PROCESSING OF DATA

Data processing is generally "the collection and manipulation of items of data to produce meaningful information". In this sense it can be considered a subset of information processing, "the change of information in any manner detectable by an observer".

CLASSIFICATION OF DATA

Classification refers to a process, wherein data is arranged based on the characteristic under consideration, into classes, or groups, as per resemblance of observations. Classification puts the data in a condensed form, as it removes unnecessary details that helps on easily comprehend data.

The data collected for the first time is raw data and so it is arranged in haphazard manner, which does not provide a clear picture. The classification of data reduces the large volume of raw data into homogeneous groups; i.e data having common characteristics or nature are placed in one group and thus, the whole data is bifurcated into a number of groups. There are four types of classification.

* Qualitative classification or Ordinal classification
* Quantitative classification
* Chronological or Temporal classification
* Geographical or Spatial classification

TABULATION OF DATA

Tabulation refers to a logical data presentation, wherein raw data is summarized and displayed in a compact form, i.e. in statistical tables. In other words, it is a systematic arrangement of data in columns and rows, that represents data in concise and attractive way. One should follow the given guidelines for tabulation.

* A serial number should be allotted to the table, in addition to the self explanatory title.
* The statistical table is required to be divided into four parts, i.e. Box, Head, Stub, caption and Body. The complete upper part of the table that contains columns and sub-columns, along with caption, is the Box Head. The left part of the table, giving description of rows is called stub. The part of table that contains numerical figures and other content is its body.
* Length and width of the table should be perfectly balanced.
* Presentation of data should be such that

it takes less time and labour to make comparison between various figures.

*Footnotes, explaining the source of data or any other thing, are to be presented at the bottom of the table.

PRESENTATION OF DATA :

Presentation of data is of utter importance nowadays. Presentation of data refers to an exhibition or putting up data in an attractive and useful manner such that it can be easily interpreted. The three main forms of presentation of data are:

*Textual presentation

*Data tables

*Diagrammatic presentation.

DIAGRAMMATIC PRESENTATION

BAR DIAGRAM

*There are two types of bar diagrams namely Horizontal Bar diagram and vertical bar diagram.

*While horizontal bar diagram is used for qualitative data or data varying over space, the vertical bar diagram is associated with quantitative data or time series data.

*Bars i.e. rectangles of equal width and usually of varying lengths are drawn either horizontally or vertically.

*We consider Multiple or Grouped Bar diagrams

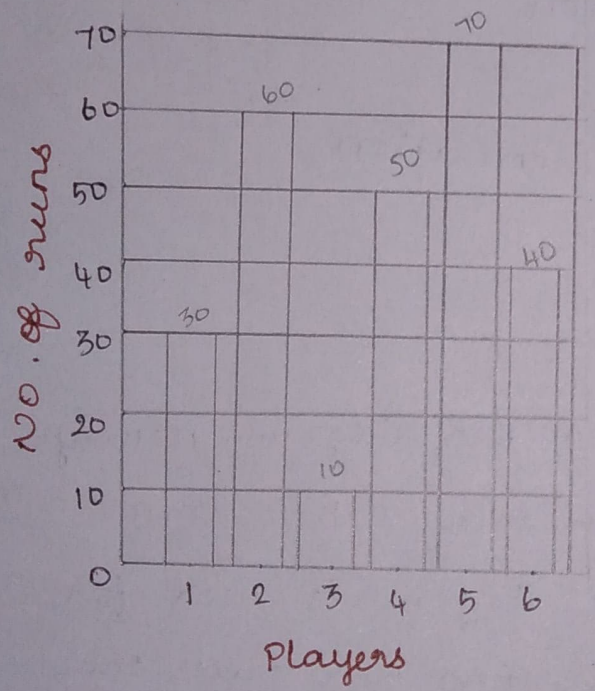
to compare related series. component or sub-divided Bar diagrams are applied for representing data divided into a number of components. Finally, we use Divided Bar charts or Percentage.

*Bar diagrams for comparing different components of a variable and also relating of the components to the whole.

Example:

The total number of runs scored by a few players in one-day match is given.

Players	1	2	3	4	5	6
No. of runs	30	60	10	50	70	40



PIE CHART

In a pie chart, the various observations or components are represented by the sectors of a circle and the whole circle represents the sum of the value of all the components. Clearly, the total angle of 360° at the center of the circle is divided according to the values of the components.

The central angle of a component is

= [value of the component / Total value] × 360°

Sometimes, the value of the components are expressed in percentages. In such cases,

The central angle of a component is

= [percentage value of the component / 100] × 360°

GRAPHICAL PRESENTATION

HISTOGRAM

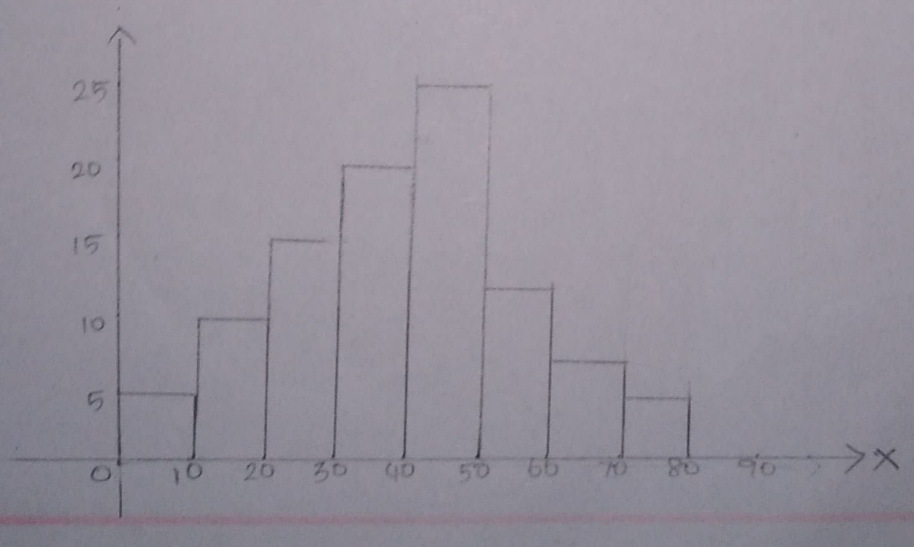
*A two dimensional graphical representation of a continuous frequency distribution is called a histogram.

*In histogram, the bars are placed continuously side by side with no gap between adjacent bars.

*That is, in histogram rectangles are erected on the class intervals of the distribution. The areas of rectangle are proportional to the frequencies.

Example

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of Students	5	10	15	20	25	12	8	5



FREQUENCY POLYGON

*Frequency polygon is another method of representing frequency distribution graphically.

*Obtain the frequency distribution and compute the mid points of each class interval.

*Represent the mid points along the X-axis and the frequencies along the y-axis.

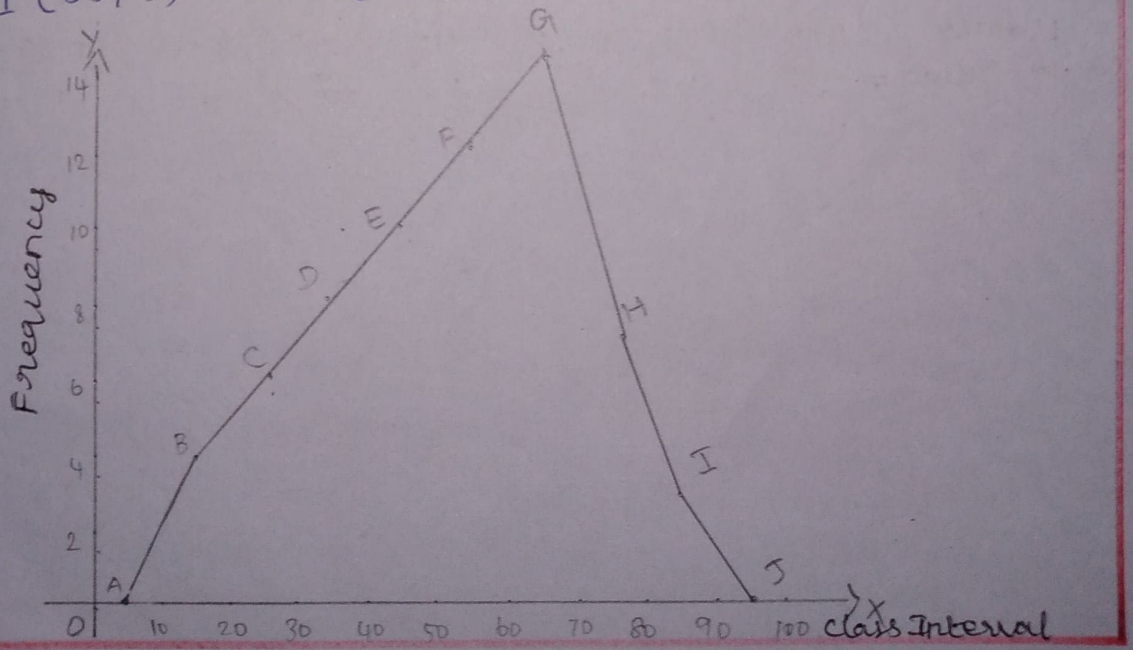
*plot the points corresponding to the frequency at each mid point.

*Join these points, by straight lines in order.

*To complete the polygon join the point at each end immediately to the lower or higher class marks on the X-axis.

Class Interval	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
Midpoints	5	15	25	35	45	55	65	75	85	95
Frequency	0	4	6	8	10	12	14	7	5	0

using the adjacent table, plot the points A(5,0) B(15,4) C(25,6) D(35,8) E(45,10) F(55,12) G(65,14) H(75,7) I(85,5) and J(95,0).



References :-

Net sources - Richard