frac{number\ of\ School\ hours}{Whole\ day} = \frac{6}{24} = \frac{1}{4}$$

So, this sector is drawn as 1/4th part of the circle. Similarly, the size of other sectors can be found.

Drawing Pie Chart



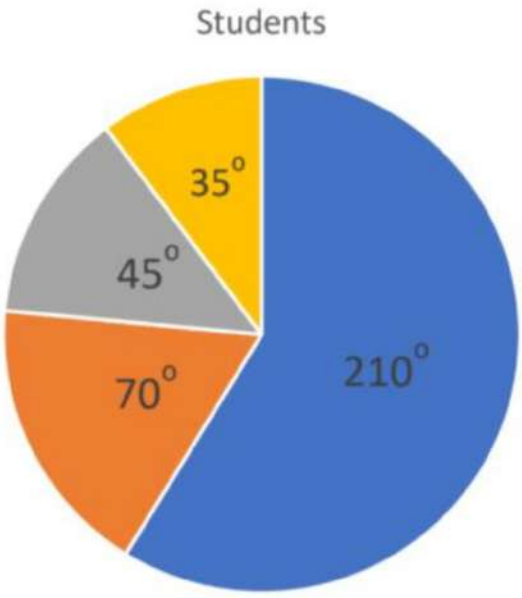
For example,

The number of students in a hostel speaking different languages is given below. Present the data in a pie chart.

Language	Hindi	English	Marathi	Tamil
Number of students	42	14	9	7

The central angle of the component = $\frac{Value\ of\ the\ components}{Sum\ of\ all\ the\ components}$ x 360°

Language	Number of students	In Fraction	Central angles
Hindi	42	$\frac{42}{72} = \frac{7}{12}$	$(\frac{7}{12} \times 360)^\circ = 210^\circ$
English	12	$\frac{14}{72} = \frac{7}{36}$	$(\frac{7}{36} \times 360)^\circ = 70^\circ$
Marathi	9	$\frac{9}{72} = \frac{1}{8}$	$(\frac{1}{8} \times 360)^\circ = 45^\circ$
Tamil	7	$\frac{7}{72}$	$(\frac{7}{72} \times 360)^\circ = 35^\circ$



Chance and Probability

Chance

There are scenarios in our life, that are certain to happen, while some that are impossible to happen and some that may or may not happen. The situation that may or

may not happen has a chance of happening.

The Chance of happening of an event may be described as one of the following:

Few terms related to probability are as follows:

a) Trial

A trial is an action that results in one or several outcomes.

For example, in throwing two coins the possible outcomes are {HH, TT, HT, TH}.

b) Random Experiments
An experiment in which the result of the trial cannot be predicted in advance.

For example, drawing a ball from a bag containing 5 red and 4 white balls.

c) Event

An event associated with a random experiment is the collection of some outcomes of the experiment.

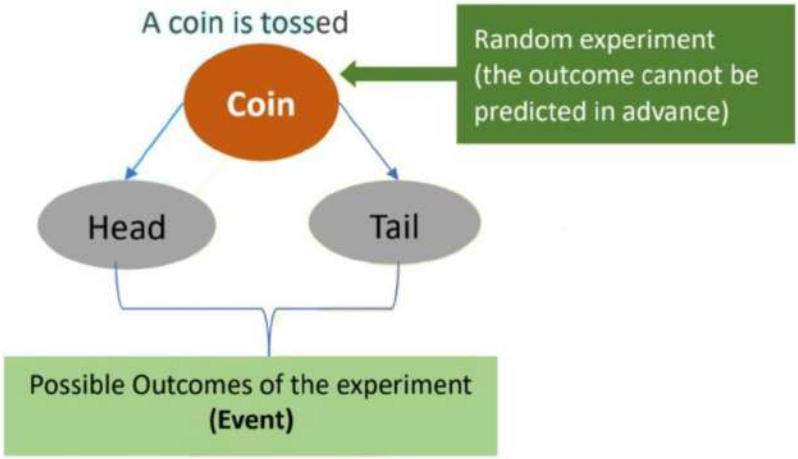
For example, on throwing a die, 'getting a prime number' is an event consisting of outcomes 2, 3, 5.

d) Equally likely outcomes

Outcomes of an experiment are equally likely if each has the same chance of occurring.

On throwing a die: 1, 2, 3, 4, 5, 6 are equally likely outcomes.

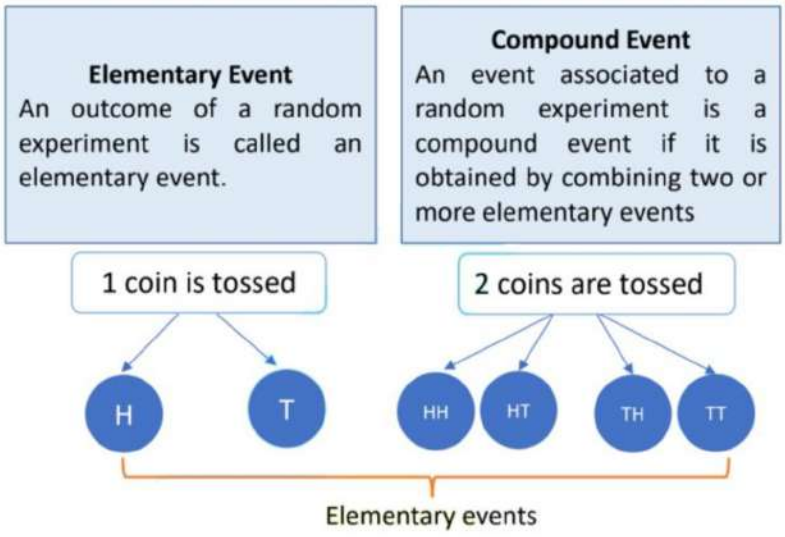
Probability

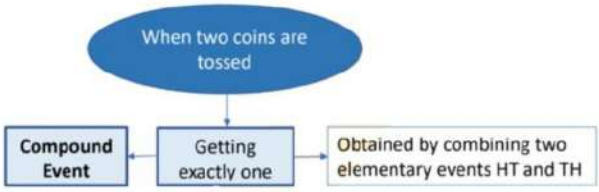


The chance of happening of an event is called the probability of the event happening. Probability tells us how likely an event is going to occur. It does not tell us what

is going to happen. We often make predictions about the chances of an event happening.

There is an even chance of an event happening if the chance of an event happening is the same as the chance of the event not happening.





Probability is nothing but the numerical value of the possibility of occurrence of an event. The value of probability lies between 0 and 1.

For example,

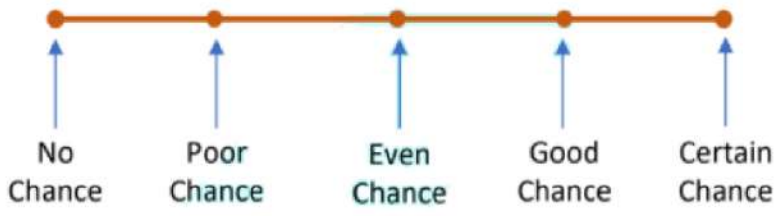
(a) If we toss a coin, then either head will come up or the tail will come up. Thus, there is an even chance of getting a head when a coin is tossed.

We say that the probability of getting Head or Tail is equal and is $\frac{1}{2}$ for each.

(b) If we throw a ball upward, then it will certainly return to the earth. Thus, the chance of happening an event can be shown on a scale with no chance at one end and

certain at other ends.

Thus, the scale is given, below:



Types of Probability

a) Experimental Probability

b) Theoretical Probability

a) Experimental Probability

Probability of an event is based on actual experiments A coin is tossed 5 times, head appears 3 times. Probability of getting a head $= \frac{3}{5}$

Probability of occurrence of an event $= \frac{Number\ of\ trials\ in\ which\ an\ event\ occurs}{Total\ number\ of\ trials}$

b) Theoretical Probability

Probability is based on what will happen without performing the actual experiment.

A box contains 4 green and 2 white balls. If a ball is picked then

Probability of getting a white ball $= \frac{2}{6} = \frac{1}{3}$

Probability of occurrence of an event $= \frac{Favourable\ number\ of\ elementary\ events}{Total\ number\ of\ elementary\ events}$

Favourable Elementary Events

An elementary event is said to be favourable to compound event A if it satisfies the definition of compound event A.

For example,

A die is thrown. What is the probability of getting?

i) a prime number?
ii) a number greater than 4?

In throwing a die, all possible outcomes are 1, 2, 3, 4, 5, 6 i.e. 6 possible outcomes

i) Prime numbers are 2, 3, 5

No. of prime numbers = 3

$$P(E) = \frac{number\ of\ favorable\ events}{total\ number\ of\ events}$$

P (getting a prime number) $= \frac{3}{6} = \frac{1}{2}$

ii) Number greater than 4 are 5, 6

$$P(E) = \frac{number\ of\ favorable\ events}{total\ number\ of\ events}$$

P (getting a number greater than 4) $= \frac{2}{6} = \frac{1}{3}$