

Introduction of Statistics

Introduction:

In the modern world of computers and information technology, the importance of statistics is very well recognized by all the disciplines. Statistics has originated as a science of statehood and found applications slowly and steadily in Agriculture, Economics, Commerce, Biology, Medicine, Industry, planning, education and so on. As on date there is no other human walk of life, where statistics cannot be applied.

Meaning of Statistics:

Statistics is concerned with scientific methods for collecting, organizing, summarizing, presenting and analyzing data as well as deriving valid conclusions and making reasonable decisions on the basis of this analysis. Statistics is concerned with the systematic collection of numerical data and its interpretation.

The word 'statistic' is used to refer to

1. Numerical facts, such as the number of people living in Particular area.
2. The study of ways of collecting, analyzing and interpreting the facts.

Definition:

Statistics may be defined as the science of collection, presentation analysis and interpretation of numerical data from the logical analysis.

We have two types for interpretation data

- (i) population
- (ii) Sample.

Population: A population is the set of all possible data value for a subject under consideration.

Sample: A sample is set of data values drawn from much larger population.

Data : Data are the facts and figures collected, analysed, and summarized for presentation and interpretation.

All the data collected in a particular study are referred to as the data set for the study.

Data can be classified as either qualitative or quantitative.

Qualitative data include labels or names used to identify an attribute of each element. Qualitative data use either the nominal or ordinal scale of measurement and may be nonnumeric or numeric.

Quantitative data require numeric values that indicate how much or how many.
Quantitative data are obtained using either the interval or ratio scale of measurement

- **Graphical representation of Data :**

- A picture is said to be more effective than words for describing a particular thing.
- A graphic representation is the geometric image of set of data.
- It is a mathematical picture.
- It enables us to think about a statistical problem in visual terms
- It is an effective and economic device for the presentation, understanding and interpretation of the collected data.

Types of Graphical representation :

(i) Ungrouped Data :

- **Line Graph**
- **Bar Graph**
- **Pie diagram**

(ii) Grouped Data :

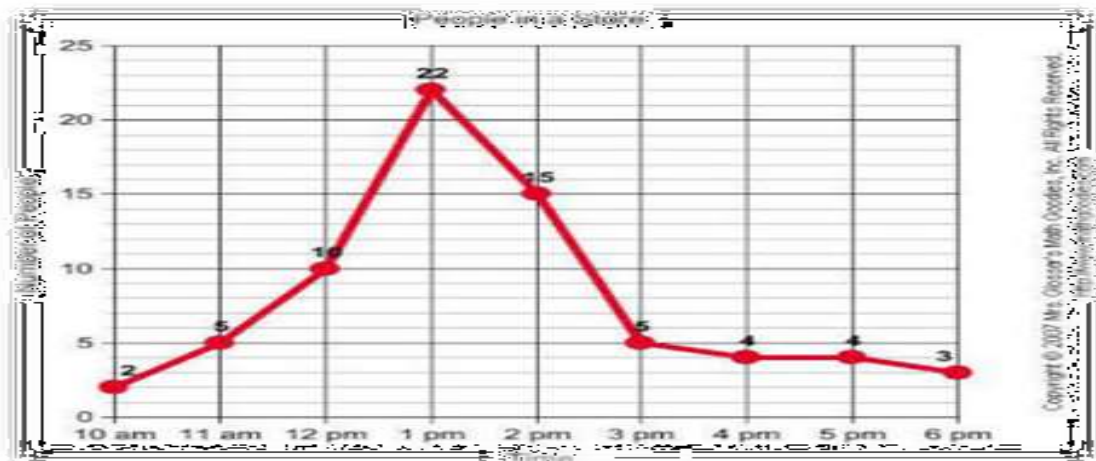
- **Histogram**
- **Frequency polygon**
- **Frequency curve**

Line graph : Line graphs are simple mathematical graph that are draw on the graph paper by plotting the data connecting one variable on the horizontal X – axis and other variable of the data on vertical Y – axis.

Example :

Time	10	11	12	1	2	3	4	5
-------------	-----------	-----------	-----------	----------	----------	----------	----------	----------

	am	am	pm	pm	pm	pm	pm	pm
No of People	2	6	10	22	15	5	4	4



- **Bar Graph** : In bar graph data are represented by bars
- The bars can be made in any direction i.e. Horizontal or Vertical.
- The bars are taken of equal weight and start from a common horizontal or a vertical line and their length indicates the corresponding values of statistical data.
- when do we use bar diagram ?

when the data are given in whole numbers

when data are to be compared easily

Example :

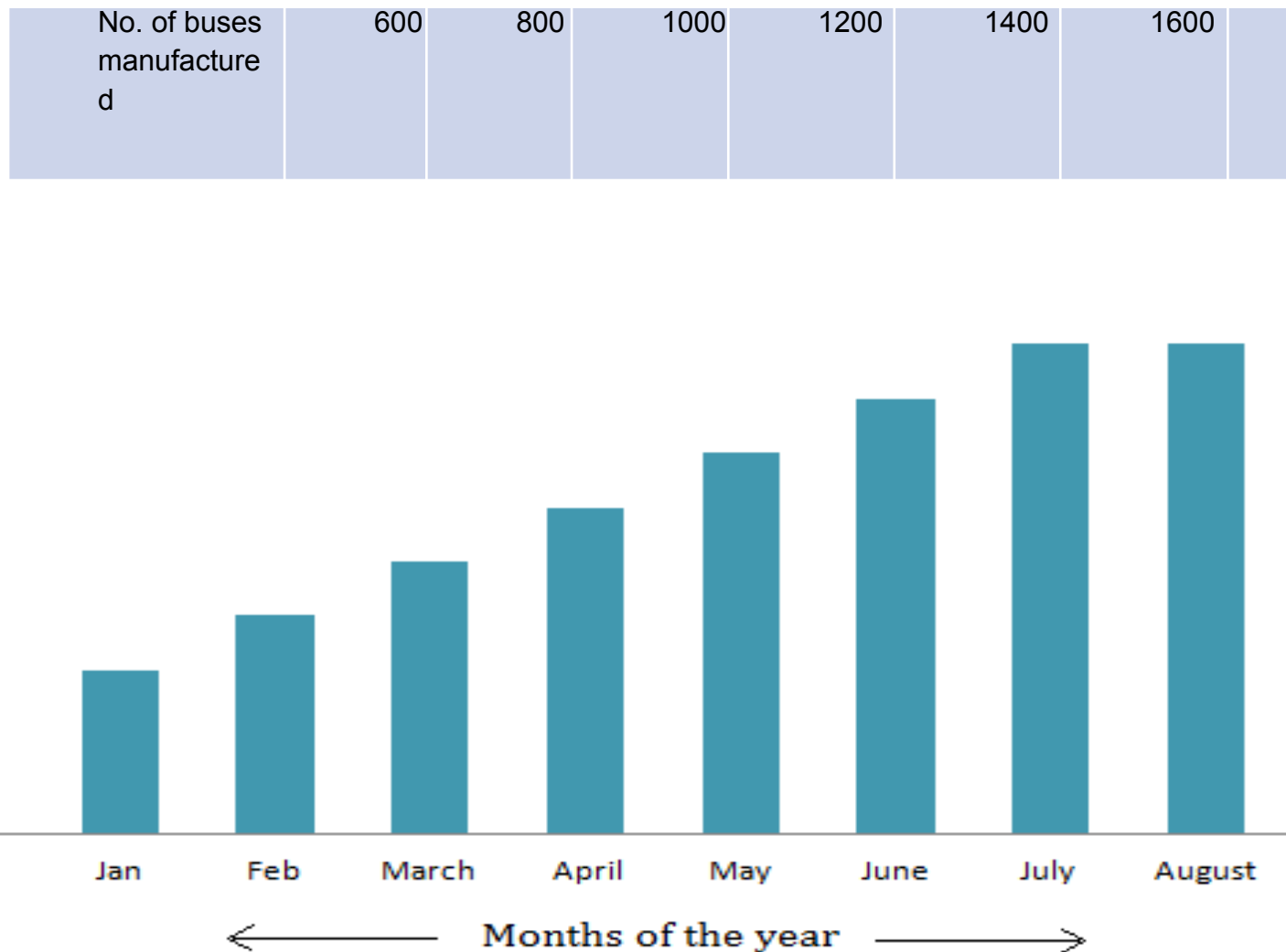
Months	Jan	Feb	Mar	Apr	May	June
--------	-----	-----	-----	-----	-----	------

PARUL UNIVERSITY - FACULTY OF IT & COMPUTER SCIENCE

Department of Applied Science & Humanities

SYLLABUS FOR 3rd SEM BCA/IMCA PROGRAMME

Statistics (05191206)



Pie diagram : It is a circle in which different components are represented through the sections or portions of the circle.

- To draw a pie diagram, first the value of each category is expressed as a percentage of the total and the angle is divided in same percentage.
- Then at the center of circle these angle are drawn simultaneously

Starting from a particular radius.

- In this way we get a set of sectional areas proportional to the values of the item.

When do we use pie diagram ?

When data are given in percentage.

PARUL UNIVERSITY - FACULTY OF IT & COMPUTER SCIENCE

Department of Applied Science & Humanities

SYLLABUS FOR 3rd SEM BCA/IMCA PROGRAMME

Statistics (05191206)

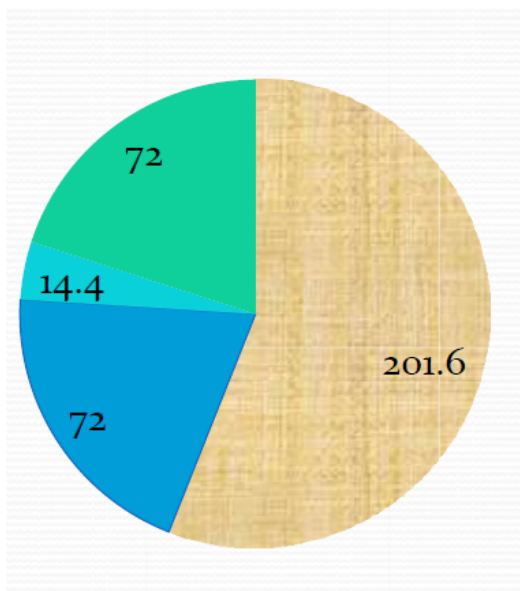
When different aspect of variable are to be displayed.

When the data are to be compared normally.

- Example : Result of class X of a school

• Marks Division	• First	• Second	• Third	• Failures
• % of student	• 20%	• 56%	• 20%	• 4%

Marks division	% of student	Approx angle in degree
First	20%	=
Second	56%	=
Third	20%	=
Failure	4%	=



PARUL UNIVERSITY - FACULTY OF IT & COMPUTER SCIENCE

Department of Applied Science & Humanities

SYLLABUS FOR 3rd SEM BCA/IMCA PROGRAMME

Statistics (05191206)

- **Histogram** : Histogram is essentially a bar graph of frequency distribution .
- It can be constructed for equal as well as unequal class intervals.
- Area of any rectangle of a histogram is proportional to the frequency of that class.

when do we use histogram..?

When data are given in form of frequencies

When class interval has to be displayed by a diagram.

When we need to calculate MODE of distribution graphically.

- Histogram for equal class width:

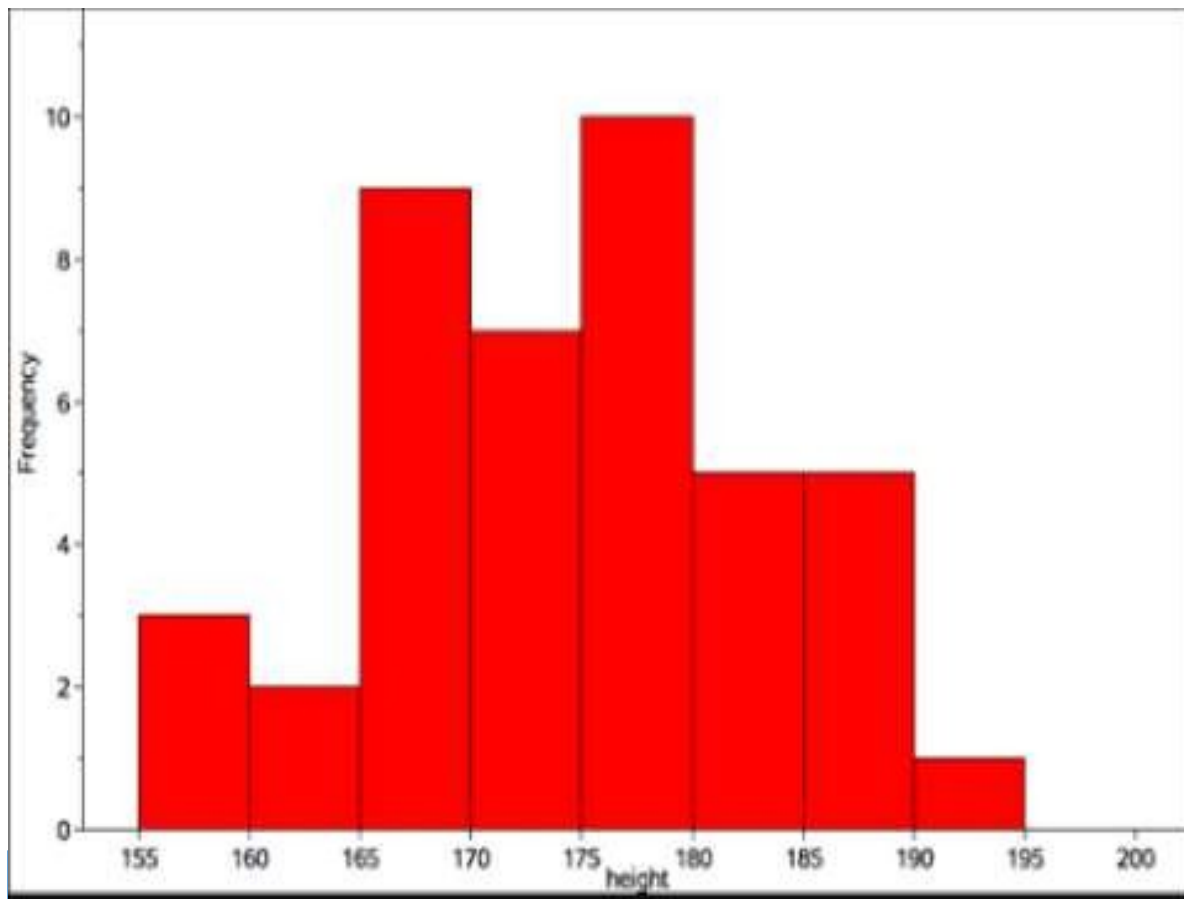
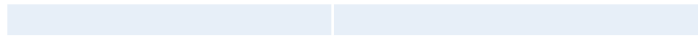
Class Interval (Height in cm)	Freq
155-160	3
160-165	2
165-170	9
170-175	7
175-180	10
180-185	5
185-190	5
190-195	1

PARUL UNIVERSITY - FACULTY OF IT & COMPUTER SCIENCE

Department of Applied Science & Humanities

SYLLABUS FOR 3rd SEM BCA/IMCA PROGRAMME

Statistics (05191206)



boundary		Width
0-10	8	10
10-15	6	5
15-20	12	5

$$8/10 = 0.8$$

$$6/5 = 1.2$$

$$12/5 = 2.4$$

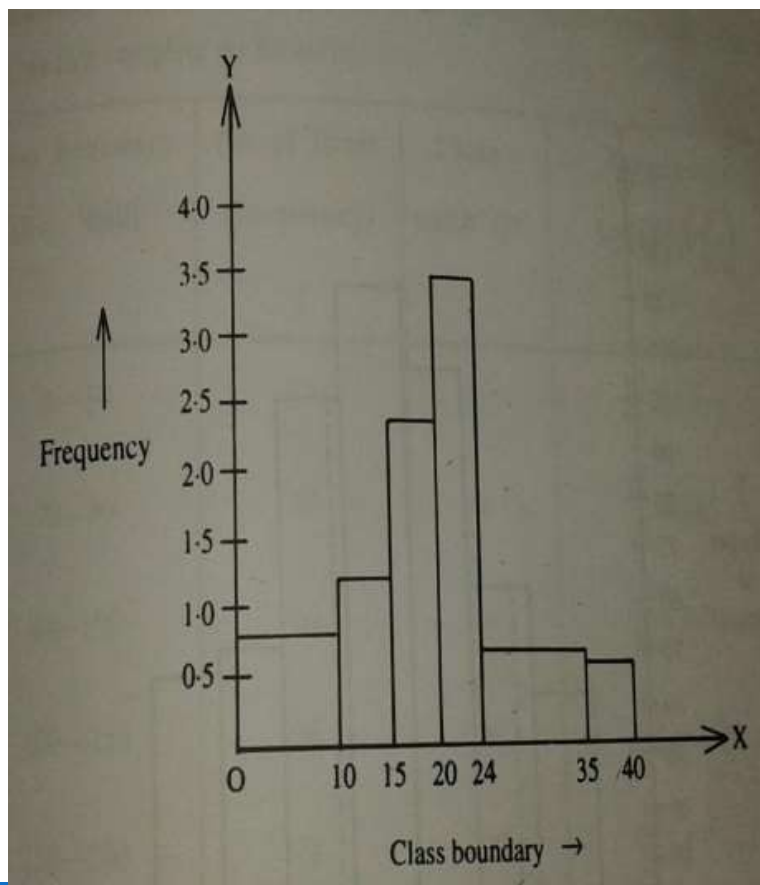
PARUL UNIVERSITY - FACULTY OF IT & COMPUTER SCIENCE

Department of Applied Science & Humanities

SYLLABUS FOR 3rd SEM BCA/IMCA PROGRAMME

Statistics (05191206)

20-24	14	4	$14/4=3.5$
24-35	7	11	$17/11=0.64$
35-40	3	5	$3/5=0.6$



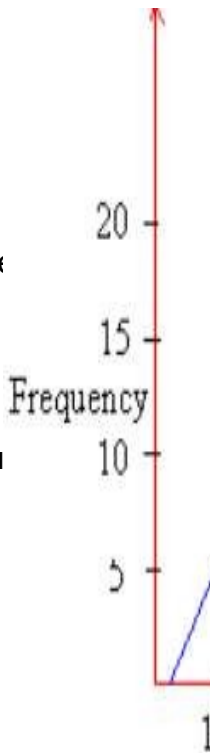
Height in Cm (class interval)	Mid value	frequency
150-154	152	10

154-158	156	15
158-162	160	20
162-166	164	12
166-170	168	8

- **Frequency Curve** : Frequency curve is another type of graphical representation of data.

- When the top points of frequency polygon are joined not by straight lines but by curved ones.

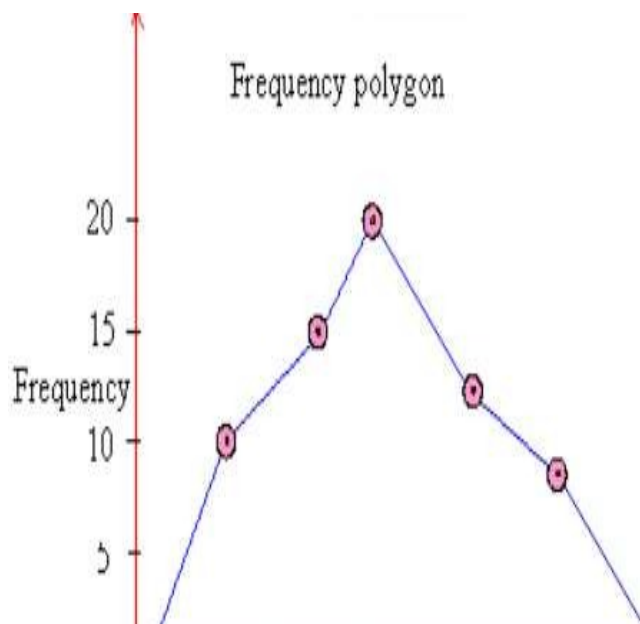
- Frequency polygon is drawn using scale, but frequency curve is drawn using free hand.

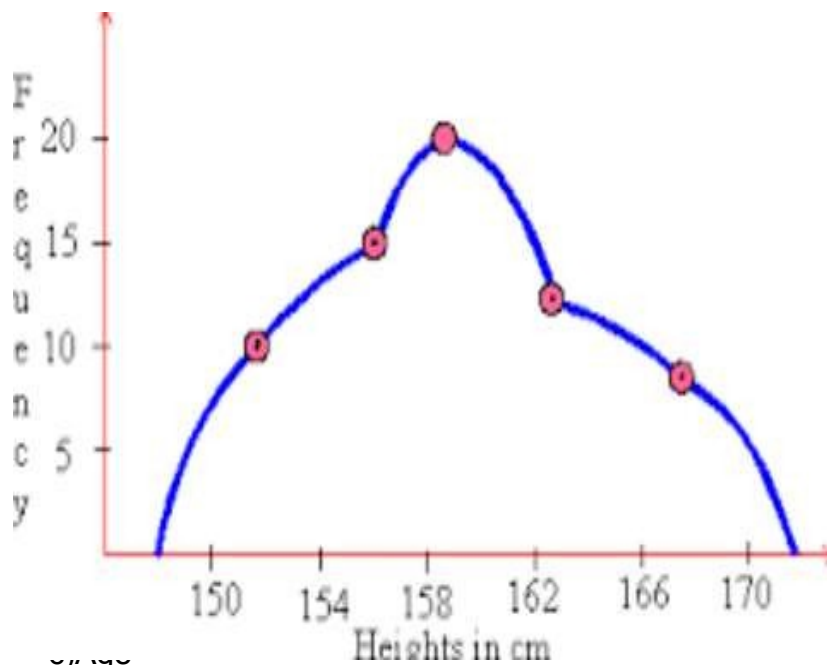


- when do we use frequency curve..?

when the number of class intervals are very large i.e. width of class intervals are very small and the total number of sample values be increased indefinitely

Frequency polygon V/S Frequency Curve:





d) Eye Colour

2) A population value or characteristic that is of interest to us and that we would like to estimate is...

- a) hypothesis
- b) statistic
- c) population

PARUL UNIVERSITY - FACULTY OF IT & COMPUTER SCIENCE

Department of Applied Science & Humanities

SYLLABUS FOR 3rd SEM BCA/IMCA PROGRAMME

Statistics (05191206)

d) parameter ...

3) What is statistical inference? ...

- a) The process of drawing conclusions about a sample based on population data
- b) The process of drawing conclusions about a statistic based on a parameter.
- c) The process of drawing conclusions about a population based on sample data.
- d) None of the above

4) If you have data on the yearly average temperature at Cape Town International Airport from 1900 to 2000, and if you are particularly interested in change over time, what is the most effective graphical display ?

- a) Histogram
- b) Bar graph
- c) Pie diagram
- d) Line graph

5) A histogram is a graphical representation of which of the following:

- a) An ogive
- b) A frequency distribution
- c) A cumulative relative frequency distribution
- d) All of the above

Answer key :

1) d 2) d 3) c 4) d 5) b

Fill in the blanks:

- 1) _____ is an orderly arrangement of data in rows and column.
- 2) _____ is a visual form of presenting tabulated data.

PARUL UNIVERSITY - FACULTY OF IT & COMPUTER SCIENCE

Department of Applied Science & Humanities

SYLLABUS FOR 3rd SEM BCA/IMCA PROGRAMME

Statistics (05191206)

- 3) Vertical rectangles representing class frequencies in a frequency distribution is called _____
- 4) A frequency polygon is a _____ line curve.
- 5) _____ diagrams are used for comparisons

Answer Key:

- 1) Table bar 2) diagram 3) Histogram 4) Straight Line 5) Multi

Questions: Solve the following Question.

- 1) Draw the frequency polygon for the following data

Class interval	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
Frequency	4	6	8	10	12	14	7	5

PARUL UNIVERSITY - FACULTY OF IT & COMPUTER SCIENCE

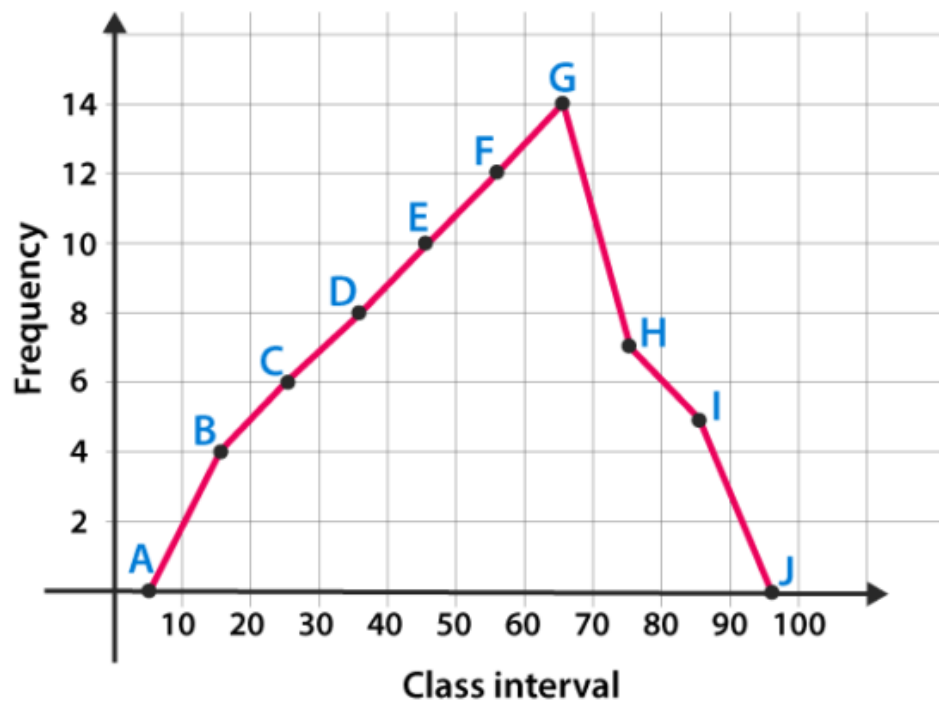
Department of Applied Science & Humanities

SYLLABUS FOR 3rd SEM BCA/IMCA PROGRAMME

Statistics (05191206)

Sol

:



2) The population of Delhi state in different census year is as given below.

Census year	1961	1971	1981	1991	2001
Population (lakh)	30	55	70	110	150

3) refer to the pie chart given below and answer the questions that follow.

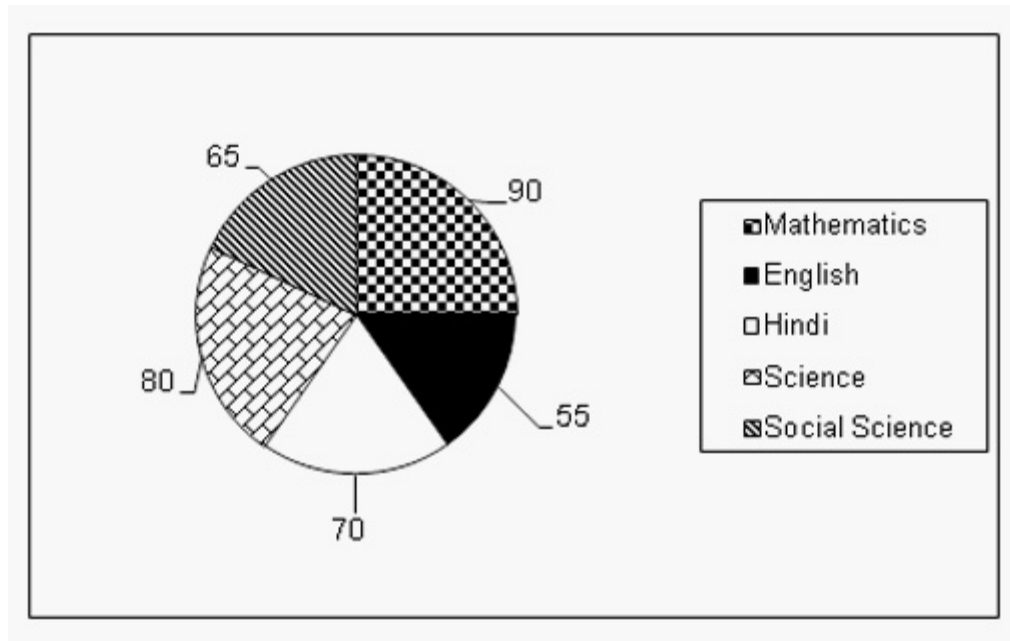
The given pie chart shows the marks scored by a student in different subjects- English, Hindi, Mathematics, Science and Social Science in an examination. The values given are in degrees.

Assumption: Total marks obtained in the examination are 900.

PARUL UNIVERSITY - FACULTY OF IT & COMPUTER SCIENCE

Department of Applied Science & Humanities

SYLLABUS FOR 3rd SEM BCA/IMCA PROGRAMME
Statistics (05191206)



1) If the total marks were 3000, then marks in Mathematics would be

1. 800
2. 750
3. 850
4. 900

2) The Marks scored in English and Mathematics is less than the marks scored in Science and Hindi by

1. 5%
2. 4.33%
3. 3.33%
4. 6%

3) 3. If the marks scored by the student are 137.5, then the subject is

1. English
2. Hindi

PARUL UNIVERSITY - FACULTY OF IT & COMPUTER SCIENCE

Department of Applied Science & Humanities

SYLLABUS FOR 3rd SEM BCA/IMCA PROGRAMME
Statistics (05191206)

3. Mathematics

4. Science

4) Total marks scored in Social Science and English is

1. 400

2. 350

3. 500

4. 300

5) The difference of marks scored in Social Science and Science is

1. 37.5

2. 40

3. 20

4. 15

Ans : 2, 3 , 1, 4, 1

4) In a batch of 400 students, the height of students is given in the following table. Represent it through a frequency polygon

Height (in cm)	No of students
140-150	74
150-160	163
160-170	135
170-180	28
Total	140

Solution :

