Sources of Data and it's Measurement

Sumon Kumar Das

Department of Geography, Netaji Nagar Day College, Kol-92

There are two sources of data in Statistics. Statistical sources refer to data that are collected for some official purposes and include censuses and officially conducted surveys. Non-statistical sources refer to the data that are collected for other administrative purposes or for the private sector.

Statistical Survey

A statistical Survey is normally conducted using a sample. It is also called Sample Survey. It is the method of collecting sample data and analyzing it using statistical methods. This is done to make estimations about population characteristics.

Census

Opposite to a sample survey, a census is based on all items of the population and then data are analyzed. Data collection happens for a specific *reference period*. For example, the Census of India is conducted every 10 years. Other censuses are conducted roughly every 5-10 years. Data is collected using questionnaires that may be mailed to the respondents.

Responses can also be collected over other modes of communication like the telephone. An advantage is that even the most remote of the units of the population get included in the census method. The major disadvantage lies in the high cost of data collection and that it is a time-consuming process.

Register

Registers are basically storehouses of statistical information from which data can be collected and analysis can be made. Registers tend to be detailed and extensive. It is beneficial to use data from here as it is reliable. Two or more registers can be linked together based on common information for even more relevant data collection. From agriculture to business, all industries maintain registers for record-keeping. Some administrative registers also serve the purpose of acting as a repository of data for other statistical bodies in a country.

Types of Data and Data Collection

Like stated earlier, there are two types of data: primary and secondary.

Primary data

As the name suggests, are first-hand information collected by the surveyor. The data so collected are pure and original and collected for a specific purpose. They have never undergone any statistical treatment before. The collected data may be published as well. The Census is an example of primary data.

Methods of primary data collection:

1. Personal investigation: The surveyor collects the data himself/herself. The data so collected is reliable but is suited for small projects.

2. Collection Via Investigators: Trained investigators are employed to contact the respondents to collect data.

3. Questionnaires: Questionnaires may be used to ask specific questions that suit the study and get responses from the respondents. These questionnaires may be mailed as well.

4. Telephonic Investigation: The collection of data is done through asking questions over the telephone. to give quick and accurate information.

Secondary data

Secondary data are opposite to primary data. They are collected and published already (by some organization, for instance). They can be used as a source of data and used by surveyors to collect data from and conduct the analysis. Secondary data are impure in the sense that they have undergone statistical treatment at least once.

Methods of secondary data collection:

1. Official publications such as the Ministry of Finance, Statistical Departments of the government, Federal Bureaus, Agricultural Statistical boards, etc. Semi-official sources include State Bank, Boards of Economic Enquiry, etc.

2. Data published by Chambers of Commerce and trade associations and boards.

3. Articles in the newspaper, from journals and technical publications.

Measurement scale of different types of data in statistics:

Four stages have been mentioned in the measurement scale: –

1. Nominal
2. Ordinal
3. Interval
4. Ratio

Details of stages:

1. Nominal

The word 'Nominal' comes from the word 'Name' or 'naming'. The method of measuring the representation of an object, etc., when it is represented by various symbols or numbers by their quality and character, is called the measure of the name index.

Example: – People are divided into two groups, male and female Mark and split.

Features

a. Qualitative quantities apply to information.

b. It is used to include objects in different categories of persons etc.

 Advantages

a. The simplest method of measuring.

b. The method only determines the character of the object. As a result, it is easy to determine which object is classified by measurement.

Difficulty

a. Analysis of the character of data. But no amount is capable of dimensional analysis.

b. To determine the order and spacing of two adjacent classes Disabled.

2. Ordinal

It can be used to represent the order of any qualitative data values. Values are important and significant, but the difference between the two is not really known.

Example: –we know that # 8 is better than # 5 or # 4 Well, and can't measure how good it is. Similarly, we cannot tell the difference between "very good" and "good."

Features

a. This scale only applies to qualitative quantity information.
b. This scale is based entirely on experience.

Advantage

a. Ordinal scale is located at the next higher level of the nominal scale.

b. It is able to maintain important religious equality.

Difficulty

a. Ordinal scale or object or position can reflect the level only. But the interval between each level is never possible to determine.

b. It is only able to determine the relative position of an object or position. So no mathematical analysis, such as – addition, subtraction, multiplication, and division is not possible by this scale.

3. Interval

"Interval" means "space in between". The interval scale is the numerical scale for representing quantitative information. On this scale, we know the exact difference between both order and values. For example, the difference between 5 and 15 cm is a measurable 10 cm. This is the same difference between 15 and 25 cm.

Features

a. scale has no absolute zero value. That is, such a measurement does not begin with zero, or Even if the scale starts from zero for some reason, it is inappropriate to assume that the value of the material cannot fall below zero. For example, in the CGS method, the unit of measurement for temperature is "C," but D ° C does not mean that there can be no temperature less than 0 ° C.
b. As the "interval scale" does not indicate absolute zero, so this scale does not have a maximum level. For this reason, the variability of a data cannot be measured by scale. For example, the temperature of two places A and B can never be said to be 20 ° C and 40 "C respectively, the temperature of B is twice that of A, because the value of O ° C on this scale is relative, again it has no maximum value.

Advantages: –

a. Methods are not only classified by order, but also by each class or order. The interval is determined by a precise unit.
b. It is known by the interval scale which position is in which position of an order.

Difficulty

a. The biggest limitation of the "interval scale" is that there is no absolute zero on this scale.

b. The exact position of the value can be determined by this scale, but the difference in intensity between the two values can never be determined. For example, a temperature of 40 ° C cannot be said to be twice the value of 20 ° c .

4. **Ratio**

Ratio scales are special measurement scales because they are accurate in our order, units say about values, and they also have an absolute zero, which allows a wide application of both descriptive and inferential statistics. Variables on the ratio scale can be meaningfully the same when multiplying or dividing the same number with them. The scale has great use of statistics.

Features

a. All types of mathematical and numerical analysis can be done by Ratio scale.
b. This scale is superior to other scales. So this scale widely uses in research work.

Advantages

a. The main advantage of the ratio scale is that it contains absolute zero.
b. Any quantitative and mathematical analysis of variables such as addition, subtraction, multiplication and division can all be done by the ratio scale.

Conclusion:

Data is a major component of statistics. Therefore, it is very important to know the nature of numeric information. The nature, advantages, disadvantages, and different types of arithmetic information help us for the explanation of a fact. The content gives us some ideas about the information and makes the reader interested to know more about it. So these are the Different types of data in statistics which is very important.

References:

1. Das,N.G.(2005), Statistical Methods, M.Das`& Sons, New Delhi.
2. Gupta, S.P. (1976) Elementary Statistical Methods, Sultan Chand & Sons, New Delhi.
3. Nadar, N.E.(2015) Statistics, PHI Learning Private Limited, Delhi.