

STATISTICS

A 3D bar chart with a red trend line and a 3D pie chart on a laptop keyboard. The bar chart has 10 bars of increasing height, colored in a gradient from light blue to orange. The pie chart is 3D and divided into five segments of different colors: red, orange, yellow, green, and blue. The background is a blurred laptop keyboard and screen.

CLASS IX
CLASS IX
MODULE 01
MODULE 01

INTRODUCTION

****The present day society is essentially information oriented. In our day to day life, we come across a lot of information in the form of facts, numerical figures, tables, graphs etc.**

****This we get from newspapers, televisions, magazines, internet, etc...In different fields of life we need information in the form of numerical figures called data.**

****This data may be related to profits of company during certain period of time, monthly wages of workers of a factory, expenditures in various sectors of a plan, batting or bowling averages of cricketers.**

DATA

- **The facts or figures which are numerical or otherwise collected with a definite process are called data.
- **Data is the plural form of the latin word datum.

FUNDAMENTALS CHARACTERISTICS OF DATA

- **Numerical figures or facts alone can represent data.
- **Qualitative characteristics like honesty intelligence poverty which cannot be expressed numerically cannot constitute data.
- **Data are collections of facts. A single observation cannot constitute data.
- **Data are collections for specific purpose. The data may not be used for some other purposes.
- **Data obtained in different expressions for a specific purpose are comparable.
- **Statistics is a branch of mathematics which deals with data with the scientific analysis of data in order to extract meaningful information from it.

STATISTICS

**It is the branch of mathematics that deals with the collection, presentation, analysis and interpretation of numerical data.

There are two types of data:
Let's take an example and understand it.

PRIMARY DATA:

The data collected by the investigator himself with a definite plan in his mind is called primary data. As primary data is collected by the investigator by personal observation according to his plan or need, it is more relevant and reliable.

SECONDARY DATA:

The data originally collected by someone else and used by the investigator for some related purpose is called secondary data. Secondary data should be used carefully as data is collected for a purpose different from that of the investigator and may not be fully relevant to the investigation in hand.

QUES 01} GIVE FIVE EXAMPLES OF DATA THAT YOU CAN COLLECT FROM YOUR DAY TO DAY LIFE.

- ** Five examples of data that we can gather from our day to day life are:
- ** Result of election held in the last year obtained from television or newspaper.
- ** Electricity bills of our school for last ten years.
- ** Production of various agricultural products in last ten years in ur country.
- ** Quantity of fuel wastage due to traffic jams in specific city.
- ** Number of vehicles in a district.

QUES 02} CLASSIFY THE DATA IN QUES 01} AS PRIMARY OR SECONDARY DATA.

Primary Data: ii} v}

Secondary Data i} iii} iv}

PRESENTATION OF DATA

After the collection of data an investigator needs to find out ways to present them in a form which is meaningful easily understood and gives its main features at a glance. Let us study the different ways of presenting the data through some illustrations.

ILLUSTRATION 1} RAW DATA

Consider the marks obtained by students in Mathematics...

11 40 48 41 62 54 53 96 40 98 44

Here searching for lower and higher scores would be less time consuming if the data is arranged in increasing or decreasing order.

ILLUSTRATION 2} ARRAY

When we arrange the above in increasing order then..

11 40 41 46 48 52 53 54 62 96 98

This data is known as array when obtained data is arranged in increasing or decreasing order.

The difference between the highest and lowest marks is known as range.

ILLUSTRATION 3} UNGROUPED FREQUENCY DISTRIBUTION TABLE

Consider the heights of 30 girls of class IX given below:

140 140 160 139 153 146 150 148 150 153
152 146 154 150 160 148 150 148 140 148
153 138 152 150 148 138 152 140 146 148

The height of girls are called **variates** and the number of girls who have particular height is called **frequency** of that variate. The term frequency implies how frequent a variate is repeated. Thus the number of times the variate is expressed is called the frequency of that variate.

In order to correctly count the number we use tally marks{1}. Then we count each tally marks and write the frequency.

The below is the ungrouped/unclassified frequency distribution table.

HEIGHT	TALLY MARKS	NUMBER OF GIRLS(FREQUENCY)
138	II	2
139	I	1
140	IIII	4
146	III	3
148	IIII I	6
150	IIII	5
152	III	3
153	III	3
154	I	1
160	II	2

**What are maximum and minimum height?

138 -- Minimum

160 -- Maximum

**How many girls are there with maximum height?

2

**Maximum girls have which height?

148

ILLUSTRATION 4} GROUP FREQUENCY DISTRIBUTION {INCLUSIVE OR DISCONTINUOUS FORM}

60 students of class IX appeared for mathematics olympiad. The marks obtained by them out of 100 are given below:

46 31 74 68 42 54 14 61 83 48
93 42 53 59 38 16 88 75 56 46
27 44 63 58 43 81 64 67 36 49
77 62 53 40 71 60 58 45 42 34
37 26 08 64 57 66 45 61 54 27
50 76 38 47 55 46 40 59 42 29

Minimum marks = 8

Maximum marks = 93

Range = $93 - 8 = 85$

Here each group is a class or class interval and its size is called class size or class width. Here we will consider 0-9,10-19.....90-99. Here least number is called lower limit and greatest number is called upper class limit.

Class intervals are inclusive because the lower and upper limits of each interval are included in that interval.

Marks Obtained	Tally Marks	Number of Students
0 - 9	I	1
10 - 19	II	2
20 - 29	IIII	4
30 - 39	IIII I	6
40 - 49	IIII IIII IIII	15
50 - 59	IIII IIII II	12
60 - 69	IIII IIII	10
70 - 79	IIII I	6
80 - 89	IIII	3
90 - 99	I	1
TOTAL		60

Here the classes are non overlapping and discontinuous because there is gaps between the upper class limits and lower class limits.

This frequency distribution is inclusive because it contains both lower limit and upper limit.

ILLUSTRATION 5} GROUP FREQUENCY DISTRIBUTION {EXCLUSIVE OR CONTINUOUS}

Weights(Kg)	Number of Students
31 - 35	9
36 - 40	5
41 - 45	14
46 - 50	3
51 - 55	1
56 - 60	2
61 - 65	2
66 - 70	1
71 - 75	1

We see that there are gaps between the class intervals, to cover this gap we have to do this into continuous.

For this we will find the difference between the upper limit and lower limit of the class. Then we will add half of this difference to the upper limits and subtract the same from the lower limits of all cases.

Therefore we will be adding 0.5 to upper class and subtract 0.5 from all lower class.

Lower Limit {36 - 40}=36

Upper limit {31 - 35}=35

Difference = 1

Half Difference = $\frac{1}{2}=0.5$

MODIFIED FREQUENCY TABLE

Weights(Kg)	Number of Students
30.5 - 35.5	9
35.5 - 40.5	$5+1=6$
40.5 - 45.5	$14+1=15$
45.5 - 50.5	3
50.5 - 55.5	1
55.5 - 60.5	2
60.5 - 65.5	2
65.5 - 70.5	1
70.5 - 75.5	1
Total	40

The frequency distribution of this type in which upper limit of each case is excluded and lower limit is included is called exclusive form of frequency distribution.

This is a continuous form of distribution as there is continuity in the classes.
The intervals are overlapping.

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

Average of upper limit and lower limit is called **class mark**.

$+ h \{a_n\}^k \quad \varphi \circ U$

B SUBBA REDDY
AECS NO 01 TARAPUR