1 Statistical Description of Data

INTRODUCTION

The word statistics has been derived either of the following
Latin word "Status" which means a political State.
Italian word "Statista"
German word "Statistik"
French word "Statistique

1. DEFINITION OF STATISTICS

The word statistics is used in two different senses-**Plural and Singular.**

Plural form: it refers to the numerical data collected in a systematic manner with some definite aim or object in view.
Example: The number of persons suffering from malaria in different colonies of Delhi or number of unemployed girls in different states of India and so on.

2. Characteristics

I. Statistics are **aggregate of facts.**
II. Statistics ate **affected by a large number of causes**
III. Statistics are always **numerically expressed**
Iv. Statistics should be **enumerated or estimated**
v. Statistics should be **collected in systematic manner**
VI. Statistics should be collected for **pre-determined purpose**
VII. Statistics should be placed in **relation to each other.**

Singular form: the word statistics means the science of statistics that deals with the principles devices or statistical methods of collecting, analyzing and interpreting numerical data.

3. IMPORTANCE AND SCOPE OF STATISTICS

I. Statistics and Economics
(a) Consumption: Statistical data of consumption enables us to find out the ways in which people in different strata of society spend their incomes.
(b) Production: The statistics of production describe the total productivity in the country. This enables us to compare ourselves with other countries of the world.
(c) Exchange: In the field of exchange. An economist studies markets laws of prices are determined by the forces of demand and supply cost of production monopoly competition banking ate. A systematic study of an all these can be made only with the help of statistics.
(d) Econometrics: With the help of econometrics. Economics has become exact science. Econometrics is the combination of economics mathematics and statistics.
(e) Public Finance: public finance studies the revenue and expenditure activities of a country revenue and expenditure activities of a country Budget.(a statistical document) fiscal policy, deficit financing etc are the concepts of economics which are based on statistics.
(f) Input-Output Analysis: The input-output analysis is based on statistical data which explain the relationship between the input and the output. Sampling, Time series, Index numbers, Probability, Correlation and Regression are some other concepts which are used in economic analysis.

II. Statistics and commerce
Statistics or statistical methods help the business establishments in analyzing the business activities such as:
(a) Organization of Business: Businessman makes extensive use of statistical data to arrive at the conclusion which guides him in establishing a new firm of business house.
(b) Production: the production department of an organization prepares the forecast regarding the production of the commodities with the help of statistical tools.
(c) Scientific Management and Business Forecasting: Better and efficient control of a business can be achieved by scientific management with the help of statistical data. The success of businessman lies on the accuracy of forecast made. **(d) Purchase:** The price statistics of different markets help businessman in arriving at the correct decisions. Raw material is purchased from those markets only where the prices are low.

III. Statistics and Business
Statistics is an indispensable tool in all aspects of business. When a man enters business he enters the profession of forecasting because success in business is always the result of precision in forecasting and failure in business is very often due to wrong expectations. Which arise in turn due to faulty reasoning and inaccurate analysis of various causes affecting a particular phenomenon. Boddingtons observes. "The successful businessman is the one whose estimate most closely approaches the accuracy.

4. LIMITATIONS OF STATISTICS

Statistics and its techniques are widely used in every branch of knowledge**. W.I. King rightly says:** ***"Science of statistics is the most useful servant. But only of great value to those who understand its proper use."***

Important limitations
1. Statistics does not deal with **individual item.**
2. Statistics deals with **quantitative data**.
3. Statistics laws are true only on **averages.**
4. Statistics does not reveal the **entire story.**
5. Statistics is liable to be **misused**.
6. Statistics data should be **uniform and homogeneous.**

5. Types of statistical data:

1. Primary data
2. Secondary data

1. Primary data: It's the data collected by a particular person or organization for his own use from the primary sources.
2. Secondary data: It is the data collected by some other person or organization for their own use but the investigator also gets it for his use.

Methods of collecting primary data	**Methods of collecting secondary Data**
1. Direct personal observation: In this method. The investigator collects the data personally and therefore, It gives reliable and correct information. **USE:** In case of natural calamity data can be collected more quickly and accurately by applying this method. **2. Indirect oral investigation** In this method. A third person is contacted who is expected to know the necessary details about the persons for whom the enquiry is meant. **USE:** if there are some practical problems in reaching the respondents directly, as in case of rail accident. **3. Estimates from the local sources and correspondence**: Here the investigator appoints agents and correspondents to collect the data. **4. Data through Questionnaire:** The data can be collected by preparing a **Question**naire and getting it filled by the persons concerned. **5. Investigations through enumerators**. This method generally employed by the Government for population census etc.	1. Information collected through **newspapers and periodicals**. 2. Information obtained from the **publications of trade associations**. 3. Information obtained from the **research papers** published by university departments or research bureaus or UG.C. 4. Information obtained from the **official publications** of the Central, State and the local governments dealing with crop statistics, Industrial statistics, Trade and transport statistics etc. 5. Information obtained from the **official publications of the foreign governments** for international organizations

.

CLASSIFICATION OF DATA

1. INTRODUCTION

Secrist *" Classification is the process of arranging data into sequences and groups according at their common characteristics or separating them into different but related parts".*

Requisites of a Good Classification

The main characteristics of a good classification are:

1. It should be exhaustive.
2. It should be unambiguous.
3. It should be mutually exclusive.
4. It should be stable.
5. It should be flexible.
6. It should be suitability.
7. It should be homogeneous.
8. It should be a revealing classification.

ADVANTAGES OF CLASSIFICATION OF DATA

i. It condenses the data and ignores unnecessary details.
ii. It facilitates comparison of data.
iii. It helps in studying the relationships between several characteristics.
iv. It facilitates further statistical treatments.

2. TYPES OF CLASSIFICATION OF DATA

i. Quantitative or Cardinal Classification
ii. Chronological or Temporal or Time series Classifying :
iii. Geographical or Spatial Classification
iv. Qualitative or Ordinal Classification

I. Quantitative Classification. When the basis of classification is according to differences in quantity the Classification is called quantitative classification.
Example: Height. Weight

II. Temporal Classification: When the basis of classification is according to differences in time the classification is called temporal or (Chronological) classification.

III. Spatial or Geographical classification: When the basis of classification is according to geographical location or place. The classification is called spatial or geographical classification.

Iv. Qualitative Classification: When the basis of classification is according to characteristics or attributes like social status etc. is called qualitative classification
Example: Married or single, Honest or dishonest, Beautiful or ugly, On the basis of religion viz. Hindu, Muslim, Sikh, Christian, etc.
Classification of this nature is of two types:
i. **Simple classification or Two-Fold Classification or One Way Classification**.
ii. **Manifold classification**.

1. Simple Classification or Two-fold Classification: If the data are classified only into two categories according to the presence or absence of only one attribute the classification is known as simple or two-fold classification or (Dichotomous.) For example the population of India may be divided into males and females: literate and illiterate etc.

2. Manifold Classification: it is the classifications in which more than one attributes are involved.

MODE OF PRESENTATION OF DATA

(a) Textual presentation:
(b) Tabular presentation or Tabulation:
(c) Diagrammatic presentation:

(A) TEXTUAL PRESENTATION

This method comprises presenting data with the help of a paragraph or a number of paragraphs.
Example: The official report of an enquiry commission, Government Reports etc.

Example 1: In 1995 out of total of 2,000 students in a college. 1,400 were for graduation and the rest for post-graduation (P.G.). Out of 1,400 Graduate
Students 100 were girls however in all there were 600 girls in the college. In 2000 number of graduate student increased to 1,700 out of which 250 were girls but the number of P.G. students fall to 500 of which only 50 were boys In 2005 out of 800 girls 650 were for graduation whereas the total number of graduates was 2,200 the number of boys and girls in P.G. classes was equal.

Example 2: Out of total number of 2,807 women, who were interviewed for employment in a textile factory, 912 were from textile areas and the rest from non-textile areas. Amongst the married women who belonged to textile areas 347 were having some work experience while for non-textile areas the Corresponding figures were 199 and 670 respectively. The total number of women having no experience was 1,841 of whom 311 resided in textile areas. Of the total number of women 1,418 were unmarried and of these the number of women having experience in the texile and non-textile areas was 254 and 166 respectively

Merits

- The merit of this mode of presentation lies in its **simplicity** and even a layman can present and understand the data by this method.
- The observations with **exact magnitude** can be presented with the help of textual presentation.
- This type of presentation can be taken as the **first step** towards the other methods of presentation.

Demerits:

- Dull
- Monotonous
- Comparison between different observations is not possible in this method
- For manifold classification this method cannot be recommended.

(b) TABULAR PRESENTATION OR TABULATION OF DATA

Tabulation may be defined as logical and systematic arrangement of statistical data in rows and Columns. It is designed to simplify the presentation of data for the purposes of analysis and statistical inferences.

Objectives of Tabulation

The purpose of tabulation is to **summarize lots of information** in such a simple manner that it can be easily analyzed and interpreted.

The **main objectives** of the Tabulation are.
1. To **simplify** the complex data.
2. To **clarify** the objective of investigation.
3. Economize space.
4. To facilitate **comparison**.
5. To depict **trend** a pattern of data
6. To help reference for **future studies**.
7. To facilitate statistical **analysis**.
8. To **detect errors** and omissions in the data.
9. To clarify the **characteristics** of data.

Essential parts of a statistical Table

1. Table Number: A table should be numbered for identification especially when there are a large number of tables in a study. The number may be put at the center, above the title or at the bottom of the table.

2. Title of the table: Every table should have a title It should be clear brief and self explanatory The title should be set in bold type so as to give it prominence.

3. Date: The date of preparation of table should always be written on the table. It enables to recollect the chronological order of the table prepared.

4. Stubs or Row designations: Each row of the table must have a heading. The designations of the rows are called stubs or stub items. Stubs clarify the figures in the rows. As far as possible the items should be condensed so that they can be included in a single row.

5. Captions or Column headings: A table has many columns, sub-headings of the columns are called captions or headings. They should be well-defined and brief.

6. Body of the Table: It is the most vital part of the table It contains the numerical information. It should be made as comprehensive as possible.

7. Unit of Measurements: The unit of measurement should always be stated along with the title if this is uniform throughout. If different units have been adopted, then they should be stated along the stubs or captions.

8. Source Notes: A note at the bottom of the table should always be given to indicate the primary source as well as the secondary source from where the data has been taken particularly. When there is more than one source.

9. Foot Notes and References: It is always placed at the bottom of the table. It is a statement which contains explanation of some specific items which cannot be understood by the reader from the title or captions and stubs.

Format of a Table

Table No. XXX

Title: [Head note or prefatory Note]

Stub Heading	Caption					Total (6)
	Sub-Heads		Sub-Heads			
	Column Head (1)	Column Head (2)	Column Head (3)	Column Head (4)	Column Head (5)	
Row 1		**Body of Table**				
Row2						
Row 3						
Total						

Foot Note

Source Note:

Number and Date: No. Dc 1172-25/2/2006

Q 1: Try to convert **Example 1** and **Example 2** given above to Textual presentation

Q 2: In a trip organized by a college there were 100 persons the average cost works out to Rs. 13.60 per head. There 80 students each of whom paid Rs. 16 Members of the teaching staff were charged at

a higher rate. The number of servants was 6 (all males) and they were not charged. The number of ladies was 20% of the total of which two were ladies staff members.
Tabulate the above information in proper tabular form.

DIFFETENCE BETWEEN TEXUAL AND TABULAR PRESENTATION

The tabulation method is usually preferred to textual presentation as:
(a) It facilitates comparison between rows and columns.
(b) Complicated data can be represented using tabulation.
(c) Without tabulation statistical analysis of data is not possible.
(d) It is a must for diagrammatic representation.

(C) DIAGRAMMATIC PRESENTTION OF DATA

- The representation of statistical data **through Charts, Diagrams and Picture** is another attractive and alternative method.
- Unlike the first two methods of representation of data diagrammatic representation can be **used for both the educated section and uneducated section** of the society.
- Furthermore, any **hidden trend** presented in the given data can be noticed only in this mode of representation.
- However, Compared to tabulation this is **less accurate**. So if there is a priority for accuracy we have to recommend tabulation.

Types of diagrams
I. Line diagram
II. Bar diagram
III. Pie chart

Rules For construction of Diagrams

1. Title: Each diagram should be given a suitable title either at the top or at the bottom. The title should be brief, self explanatory, clear and un-ambiguous.
2. Foot Notes: Foot notes are given at the bottom of each diagram in order to clarify some points relating to the diagram
3. Source Note: Wherever possible should be appended at the bottom of the diagram. It is given to indicate the source from where the data have been taken.
4. Neat and Attractive: Diagrams should be made very neat, Clean and attractive be proper size and lettering.
5. Size of diagram: The size of diagram would depend on the magnitude of data which is to be shown. The size should be such that all the important characteristics of the data are properly emphasized and can be understood by a look at the diagram.
6. Selection of scale: scale of presentation should be consistent with the size of the paper and the size of the observations to be displayed so that the diagram obtained is neither too small nor too big. The lettering of the scale must be very clear and legible.
7. Index: A brief index explaining various types of shades, Colours, lines and designs used in the diagram should be given for clear understanding of the diagram.
8. Simplicity: The diagrams should be as simple as possible so that they are easily understood even by a lay-man who does not have any mathematical background.
9. Proportion between length and breadth: the length and breadth to the diagram should be consistent with the space available.
10. Choice of a diagram: The choice of a diagram for the given data (or data under consideration) be made with utmost caution and care.

(I) LINE DIAGRAM OR HISTORIGRAM

- It is the simplest of all the diagrams.
- With the help of such graphs the effect to one variable upon another variable during an experimental or normative study may be clearly demonstrated.
- This diagram facilitates comparison, even a time series data can be presented by a line diagram.

- When we observe the values of a variable at different period of time the series so formed is known as **Time series.**
- This is the simplest and easiest graph and no technical skill is needed in its construction.
- Moreover, it enables an individual to present more information in a more precise form than any more other kind of chart can do.
- Two or more variable can be shown on the same graph and thus comparison becomes easy.
- Graphs of Time series can be constructed either on a natural scale or on a ratio scale.

In **natural scale, Absolute changes** from one period to another are shown whereas in a **ratio scale the relative changes** are shown.

Q 3: A word-nonsense Syllables Association test was administered on students of a class X to demonstrate the effect of practice on learning. The data so obtained may be studied from the following table.

Trial No: 1 2 3 4 5 6 7 8 9 10 11 12
Score: 4 5 8 8 10 13 12 12 14 16 16 16

Draw a line graph for the representation and interpretation of the above data.

(II) BAR DIAGRAM

- It is especially useful in comparing qualitative data or quantitative data of discrete type.
- A bar diagram is a graph on which the data are represented in the form of bars. It consists of a number of bars or rectangles which are of uniform width with equal space between them on the X-axis.
- The length of the bar is proportional to the value it represents. It should be seen that the bars are neither too short nor too long.
- The scale should be clearly indicated and base line be clearly shown.
- Bars may be drawn either horizontally or vertically. A good rule to use in determining the direction is that if the legend describing the bar can be written under the bars when drawn vertically Vertical bars should be used, when it cannot be Horizontal ones must be used.
- The descriptive legend should not be written at the ends of the bars or within the bars since such writing may distort the comparison.
- Usually the diagram will be more attractive if the bars are wider than the spaces between them
- The width of bars is not governed by any set rules.

Types of Bar diagram

(I) Simple Bar Diagram:
(II) Multiple or Grouped Bar Diagram:
(III) Subdivided or component Bar Diagram.
(Iv) Percentage subdivided Bar Diagram:
(V) Deviation or Bilateral Bar Diagram:
(VI) Broken Bars.

(I) Simple Bar Diagram

- It is used to compare two or more items related to a variable.
- In this case. The data are presented with the help of bars.
- These bars are usually arranged according to relative magnitude of bars.
- The length of bat is determined by the value or the amount of the variable.
- A limitation of simple Bar Diagram is that only one variable can be represented on it.

(II) Multiple or Grouped Bar Diagram

- A multiple or grouped bar diagram is used when a number of items are to be compared in respect of two, three or more values.
- In this case the numerical values of major categories ate arranged in ascending or descending order so that the categories can be readily distinguished.

(II) Sub-divided or Component Bar Diagram

- A Component bar diagram is one which is formed by dividing a single bar into several component parts.

- A single bar represents the aggregate value whereas the component parts represent the component values of the aggregate value.
- It shows the relationship among the different parts and also between the different parts and the main bar.

(V) Deviation Bar

- Deviation bars are popularly used for graphic presentation of net deviation in different values which have both positive and negative value.
- The positive values are shown above the baseline and negative values are shown below the baseline.

(VI) Broken Bars

- Sometimes the data contains very wide variation in values i.e. very small or very large.
- In order to provide adequate and reasonable shape to the smaller bars. Larger bars may be broke at top which is known as **Broken** bars.

PIE DIAGRAM OR ANGULAR DIAGRAM

- A pie diagram is a circular graph which represents the total value with its components.
- The area of a circle represents the total value and the different sectors of the circle represent the different parts.
- The circle is divided into sectors by radii and the areas of the sectors are proportional to the angles at the center.
- It is generally used for comparing the relation between various components of a value and between components and the total value.
- In pie diagram the data are expressed as percentage of the total value. A pie diagram is also known as angular diagram.

Q: 8 120 students of a college were asked to opt for different work experiences, the details of these Housing	12
Fuel and lighting	5
Miscellaneous	8

PRESENTATION OF DTATA BY FREQUENCY DISTRIBUTION

- The methods to condense the data (Collected by an investigator) in the tabular form to study their salient features are called the presentation of data.
- The raw data can be arranged in anyone of the following ways:

I serial order or Alphabetical order
II. Ascending order or Descending order
III. Tables or charts
IV . Groups or Class intervals.

Example: The marks obtained by 30 students in a class test out of 50 marks. According to their roll numbers are:
41, 25,5,33,12,21,19,39,19,21,1,19,12,
19,17,12,17,17,41,19,41,33,12,21,33,5
1,21

Array: An array is an arrangement of data in order of magnitude either in ascending order or in descending order.
The arrayed data do not tell us much except the maximum(s) or minimum (s) of data.

TALLY BARS AND FREQUENCY

In order to make the data easily understandable we tabulate the data in the form of tables or charts. A table has three columns.
(i) Variable:
(ii) Tally marks:

(iii) Frequency.

(i) Variable: Any character which can vary from one individual to another is called a variable or a variate.
Example: age, income, height, intelligence, Colour etc.

Variables or observation with numbers as possible values are called **quantitative variables**.
Example: age, Height, temperature, etc.

Variable with names of places, quality, things etc. as possible values are called **qualitative Variables or attributes**.
Example: Colour, Intelligence etc.

Variables are of two **types**
(I) Continuous (II) Discontinuous or Discrete

Quantities which can take all numerical values **within a certain interva**l are called **continuous variables**.
A discrete variable can assume only **integral values** and is capable exact measurement. In other words, those variables which can take only **a finite set of values** are called discrete variables.
Example: number of students in a particular class. Number of sections in a school etc.

(ii) Tally: It is a method of keeping count in blocks of five.
Tally Bars: These are the straight bars used in the Tally.

(iii) Frequency: The number of times an observation occurs in the given data is called the frequency of the observation.

Frequency Distribution: A frequency distribution is the arrangement of the given data in the form of a table showing frequency with which each variable occurs.

TYPES OF FREQUENCY DISTRIBUTIONS
Frequency distribution is of two types:
(I) Discrete frequency Distribution
(II) Grouped (or Continuous) Frequency Distribution

(I) Discrete Frequency Distribution
The construction of discrete frequency distribution from the given raw data is done by the method of tally marks as explained earlier.
Construction of Discrete Frequency Distribution Table
The frequency distribution table has three columns headed by
1. Variables (or Classes) 2. Tally marks or Bars 3. Frequency

USE: This method is convenient only when the values in the raw data are largely repeating and the difference between the smallest and largest observation is not very large.

(II) Continuous or Grouped Frequency Distribution

Continuous series: When the continuous variables
are arranged in the form series, it is called continuous series or exclusive series.

Discrete or Discontinuous series: when the discrete variables are arranged in form of a series. It is called a discrete or discontinuous series.
Range: it is the difference between the largest and the smallest number
R= L- S, where **L is Largest** value and **S is smallest** value of Data.

Class, Class Interval and Class limits:

Class: If the observations of a series are divided into groups and the groups are bounded by limits, then each group is **called class**.

Class Limits: The end values of the class are called **class limits**.
The smaller value of the class limit is called the **lower limit** and the higher value of class the same is called the **upper limit** of the class. These two class limits are sometimes called the **stated class limits**.

Class Interval: The **difference** between the lower limit (L) and the upper limit (U) of the class is known as class interval (I)
I = U - L In other words. The **range of a class** called its class interval

Actual class Limit or Class Boundaries: If there a gap between the limits of any two adjacent class then those type of continuous series are known as
Inclusive Series. This gap may be filled up by extending the two limits of each class by half of the value of the gap thus.

Lower class boundary = Lower class limit – ½ of the gap
Upper class boundary= upper class limit + ½ of the gap.

Actual class limits or True class limits: The class boundaries are the limits up to which the two limits (inclusive series) of each class may be extended to fill up the gap that exists between the classes, the class boundaries so obtained are known as Actual class limits or True class limits.

Mid Point, Mid-value or Class Mark: It is the arithmetic mean of the lower class and upper class limit of the same class.
Mid-value of class $= \frac{L+U}{2}$

Class mark $= \frac{\text{True upper class limit} + \text{True lower class limit}}{2}$

Class Magnitude: It is the difference between the upper class boundary and the lower class boundary of the class.

Inclusive and Exclusive series

Inclusive series: When the class-intervals are so fixed that the upper limit of the class is included in that class it is known as inclusive method of classification.
Example: 0 – 5 6 – 10 11 – 15 16 - 20

In the inclusive series the upper limit and lower limit are included in that class interval.

Exclusive or continuous series: In this series the upper limit of one class is the lower limit of the other class. The common point of the two classes is included in the higher class. When the class-intervals are so fixed that the upper limit of one class is the lower limit of the next class, it is known as **Exclusive method of classification**.

CUMULAIIVE FREQUENCY DISTRIBUTION

Cumulative frequency corresponding to a class is the **sum of all the frequencies** up to and including that class.

Cumulative frequency series are of two types.
(I) Less than series (II) More than series

Suppose. We are given the following discrete series of marks obtained by 100 students with the help of this series. We shall form the Less than and More than series.

Marks obtained:	30-40	40-50	50-60	60-70	70-80
No of students:	8	12	20	25	18

PRACTICAL PROBLEMS IN FORMING A FREQUENCY DISTRIBUTION

The following practical problems are generally faced while forming a frequency distribution:

1. The number of classes should **neither be too small nor too large**. As far as possible open-ended classes should be avoided, as these create problems in analyzing and interpretation.
2. Class interval is obtained by **dividing the range with the number of Classes**. Again this may create problem unless the classes are of equal width. Sometimes however, Unequal class intervals become necessary.
3. While recording the data in frequency tally, one has to be careful in putting the tally marks. Especially when the number of observations is large
4. Once tally marks are obtained, the frequencies are obtained by counting the tallies. One should be careful while counting the tallies.

RELATIVE FREQUENCY AND PERCENTAGE FREQUENCY OF A CLASS INTERVAL

Relative Frequency

- Frequency of each class can also be expressed as a fraction or percentage terms. These are known as relative frequencies.
- In other words, a relative frequency is the class frequency expressed as a ratio of the total frequency

$$\textbf{Relative Frequency} = \frac{\textbf{Class Frequency}}{\textbf{Total frequency}}$$

Percentage Frequency

Percentage frequency of a class interval may be defined as the ratio of the class frequency to the total frequency expressed as a percentage.

$$\textbf{Percentage Frequency} = \frac{\textbf{Class Frequency}}{\textbf{Total frequency}} \text{ X } 100$$

BIVARIATE OR TWO-WAY FREQUENCY DISTRIBURRION

In many situations simultaneous study of two variables becomes necessary. The frequency distribution so obtained a result of this cross classification of two variables give rise to bivariate frequency distribution and this can be summarized in the form of a two-way table called Bi-variate frequency Table.

GRAPH OF FREQUENCY DISTRIBUTION

The graphs of frequency distribution are designed to present the characteristic features of a frequency data. They facilitate comparative study of two or more frequency distributions regarding their shape and pattern.

The most **commonly used graphs** are:
1. Histogram
2. Frequency polygon
3. Frequency curve

4. Cumulative Frequency curve or Ogive

Histogram

- A Histogram is a graph containing a set of **class interval by its width** and the **frequency in each class interval by its height**.
- The area of each rectangle is proportional to the frequency in the respective class-interval and the total area of the histogram is proportional to the total frequency.
- A histogram is used to depict a frequency distribution.
- In constructing a histogram the class-boundaries are erected side by side on the graph. The vertical rectangles are erected side by side on the basis of the frequencies over the class-boundaries.
- This type of diagram is used exclusively for showing frequency distributions of quantitative data that are **continuous in nature**.

REMARKS:

1. It is not necessary that the scale on X-axis and y-axis be the same. Different scales may be taken on the two axes.
2. The position of origin on the y-axis is according to scale which may not be so on the x-axis. If class intervals are not starting from zero then it is indicated by drawing a Kink mark (-√ \/ \/) on the X-axis near the origin. If necessary the Kink mark may be made on y-axis or on both the axes.
3. A histogram may not be confused with a bar diagram. In a bar diagram the height of each bar matters and not its width whereas in a histogram height as well as the width of each rectangle matters.

Types of Histograms

(a) Histograms with equal class intervals
(b) Histograms with unequal class intervals.

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.

The curve is drawn free hand in such a way that the area included under the curve is approximately the same as that of the frequency polygon.

The basic object of drawing a frequency curve is to present graphically the area covered by histogram in a more symmetrical manner. It is also known as smooth frequency curve.

CUMULATIVE FREQUENCY CURVE OR OGIVE

It is a graph which represents the data of a cumulative frequency distribution.

Types
Cumulated downward - Less than Ogive
Cumulated upward - More than Ogive

Uses

- All Ogive is used to find median, quartiles, deciles and Percentiles etc.
- It is also used to find the number of observations which are expected to lie between two given values.