

① Define statistics as a subject of study.

⇒ Statistics is the collection of concepts and methods used for collection, analysis and interpretation of statistical data.

Note:- Statistics also means numerical information and the word statistics means a constant computed on the basis of observations of a sample.

② What are the limitations of statistics?

⇒ (i) Statistics doesn't deal with individual item but with sum/aggregate of items.

(ii) Statistics deals with quantitative data only. Quantitative data are studied after converting them to numerical data by method of ranking, scoring or scaling etc.

③ Explain what is meant by statistical data?

⇒ Numerical data obtained in the face of uncertainty and variability ~~are two~~ constitute statistical data. Uncertainty and variability are two characteristics of statistical data. Data are used for making deductions, generating informations, reaching conclusions and thus form the basis of interference.

4) State the difference betⁿ Quantitative and Qualitative statistical data.

⇒ Data obtained from phenomena which can't be measured quantitatively such as beauty, honesty, poverty etc. are called qualitative data. Quantitative data are also known as statistics of attributes. Since direct numerical measurement isn't possible, attributes are studied by observing its presence or absence, by counting number of units possessing a given attribute, by finding the proportion or percentage of units possessing a given attribute.

Data obtained from phenomena which can be quantitatively measured are known as quantitative data. For example heights, weights, length, etc. Quantitative data are also known as statistics of variables.

5) What are primary data?

⇒ Data collected from the source directly from the first time are called primary data. e.g. Data collected by a market research group from the market directly constitutes primary data.

6) What are the secondary data?

⇒ Data collected from a secondary source constitutes secondary data i.e. Data collected from some records or individual or organisations where data were already available. For example Data collected from the market research group by a company launching a new product constitutes secondary data.

⑦ Distinction between primary data and secondary data.

- (i) Primary data is actually a raw data as it is collected for the 1st time but secondary data is process ~~present~~ data.
- (ii) Primary data is collected from the field of enquiry but the secondary data is collected from the previous records.
- (iii) Primary data fulfill's the desire of the investigators. But secondary data may not serve all purposes as desire.
- (iv) Primary data is used only by investigators. But secondary data may be used by anyone.
- (v) All secondary data was initially primary. But primary data may not be secondary data.

⑧ Census enumeration and sample enumeration.

- When all the data are collected then it is known as complete or census enumeration. As in the case of census in our country we collect all the data of human being, their age, literacy etc.
- Instead of taking total information, if we select randomly some parts of it and collect information from the parts, it is known as sample enumeration. When it becomes difficult to collect all the individual we used this sample enumeration.

⑨ Comparison betⁿ census and sample enumeration.

Census	Sample
(i) It is costly.	(i) It costs less.
(ii) It is time consuming	(ii) It is time saving.
(iii) It needs more man-power.	(iii) It needs less man-power.
(iv) Training for enumerators should be regularly frame.	(iv) Some training is necessary.
(v) In case of doubt chance of verification is almost nil.	(v) In doubtful case verification is possible.
(vi) Informants pay less interest for wider zone	(vi) Informants pay more interest for closer zone.

⑩ Classification: —

After the collection of data, we find that there is a cluster of haphazard figures. At a glance we can't find any details about data, and naturally no inference can be drawn from that list. This is ~~the~~ that is why classification is essential. So to say the classification is the arrangement of data in groups or classes in accordance with their resemblances.

① Objects of classification:—

Elimination of unnecessary details, clear division into similar and dissimilar quantities, settling up mentally for further analysis and drawing inference regarding similar objects on comparison with others are the main objects of classification.

② Basis of classification:—

Statistical data collected are divided into two broad basis:—

(i) Descriptive

(ii) Numerical

(i) Descriptive basis are those which can't be expressed by numerical measurement. Examples of these are ~~age~~, gender, nationality, marital status, literacy etc.

(ii) Nu These are all the qualitative characters of the individuals and they are called attributes

(ii) Numerical basis are those which are expressed by numerical measurements such as height, weight, temperatures etc. These are all quantitative measurement and are called variables.

⑫ Classification of attributes.

→ ① Simple Classification:—

The classification is of 'yes' or 'no' type nature. For any character say literacy classification can be done as literate or illiterate. For any such character it will be either 'yes' or 'no'. So classification involving one attribute is the simple classification.

② Many fold classification:—

When two or more characters are studied such as literacy and gender, then both divisions are possible like —

(i) male - literate

(ii) Female "

(iii) Male - illiterate

(iv) Female "

So the classification involving two or more attributes are called many fold classification.

⑬ Classification of variables.

Variables are divided into two categories

(i) Continuous

(ii) Discrete

(i) The variable which can take any fractional value within a specified range is called continuous variable. Eg. height, weight, distance

(ii) The variable which can take some distinct values maintaining some gap between the values is called the discrete variable. Eg. the runs scored by the batsmen in different matches.

15) Tabulation:— If we say classification is the theoretical idea, if we then tabulation is the practical part of classification. It is the systematic arrangement of data into rows and columns. It is actually the final step of the collection and presentation of data, from which the stepping of analysis efforts. So wise tabulation serves the serve the purpose the investigation properly.

16) Advantages of tabulation:—

- (i) It becomes a nice display of data through tabulation.
- (ii) Quick comparison in between the similar type of data is possible.
- (iii) Easy detection of errors and omissions is possible.
- (iv) Creates ~~impression~~ ^{impression} over the investigations.
- (v) Proper tabulation needs no such further explanation, description regarding the collection of data.

17) Main features of tabulation:—
(Different parts of a table)

The proper tabulation should possess

- (i) Title: It contains a brief description or explanation regarding the collection of data.
- (ii) Captions: Descriptions of the columns are given in the captions. It may contain sub-captions also.
- (iii) Stubs: Description about the rows are given in stub portion. It is placed in extreme left portion of the table.
- (iv) Body: All the numerical data are placed in the body portion.

(vi) Head note: It is placed just under the title which contains detail description if any other than those in the title.

(vii) Footnote:— Any additional as well as striking information regarding data may be placed under the table.

(viii) Source note:— For any uncommon information, the source, where the data is collected should be given. It helps the investigators for further query and also for verification if necessary.

18) Diagrammatic representation of a Table.

→ Title: The number of students admitted in the college in the year 2000-01.
 Head note: (In Science/Humanities/Commerce Section)

← STUBS →	← CAPTIONS →
College Name	General OBC SC/ST M F M F M F ← (Sub captions)
College — A	
College — B	B O D Y
.....	
College — N	

Footnote: (i) The students of General and Honours courses are taken together.
 (ii) The transferred candidates are not counted.

Source: The admission register of the corresponding colleges.

(19) Diagrammatic representation:—

Diagrammatic representation is one kind of process to display statistical results, obtain from collected and tabulated data, visually.

Advantages:—

- (i) It needs no mental taxation to draw inference regarding comparison.
- (ii) The person who are not acquainted ~~the~~ with the theory and practice of statistics, may also ~~involved~~ be convinced through the diagram.
- (iii) It has more attraction than the figures to the common people.
- (iv) Quick comparison between two similar types of data is possible just at a glance.
- (v) The nature and trend of data and particularly the difference between the two data become more memorable through diagrams rather than the figures.
- (vi) For publicity purpose it is ~~unparallel~~, unparallel.

Disadvantage:-

- (i) It doesn't help in tracing statistical analysis.
- (ii) The appeal of diagram is mostly confined to the comparison to the data, so when comparison isn't the object the diagram helps a little.
- (iii) Though the diagram originates from the tabulation it can't be an alternative to the tabulation, so it can't be a substitute of statistical data.
- (iv) When two data are large and differ slightly, diagram can't exhibit their difference clearly.
- (v) Diagrams can't clarify all other statistical results except comparison.

② Different types of Diagrams.

→ (a) Line Diagrams:-

Line diagrams are useful for representation of ~~kind~~ ^{time} series of data. Here the time variable is taken along X-axis and the corresponding values along the Y-axis and then the plotted points are joined by the line segment to get the line diagram.

(b) Bar Diagram:-

Here a number of bars of equal width having height proportional to the data values are drawn on the same base side by side at equal distance from one another.

(iii) Pie diagram: - A pie diagram is constructed by dividing a circle into sectors with areas or central angles proportional to the different components. The angles are the relative proportion of the magnitude of the various components.

(iv) Pictogram: - It is a number of pictures of some size each representing a number of units are drawn side by side horizontally to represent the statistical data. This type of diagrams are useful but because of their natural appeal and when the targeted audience with general public, not specialists.

(v) Histogram: - When the data are classified based on the class intervals, they can be represented by a histogram. Histogram is just like a simple bar diagram with minor differences. There is no gap between the bars since the classes are continuous. The bars are drawn only in outline without colouring or making as in the case of bar diagram.

(vi) Frequency polygon: - The type of data required for frequency polygon is the same as that for histogram. The frequency of the classes are plotted by dots against the midpoints of each class. The adjacent dots are then joined by straight line. The resulting graph is known as frequency polygon.

(vii) Frequency curve

The procedure for drawing frequency curve is same as for frequency polygon but the points are joint by smooth or free hand curve. The curve may not pass through all points as in the case of frequency polygon.

(viii) Ogive :- Here instead of plotting frequencies, cumulative frequencies (\angle type & \lrcorner type) are plotted by dots and these dots are joint by a free hand curve or by a line.