MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI

PG & RESEARCH DEPARTMENT OF MATHEMATICS

CLASS : I – B.A. ECONOMICS

SUBJECT CODE : 23UEC12

SUBJECT NAME : STATISTICS FOR ECONOMICS –I

SYLLABUS

UNIT- II

Classification and Presentation of Data

Classification and Tabulation of Data– Types - Frequency Distribution – – Cumulative Frequency Distribution- Class Interval – Diagrams – Types- Graphical Representation– Histogram – Frequency Polygon - Ogive Curve - Lorenz Curve.

Unit - II

Classification and Tabulation :

The collected data in any Statistical investigation are known as raw data. They are huge and confusive. As such they cannot be easily understood by person and are not fit for further analysis and interpretation.

Define classification :

Classification is the process of arranging the available facts into homogeneous group or classes arounding to resemblance and similarities.

"Classification is the process of arranging things (either actually or normally) in group or classes according to their resemblances & affinities given expressions to the unity of attribute that may subsist amongst a diversity of individual."

Characteristic of classification :-

⇒ All the facts are classified into homogeneous groups by the process of classification

⇒ The Basis of classification is unity in diversity.

⇒ classification may be either real or imaginary.

⇒ The classification may be according to either similarities or dissimilarities

⇒ It should be flexible to accommodate adjustments.

Object Types of classification :-

The classification of data primarily depends on the purpose and objectives of the enquiry. There are four importants types of classification.

* Geographical (area wise or region wise or district wise)
* Chronological or historical (on a basis of time)
* Qualitative by character or by attribute
* Quantitative or numerical or magnitude

Geographical classification:

Chronological classification:-

This types of statistical data is classified according to the time of its occurrence such as years, months, weeks, days, hours etc..

For ex: Census data are expressed in decades, national income is expressed every month or week. chronological classification is illustrated below:

Population of india from 1921 to 2001

Years	Population (in million)
1921	248
1931	276
1941	313
1951	357
1961	438

Qualitative classification:-

When the data are classified according to some quality are attributes

such as sex, honesty, intelligence, literacy, colour, religion, marital status, etc..

The classification is termed as qualitative or descriptive attributes. In this type we can only find out the presence or absence of the attribute.

Two types of qualitative classification:

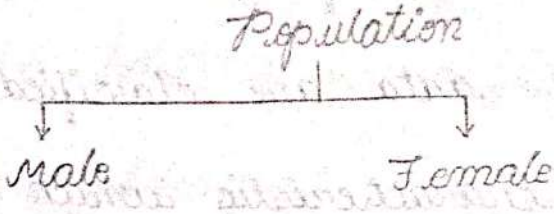
i) Simple classification

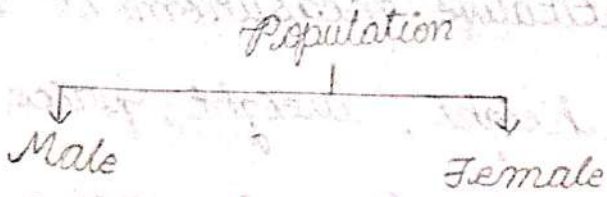
ii) Manifold classification

Simple classification:

If the data are classified into only two classes, such as literate and illiterate or honest and dishonest or skilled and unskilled, the classification is termed as Simple classification. This classification is normally dichotomy or twofold.

Ex :



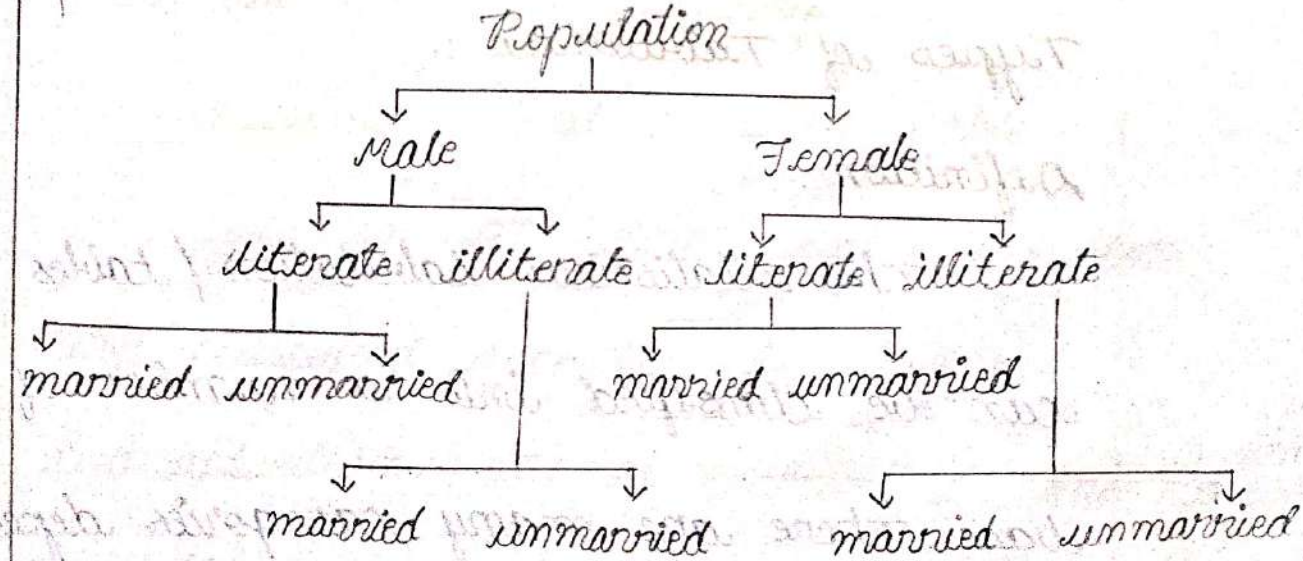


Manifold classification :.

In manifold classification, the universe is classified on the basis of more than one attribute at a time.

Ex :

We may first divided the population into males and females on the attribute of Sex, then further divided them on the basis of literacy and so on.



Quantitative classification :-

If the data are classified according to some characteristic which is capable of quantitative measurement like age, income, height, weight, price, production sales, profits, etc.. is called Quantitative classification or classification according to variables.

Ex :-

Marks	No. of. Students
10 - 20	10
20 - 30	7
30 - 40	13
40 - 50	18
50 - 60	12

Types of Tabulation :-

Definition :-

The Statistical tabulation / tables can be classified into a number of base. There are many categories depends

upon.

1.) The basis of coverage which can be further classified into simple table & complex table. A complex table can be classified into two-fold, three fold or manifold.

2.) The basis of objectives or purpose. This can be further classified into general purpose table or reference table and special purpose table or summary table.

3.) The basis of nature of enquiry, which can further be classified into original or primary table and derived or derivative table.

1.) On the basis of coverage:.

i) Simple and complex table

In a simple table the data are classified according to only one characteristic. It is termed as one way or single table and it takes form of

frequency table.

In a complex table two or more characteristic are shown. It is more popular, because it help appropriate consideration of all related facts.

Ex: Simple table:

Distribution of marks

class marks	No. of. Students
20 - 30	10
30 - 40	18
40 - 50	22
Total	50

ex: complex table:.

i) Two way table:.

class marks	No. of. students		
	Boys	Girls	Total
20 - 30	6	4	10
30 - 40	8	10	18
40 - 50	10	12	22
Total	24	26	50

ii) Three way table :

In this types of table three characteristics are shown. It gives information regarding three interrelated characteristic of a phenomenon.

Ex.:

Distribution of population by age, sex, literacy

Age group (years)	Males			Females			Total		
	literate	illiterate	Total	literate	illiterate	Total	literate	Illiterate	Total
0 - 18	10	20	30	11	12	23	21	32	53
18 - 25	30	10	40	12	13	25	42	23	65
25-35	10	20	30	13	13	26	23	33	56
35-45	20	30	50	12	12	24	32	42	74

iii) Manifold (or) Higher order Table

Ex:-

No. of. Students in M.K University

(according to faculty, age, sex and residence)

Faculty age group (year)	Students						Total			Total
	Boys			Girls						
	Hosteller	Days scholar	Total	Hosteller	Days scholar	Total	Hosteller	Days scholar	Total	
Maths										
20 - 25	10	20	30	33	10	43	43	30	73	73
25 - 30	10	40	50	12	10	22	22	50	72	72
above 30	10	50	60	10	10	20	20	60	80	80

2.) On the basis of Objectives (purpose) :.

i) General purpose table :.

It is also known as informative table. And provide information for general use and usually in chronological order. The detailed use table in the census reports are of this kind. The government agencies prepare this types of tables. Thus are used by research workers and Statisticians. Thus placed in the appendix of a report for reference.

ii) Special purpose table :.

It is also called a Summary table or text table or analytical table or derivative table or derived table. [It presents the data relating to a particular or a Special purpose ratios, percentages, etc...] are used to facilitate comparison.

iii) On the basis of originality :

The statistical table may be classified into primary table and derived table.

In primary table (original), the statistical forms are expressed in origina It contains actual & absolute figures.

In a derived table, figures & ~~table~~ results are derived from the primary data.

It presents totals, percentages, ratios, averages, dispersion, coefficient of correlation.

Both primary and derived tables are generally used in practice.

Cumulative Frequency Distribution :-

A Frequency Distribution shows how the frequency of a particular value (variable or class) in occurring cumulative Frequencies (C.F) are derived

by the cumulation of the frequencies of successive values. C F of a given variable or class represents the total frequency of all previous variables including the variable or the class.

Less than C.F of any value of the variable or classes are obtained by adding successively the frequencies of all previous variable including the variable or class against which it is written. The cumulation is started from the lowest Size to the highest Size.

More than C.F distribution is obtained by finding the cumulation total of frequencies starting from the highest to the lowest variable or class.

Example :.

Ex:

continuous series
class
f

Marks	Frequency	Cumulative Frequency "less than"		Cumulative Frequency "More than"	
20-30	5		5		200
30-40	18	(5+18)	23	(200 - 5)	195
40-50	20	(23+20)	43	(195 -18)	177
50-60	15	(43+15)	58	(177 -20)	157
60-70	18	(58+18)	76	(157 -15)	142
70-80	40	(76+40)	116	(142 -18)	124
80-90	68	(116+68)	184	(124 -40)	84
90-100	16	(184+16)	200	(84 - 68)	16
	Total = 200 $\leq f = 200$				

Class Interval:

Diff. between the lowest limit and the upper limit of the class is known as the class interval. For example, In the class 10 - 20, the class interval is 10. The formula to find the class interval of a given problem is

$$i = \frac{L - S}{K}$$

where, i = class interval

L = Largest item

S = Smallest item

K = Number of classes

Example:

If the marks of 50 students are varied between 10 and 80 and if we want to form 7 classes, then the class interval could be

$$i = \frac{L - S}{K} \qquad i = \frac{80 - 10}{7} \qquad i = 10$$

Therefore the class interval could be 10-20, 20-30, 30-40, 40-50, 50-60, 60-70, 70-80,

Two methods of forming class interval:-

i) Exclusive method (Ex: 10-20, 20-30, 30-40, 40-50)

ii) Inclusive method

Ex:	10 - 19	9.5 - 19.5
	20 - 29	19.5 - 29.5
	30 - 39	29.5 - 39.5
	40 - 49	39.5 - 49.5

Frequency distribution:-

A classification according to the number possessing the small values of the variables.

It is simply a table in which the data are grouped into classes and the number of cases which falls in each class recorded.

Frequency Distribution can be of two kinds i) univariate frequency Distribution

ii) Bivariate Frequency Distribution (two way frequency distribution)

i) Univariate (types)

→ Series of individual observation

→ Discrete frequency distribution

→ Continuous frequency distribution

ii) Bivariate (types)

→ Two way frequency distribution

Example:

a) Make a frequency distribution with intervals of 10 from the following data.

b) Also prepare less than C.F distribution

c) And prepare more than C.F distribution

40	36	43	57	81	90	92	74	66	85
41	57	34	63	84	93	71	55	56	63
39	44	59	43	90	82	88	72	73	45
53	64	79	85	95	68	65	69	83	80

a) Frequency Distribution

Marks	Tally marks	Frequency (f)
30 - 40	III	3
40 - 50	𝍸 I	6
50 - 60	𝍸 I	6
60 - 70	𝍸 II	7
70 - 80	𝍸	5
80 - 90	𝍸 III	8
90 - 100	𝍸	5
		$\Sigma f = 40$

b) Less than C.F Distribution

Marks	Frequency	L.C.F Distribution	
30 - 40	3	(0 + 3)	3
40 - 50	6	(3 + 6)	09
50 - 60	6	(9 + 6)	15
60 - 70	7	(15 + 7)	22
70 - 80	5	(22 + 5)	27
80 - 90	8	(27 + 8)	35
90 - 100	5	(35 + 5)	40
	$\Sigma f = 40$		

c) more than C.F Distribution

Marks	Frequency	M.C.F Distribution	
30-40	3		40
40-50	6	(40-3)	37
50-60	6	(37-6)	31
60-70	7	(31-6)	25
70-80	5	(25-7)	18
80-90	8	(18-5)	13
90-100	5	(13-8)	5
	$\Sigma f = 40$		

Diagram :-

A Diagram is a visual form for presentation of Statistical data Diagram refers to the various types of devices such as Bars, circles, maps, cartograms, pictorials, etc.. Thus devices can take many attractive forms.

Types of Diagram:

There are various diagrametic devices by which statistical data can be presented the common type of diagrams,

→ One dimensional diagram (line & Bar)

→ Two dimensional diagram (Histogra Rectangle, Square, circle etc...)

→ Three dimensional diagram (cube, Sphere, cylinder etc...)

→ Pictogram

→ Cartogram

a) One dimensional diagram:

In one dimensional diagram the length of the lines are bars is considered and the width of the bar is not taken into consideration. The term 'Bar' means a thick, wide line.

This diagram is not attractive; hence it is less important.

One dimensional diagram

a) line diagram

b) Simple bar diagram

c) Multi bar diagram

d) Sub divided bar diagram

e) Percentage bar diagram

f) Other Percentage bar diagram

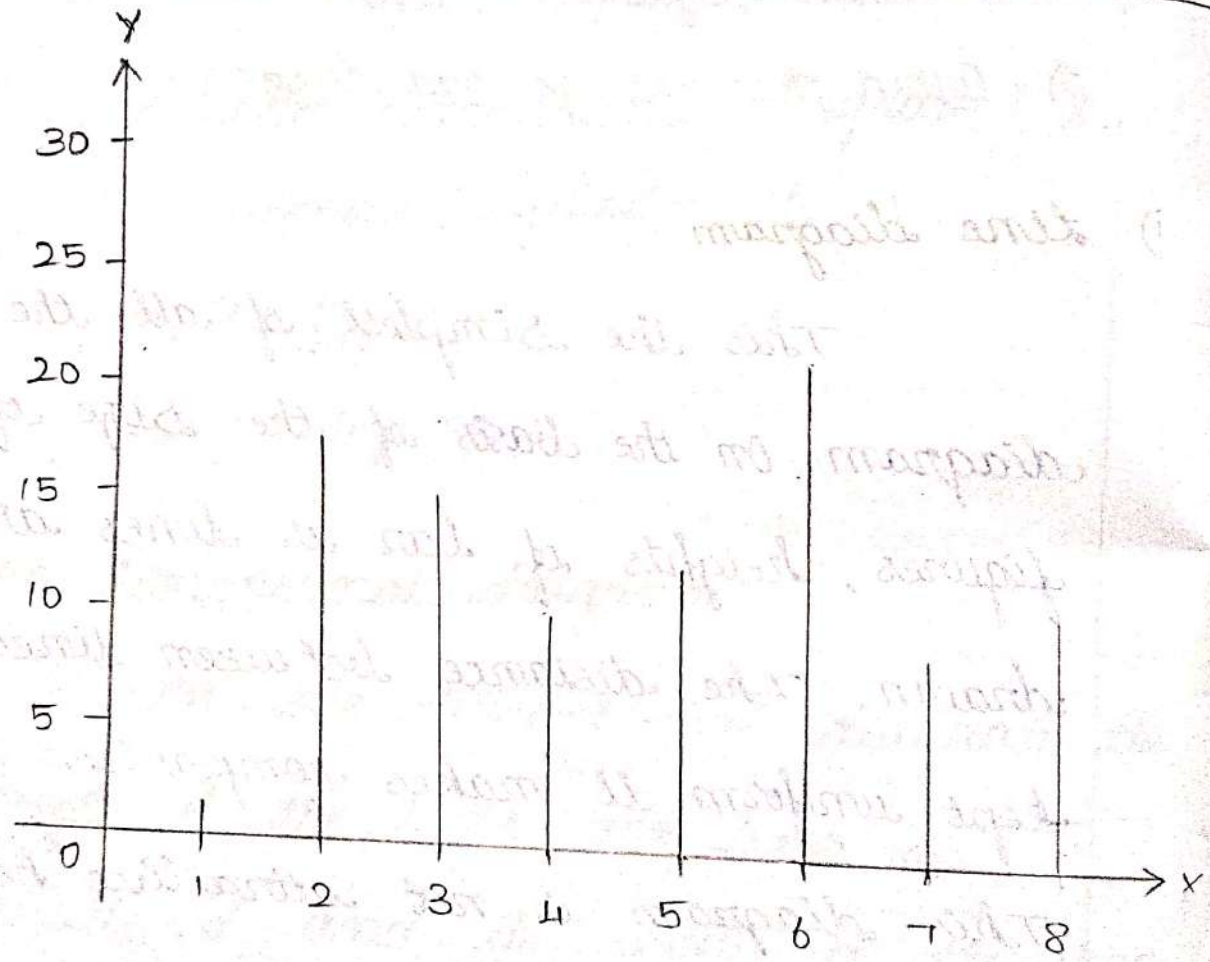
i) Line diagram :.

This the simplest of all the diagram. On the basis of the size of the figures, heights of bar or lines are drawn. The distance between lines is kept uniform it makes comparison easy. This diagram is not attractive hence it is less important.

Ex :

The following data show the number of accidents sustained by 100 drivers of a company in a particular year. Draw a suitable diagram.

No. of accidents	1	2	3	4	5	6	7	8
No. of drivers	2	18	15	10	13	22	9	11



ii) Simple bar diagram :-

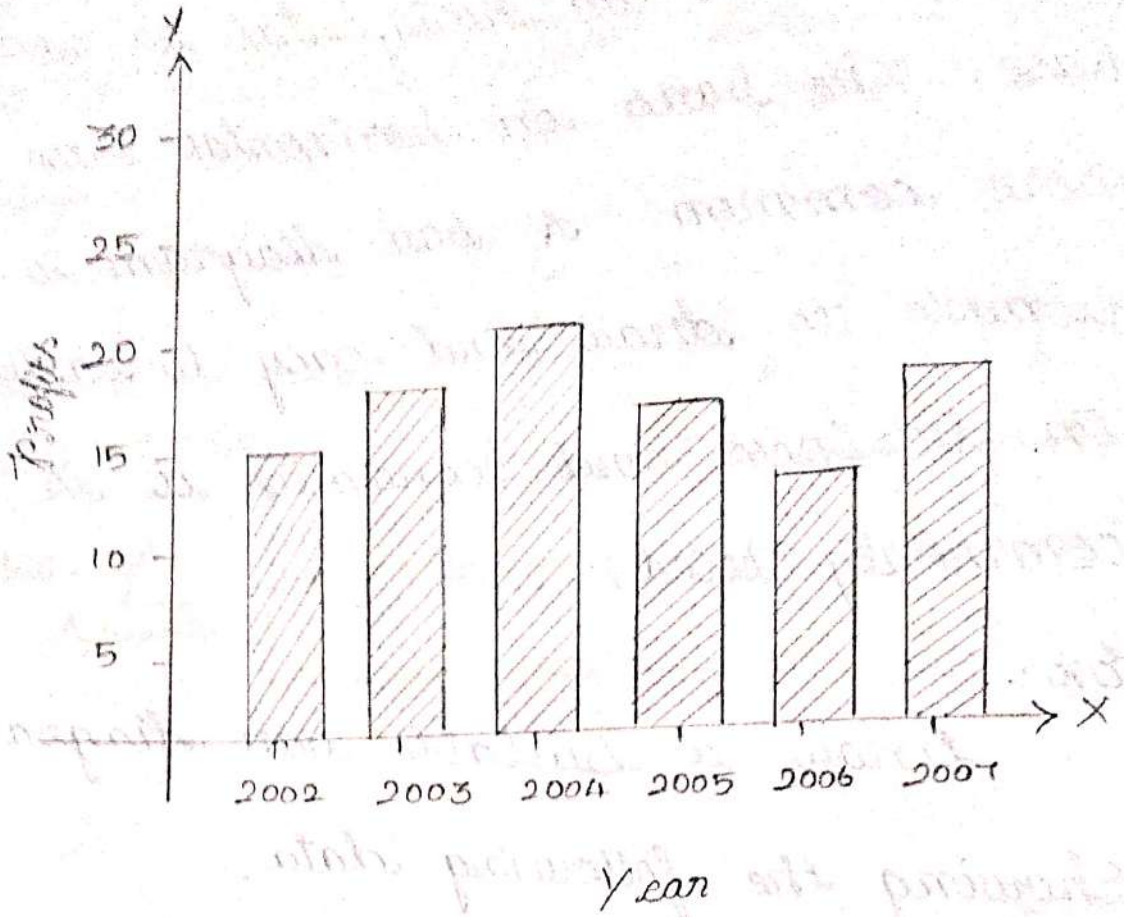
A Simple bar diagram can be drawn either on horizontal or vertical base. The bars on horizontal base or more common. A bar diagram is simple to draw and easy to understand. In business and economics it is commonly used.

Ex :-

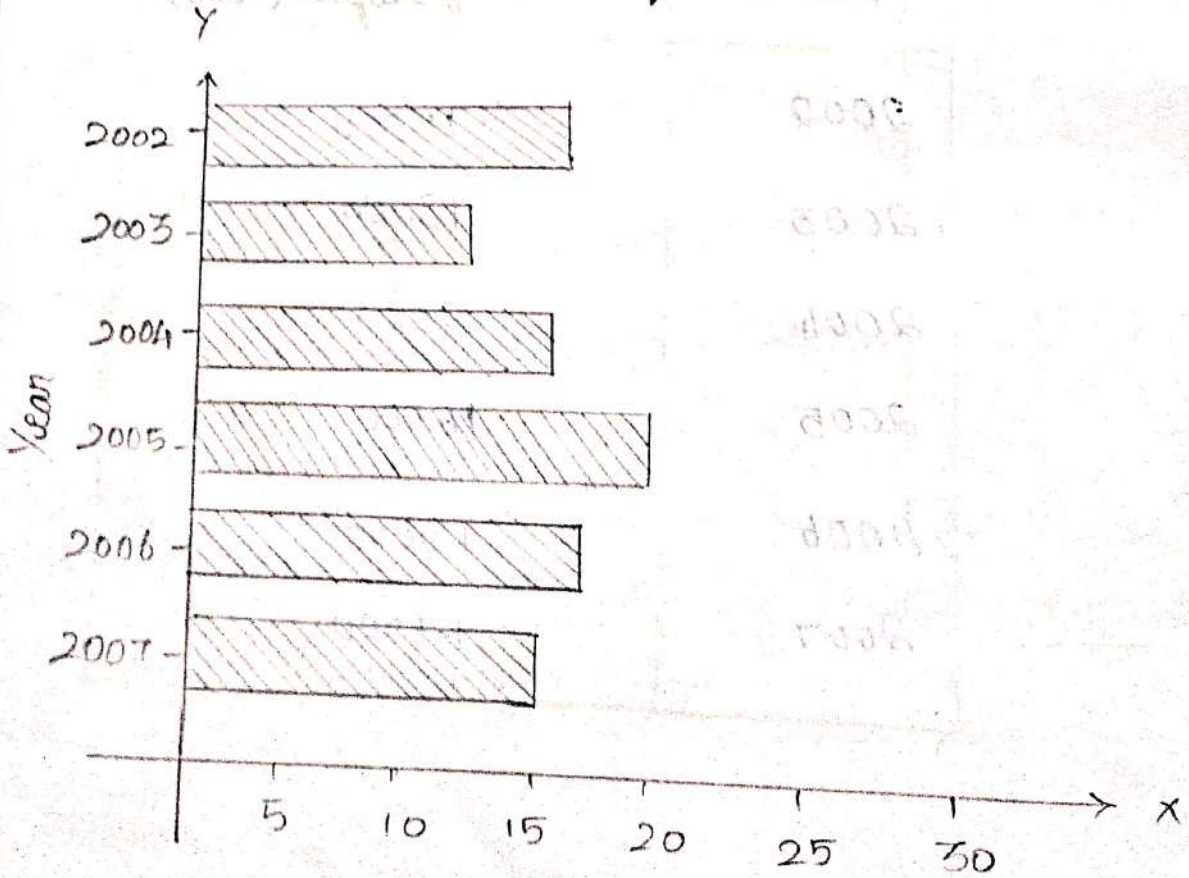
Draw a suitable bar diagram showing the following data.

Year	profits ('000)
2002	15000
2003	18000
2004	20000
2005	16000
2006	13000
2007	17000

a) Vertical bar diagram



b) horizontal bar diagram



iii) Multiple bar diagram (compound bar diagram)

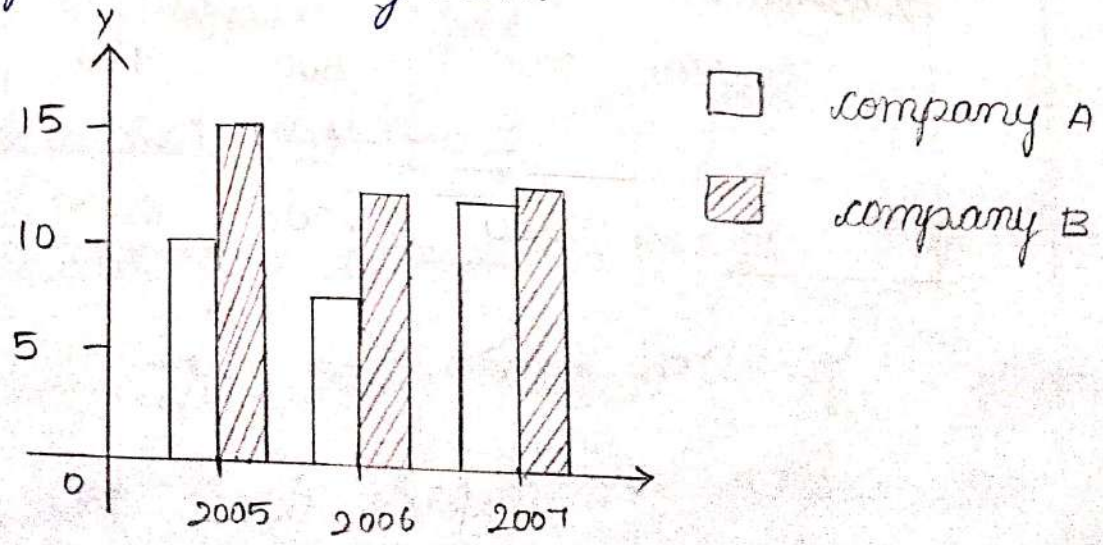
Multiple bar diagrams are used to denote more than on ~~phonemenon~~ Phenomenon. Example, for import and export trend. Multiple bars are useful for direct comparison between two values. The bars are drawn side by side.

Ex:

The data below gives the early profits of two companies A and B

Years	Profits	
	A	B
2005	10,000	15,000
2006	8,000	13,000
2007	13,000	14,000

Represent the data by means of a multiple bar diagram.



iv) Sub divided bar diagram: (Component bar diagram)

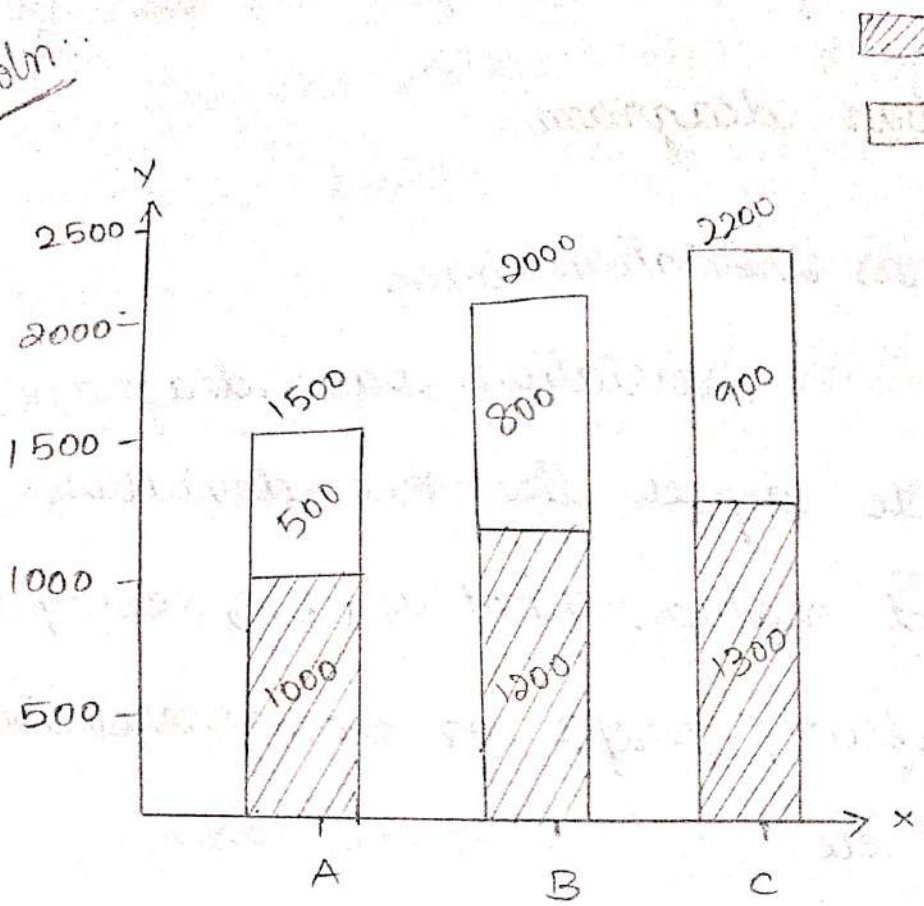
The bar is sub divided into various parts in proportion to the values given in the data & may be drawn on absolute figures or percentages. Each components occupies a part of the bar proportional to its share in the total.

Ex:

Represent the following data in a suitable diagram.

Districts	A	B	C
Male population	1000	1200	1300
Female	500	800	900
	1500	2000	2200

Soln:



v) Percentage Sub divided bar diagram:-

The above mentioned diagram have been used to represent absolute value. But comparing this method on a relative basis. The varies components are expressed as percentage to the total. For dividing the bar thus percentages are cumulated. In this case. The bars are all of equal height. Each segment shows the

percentage to the total.

vi) Other bar diagram :-

a) Deviation bars

Deviation bar diagram is used to depict the net deviations in different values. That is (i.e) Surplus or deficit, profit or loss, Net import or export etc..

b) Broken bars

In certain cases, we may come across data which contain very wide variation in values - very small or very large. In order to provide adequate & reasonable shape to the smaller bars, the larger bar may be broken at the top. The value of each bar is written at the top of the bar.

b) Two dimensional diagram :

In one dimensional diagram, only length is taken into account. In two dimensional diagram, the area of the diagram represents the data, i.e, the length and breadth are considered.

The important types are,

a) Rectangles →

The Rectangles are used when two or more magnitudes with different components have to be compared. The area of the rectangles are kept in proportion to the values. It may of two types,

★ Percentage Sub divided rectangular diagram

In such a diagram, the width of the rectangle is kept according to the proportion of the values. The

Various components of the values are converted into percentage & rectangle divided according to them.

★ Sub divided rectangle

Such diagrams are used to show some related phenomenon. Ex: cost per unit, quantity of production etc..

Example:

Draw a two dimensional diagram to represent the following data.

Item of expenditure	Expenditure in rupees	
	Family A	Family B
1. Food	200	300
2. Clothing	48	75
3. Education	32	40
4. House Rent	110	75
5. Miscellaneous	80	110
	400	600

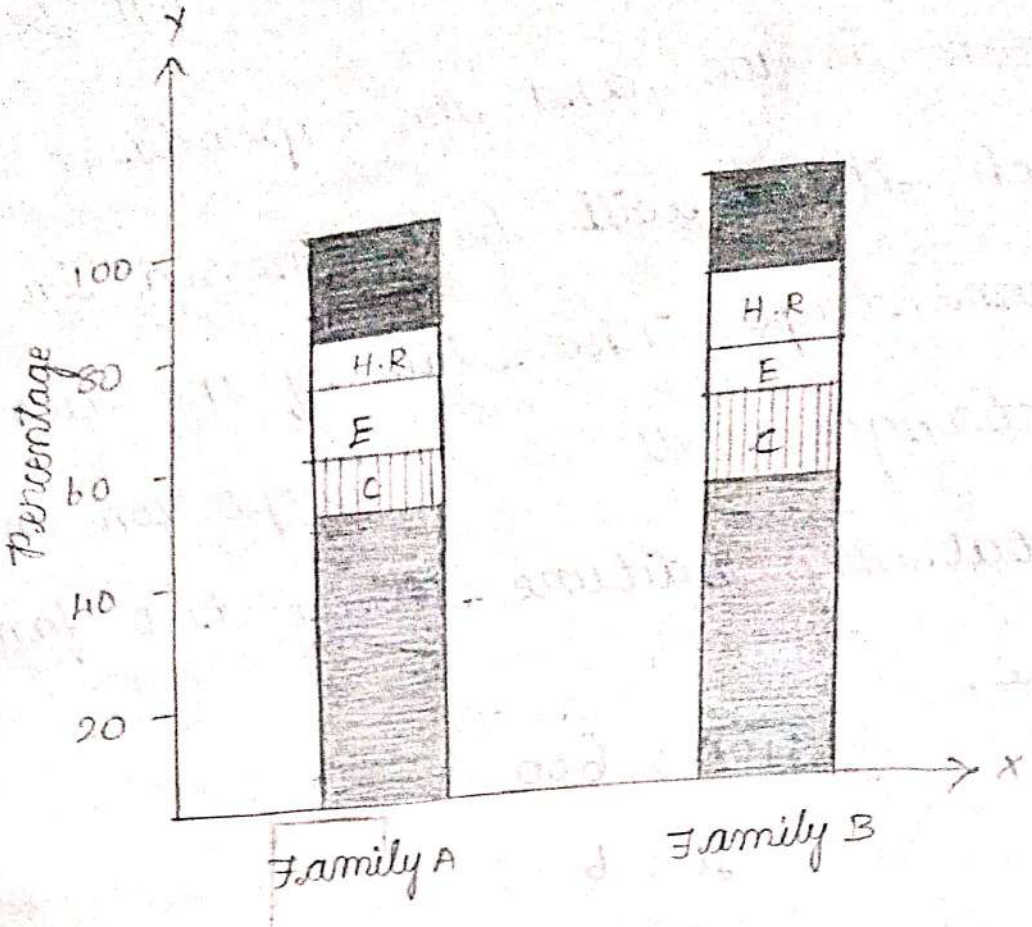
Soln:-
The total expenditure will be taken as 100 and the expenditure on each item will be expressed in percentage. The width of the two rectangle will be in proportion to the total expenditure of the two families, i.e,

400 : 600

4 : 6

2 : 3

Items of expenditure	Monthly Expenditure					
	Family A (Rs. 400)			Family B (Rs. 600		
	Rs	%	cumulative %	Rs	%	cumulative %
Food	200	50	50	300	50	50
Clothing	48	12	62	75	12.5	62.5
Education	32	8	70	40	6.67	69.17
House Rent	40	10	80	75	12.5	81.67
Miscellaneous	80	20	100	110	18.33	100
	400	100		600	100	



b) Square →

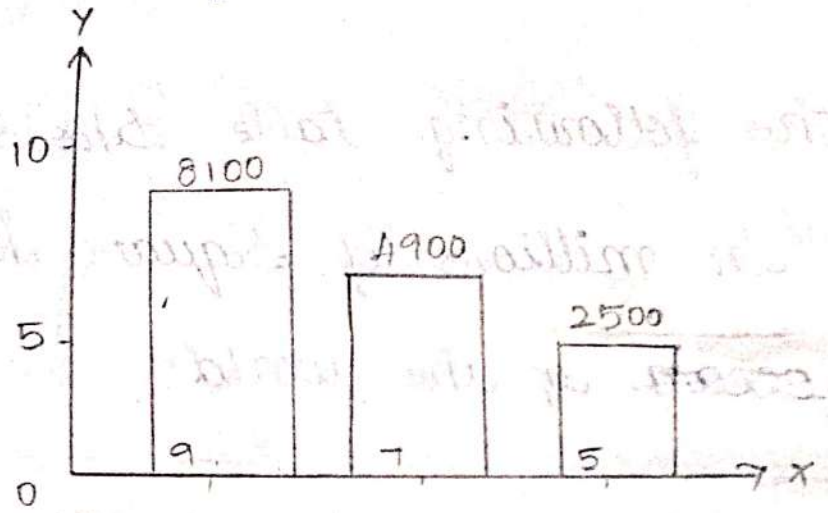
While preparing Squares we have to bear in mind that the ratio it to be maintained according to the area of Squares. To draw a Square diagram, the Square root is taken of the values of the varies items to be shown in the diagram

Ex :

Draw a Square diagram to represent the following data.

8100 4900 2500

Soln :-
The square roots of the figure they are 90, 70, 50. Further the roots are divided by 10. Thus we get 9, 7, 5.



c) circle →

circle diagram are alternative to square diagram. Steps are similar to the above. The side of the square will become the radius of the circle.

d) pie chart (or) angular →

The pie chart diagram runs high in understanding. Just as divide a bar or a rectangle to show its component, a circle can also be

divided into sectors. As there are 360° at the centre, proportionate sector or cut taking the whole data equal to 360°.

Ex:

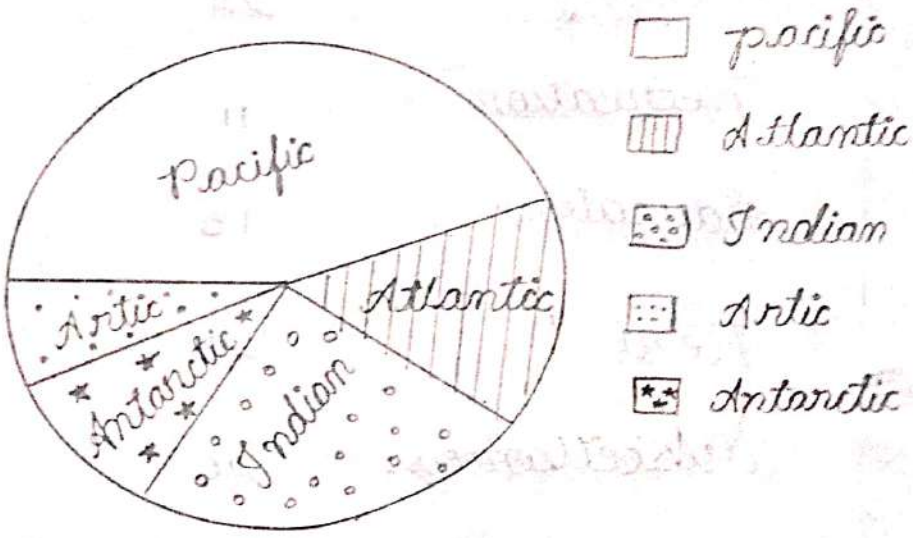
1. The following table shows the area in millions of square km of the ocean of the world.

Ocean	Area (million Sq. km)	
Pacific	70.8	Degrees 167
Atlantic	41.2	97
Indian	28.5	67
Antartic	7.6	18
Artic	4.8	11
	152.9	

Draw a pie diagram to represent the data.

Soln:

Ocean	Area (million sq. km)	Degrees
Pacific	70.8	167
Atlantic	41.2	97
Indian	28.5	67
Antarctic	7.6	18
Artic	4.8	11
	Area total = 152.9	360



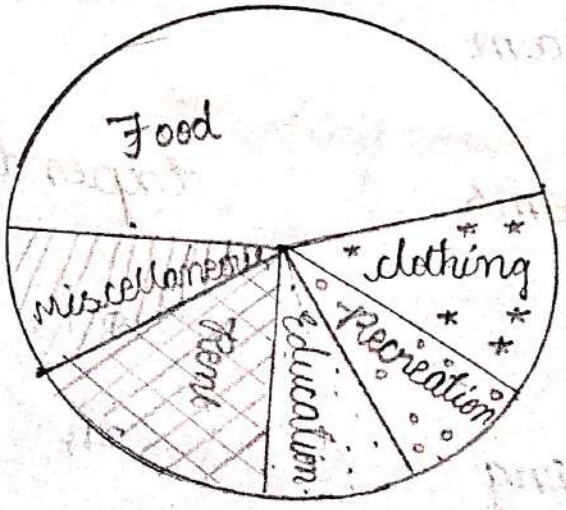
2. Represent the following by a pie chart diagram

Items	Expenditure
Food	87
clothing	24

Recreation	11
Education	13
Rent	25
Miscellaneous	20

Soln:-

Items	Expenditure	Degrees
Food	87	174
clothing	24	48
Recreation	11	22
Education	13	26
Rent	25	50
Miscellaneous	20	40
	Total = 180	360



c) Three dimensional diagram :-

The Square, the rectangle, circle etc.. may failed to represent to the data if the quantities to be represented or awfully diverse. In such case three diagrams are drawn. They are called length, height, and width or depth are considered and thus comprise of cubes, Spheres, cylinders etc...

d) Pictogram:

Pictogram is a device of representing Statistical data in pictures. Thus are very useful in attracting the attention. They are easily understood. For the purpose of propaganda, the pictorial presentation of facts are quite popular and find place in exhibition. They are extensively used by government

organisation as well as by private institution.

e) cartogram

In cartogram, statistical facts are presented through maps accompanied by various types of diagrametic representa It presents the numerical facts in a pictorial form in a geomeptrical or special distribution. cartograms are simple and easy to understand.

Graphical Representation (or) presentation

Graphic presentation of numerical data is becoming popular because of various merits a graph is a visual form of presentation. Graphs are drawn on a special types of paper known as Graph paper. Each Graph praper has thick horizontal & vertical line of each division of a

centimeter and thin lines of smaller part of the same. A Graph is divided into four quadrant but normally the 1st quadrant is used.

Graph of frequency Distribution:

Graphical Representation can be advantagious employeed to bring out clearly the statistical nature of frequency distribution which may be discrete or continuous.

The most commonly used graphs are,

i) histogram
ii) Frequency polygon
iii) Frequency curve (or) Lorenz curve
iv) Ogives curve (or) cumulative frequency curve.

Histogram :-

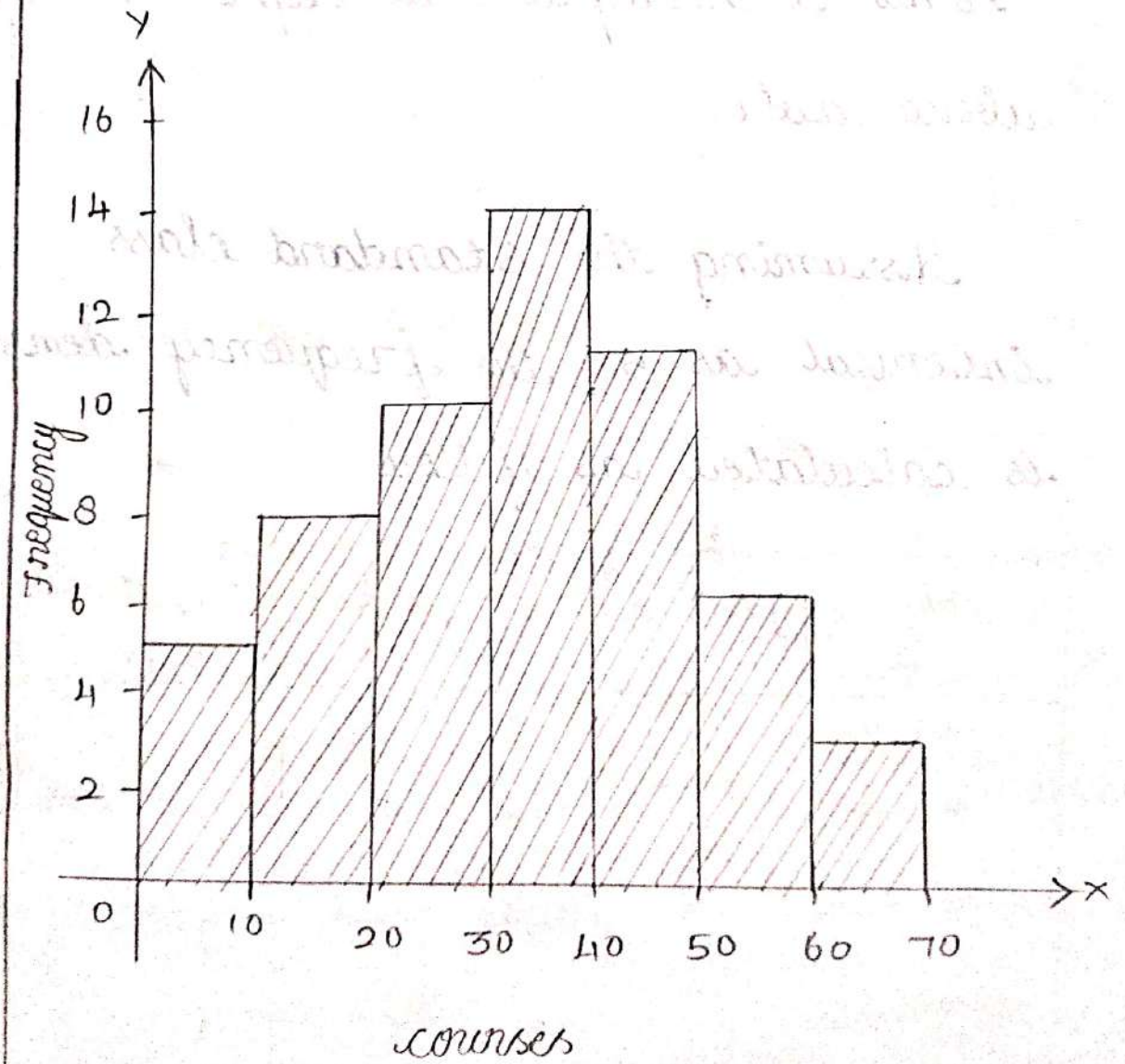
One of the most important and useful methods of presenting FD of continuous series is known as histogram In this, the magnitude of the class interval is plotted along the horizontal axis and the frequency on the vertical axis. Each class has lower and upper values. This gives us two equal vertical lines Representing the frequency. Upper ends of the lines are joined together. This process will gives us rectangles, as they are classes, and the heights of Rectangles are proportion to the frequency. Histogram is also known as block diagram or Stair case chart.

Ex:

From the following data draw a histogram.

Course	Frequency
0 - 10	5
10 - 20	8
20 - 30	10
30 - 40	14
40 - 50	11
50 - 60	6
60 - 70	3

Soln:



Ex:

Weekly wages	No. of workers
10 - 15	7
15 - 20	19
20 - 25	27
25 - 30	15
30 - 40	12
40 - 50	12
60 - 80	8

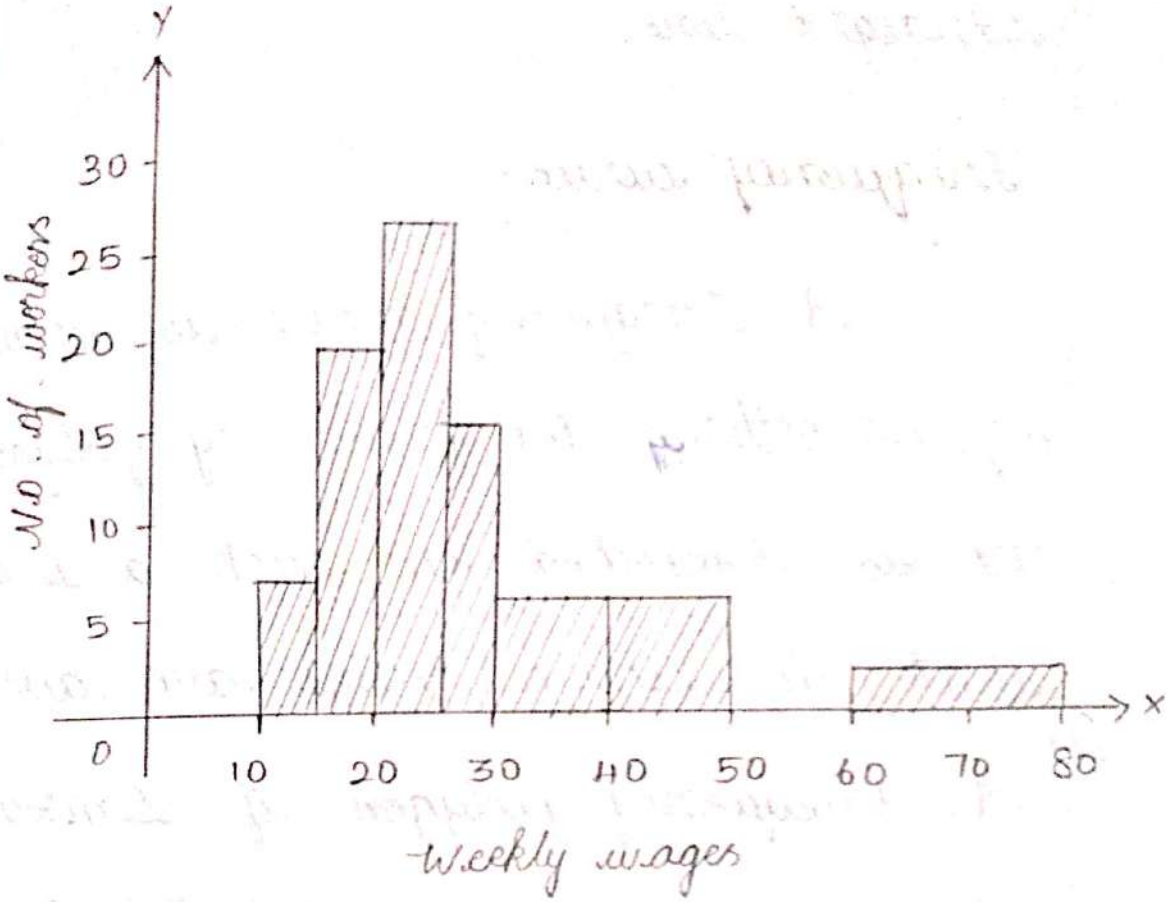
Draw a histogram to represent the above data.

Soln:

Assuming the standard class interval as 5, the frequency density is calculated as follow

Weekly wages	No. of workers	Adjustment (5=1)	Freq density
10 - 15	7	1	$7/1 = 7$
15 - 20	19	1	$19/1 = 19$
20 - 25	27	1	$27/1 = 27$
25 - 30	15	1	$15/1 = 15$

30 - 40	12	2	$12/2 = 6$
40 - 50	12	2	$12/2 = 6$
60 - 80	8	4	$8/4 = 2$



Frequency polygon:.

A grouped Frequency Distribution can be represented by a histogram. A simple method of smoothing the histogram is to draw a frequency polygon. i.e done by connecting the

mid point of the top of each rectangle with the mid point of the top of each adjacent rectangle by straight line.

Frequency curve :-

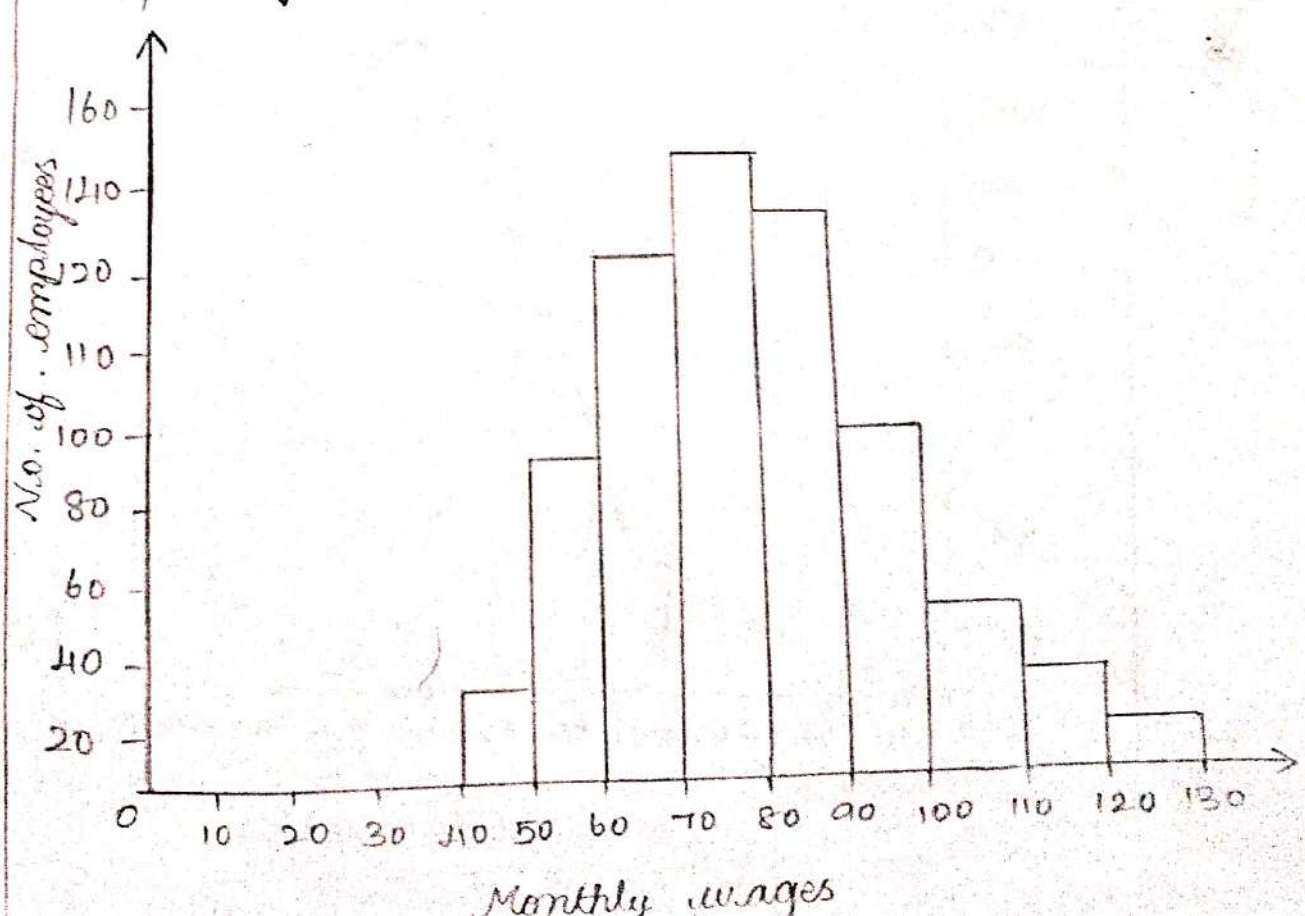
A Frequency curve is drawn by smoothing the frequency polygon. It is smoothed in such a way that the sharp turns are avoided. A Frequency polygon, if smoothed further so as to minimize sudden changes, results into a continuous smooth curve. known as frequency or smooth frequency curve. The curve should begin and end at the base line.

Ex:

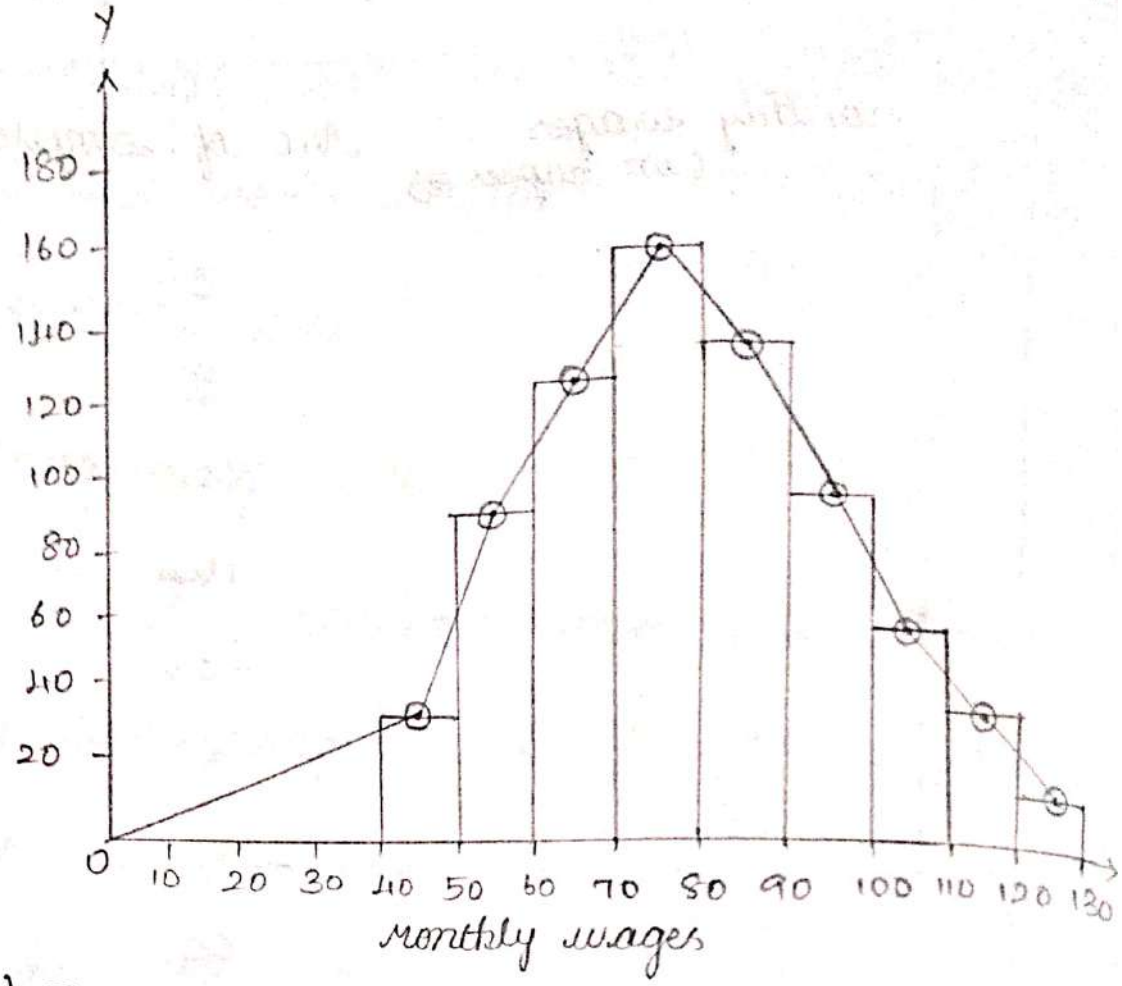
Monthly wages (in rupees ₹)	No. of employees
40 - 50	36
50 - 60	87
60 - 70	121
70 - 80	154
80 - 90	133
90 - 100	95
100 - 110	50
110 - 120	30
120 - 130	10

Soln:

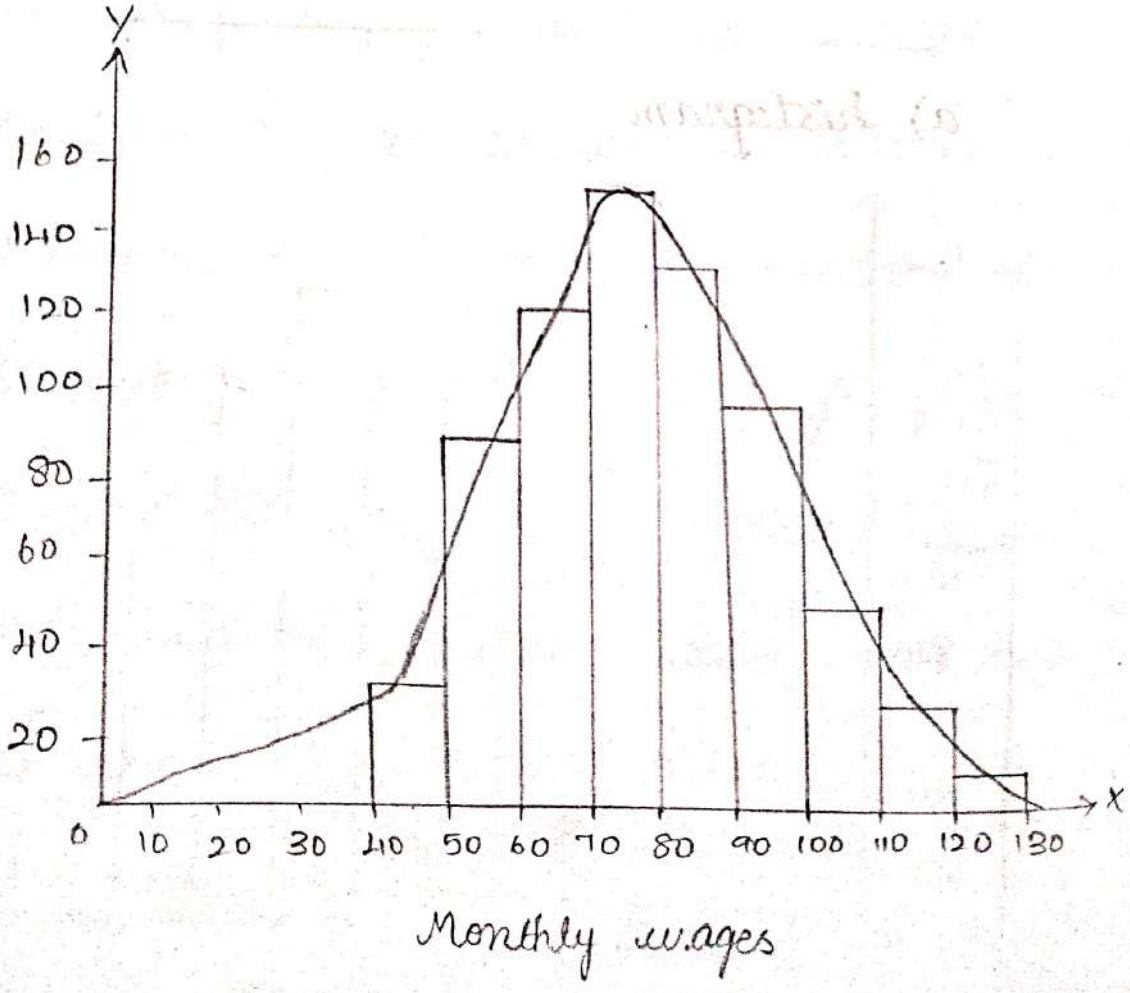
a) Histogram:



b) Frequency polygon: only straight lines ∧



c) Frequency curve: only curve lines ∩



Ogive curve :-

Ogive when cumulative frequencies are plotted on a graph, then the frequency curve obtained is called ogive or cumulative frequency curve. There are two methods of constructing ogive that is,

i) Less than ogive

ii) More than ogive

Ex :-

Draw less than and more than cumulative frequency curve from the following.

Marks	Frequency
0 - 10	3
10 - 20	9
20 - 30	15
30 - 40	30
40 - 50	18
50 - 60	5

Soln:

Marks	Frequency	cumulative Frequency 'less than'		cumulative Frequency 'more than'	
0 - 10	3		3		80
10 - 20	9	(3+9)	12	(80 - 3)	77
20 - 30	15	(12+15)	27	(77 - 9)	68
30 - 40	30	(27+30)	57	(68 - 15)	53
40 - 50	18	(57+18)	75	(53 - 30)	23
50 - 60	5	(75+5)	80	(23 - 18)	5
	$\Sigma f = 80$				