

Study Material for the students of:

1. Ph.D. course work

2. M.A. (4th semester)

DATA, TYPES AND TECHNIQUES OF COLLECTION

Prepared by:

Dr. Sangram Bhushan

INTRODUCTION

A sound structure of statistical investigation is based on a systematic collection of data. The task of data collection begins after a research problem has been defined and research design plan chalk out. While deciding about the method of data collection to be used for the study the researcher should keep in mind two types of data viz primary and secondary data the nature, object and scope of the enquiry are the most important things on which the selection of the method depends. The methods selected for the collection of data should be such, that it suits the type of enquiry that is being conducted.

How to define Data?

Data is a collection of facts, figures and statistics related to an object from which conclusion may be drawn. The term data means group of information that represent qualitative or quantitative attributes or set of variables. Data can be processed to create useful information. It is the base on which the superstructure of statistical investigation is made. Collection of data is the first step in any statistical investigation the success and failure of investigation mainly depend up on the quality of data.

TYPES OF DATA

Data is generally classified in to two groups, viz:

- a. Primary Data
- b. Secondary Data

Primary data:

Primary data are those which are collected a fresh and for the first time and thus happen to be original in character. Primary data are the actual information which is received by researcher for the study from the actual field of research. Primary data are in the shape of raw materials to which statistical methods are applied for the purpose of analysis and interpretation.

Secondary data:

The secondary data, on the other hand are those which have already been collected by some one else and which have already been passed through the statistical process. These data are usually in the shape of finished products since they have been treated statistically. Primary data involves large expenses in terms of time energy and money while secondary data is relatively a less costly method.

SELECTION OF APPROPRIATE METHOD FOR DATA COLLECTION

There are various methods of data collection. As such the researcher must judiciously select the method for his own study, keeping in view the following factors:

1. Nature, scope and object of enquiry:

This constitutes the most important factor affecting the choice of a particular method. The method selected should be such that it suits the type of enquiry that is to be conducted by the researcher.

2. Availability of funds:

Availability of funds for the research project determines to a large extent the method to be used for the collection of data. When funds at the disposal of researcher are very limited he will have to select a comparatively cheaper method which may be efficient and effect as compared to other methods.

3. Time factor:

Availability of time has also to be taken in to account in deciding a particular method of data collection some methods take relatively more time, while as with others the data can be collected in a comparatively shorter duration. The time at the disposal of the researcher, thus affects the selection of the method by which the data are to be collected.

4. Precision required:

Precision required is yet another important factor to be considered at the time of selecting the method of collection of data

TECHNIQUES OF COLLECTION PRIMARY DATA

The following methods of collection of primary data are in common use.

- a. Observation method
- b. Interview method
- c. Through questionnaires
- d. Local reports
- e. Audits
- f. Through schedules

Observation method:

The observation method is the most commonly used method especially in studies relating to behavioural science. Observation becomes a scientific tool and the method of data collection for the researcher when it serves as a formulated research purpose, is systematically planned and recorded and is subjected to checks and controls on validity and reliability. Under the observation method, the information is sought by way of investigators own direct observation without asking from the respondent. For instance in a study relating to consumers behaviour the investigator instead of asking the brand of wrist watch used by the respondent may himself look at the watch. The main advantage of this method is that the subject bias is eliminated if observation is done accurately.

Advantages of observation method:

1. The main advantage of observation method is that the subjective bias is eliminated.
2. This method is independent respondent's willingness to respond.
3. The information obtained under this method relates to what is currently happening is not complicated by either the past behaviour or further intentions or attitudes.

Limitations:

1. Observation method is an expensive method.
2. The information provided by this method is very limited.

3. Sometimes unforeseen factors may interfere with the observation task.

Participant and non-participant observation:

If the observer observes by making himself more or less a member of the group he is observing so that he can experience what the members of the group experiences the observation is called as the participant observation. But when the observer observes as a detached emissary on his part to experience through participation what other feed, the observation of this type is often termed as non-participant observation.

b. INTERVIEW METHOD

The interview method of collecting data involves presentation of oral verbal stimuli and reply in terms of oral-verbal responses. This method can be used through personal interviews and if possible through telephone interviews.

a. Personal interviews: personal interview method requires a person known as the interviewer asking questions generally in a face to face contact to the other persons. This sort of interview may be in the form of direct personal investigation or it may be indirect oral investigation. The interview has to collect the information personally from the sources concerned.

The method of collecting information through personal interviews is usually carried out in a structured way. As such we call the interviews as structured interviews involve the use of a set of predetermined question and of highly standardized techniques of recording so a rigid procedure is laid down asking question in a form and order prescribed. The unstructured interviews are characterized by a flexibility of approach to questioning.

Merits:

1. More information and that too in greater depth can be obtained.
2. There is greater flexibility under this method as the opportunity to restructure questions is always there.

3. It can be applied to recording verbal answers to various questions.
4. Samples can be controlled more effectively as the missing returns, non-response generally remains very low.

Demerits:

1. It is very expensive method, especially when large and widely spread geographical sample is taken.
 2. There remains the possibility of bias of the interviews as well as that of respondent.
 3. This method is relatively more time consuming method.
 4. The presence of the interviewer on the spot may over stimulate the respondent.
- b. Telephone interview:** this method of collecting information consists in containing respondents on telephone itself. It is not very widely used method, thus plays important part in industrial surveys particularly in development regions.

Merits:

1. It is more flexible in comparison to mailing method.
2. It is cheaper than interview method.
3. It is faster than other methods.
4. No field staff is required.
5. There is a higher rate of response than what we have in mailing method.

Demerits:

1. Little time is given to the respondent for considered answers.
2. Surveys are restricted to respondents who have telephone facilities.
3. It is not suitable for intensive surveys where comprehensive answers are required.
4. Possibility of bias of the interviewer is relatively more.

c. COLLECTION OF DAT THROUGH QUESTIONNAIRES

This method of collection data is quite popular particularly in case of big enquiries. It is being adopted by private individuals, research workers, private and public organization

and even by governments. A questionnaire consists of a number of questions printed or typed in a definite order on a form or set of forms. The questionnaire is mailed to respondents who understand the questions and write down the reply in the space meant for the purpose in the questionnaire

It is an important instrument in normative survey research, being used to gather information from widely scattered sources.

Purposes of the questionnaire are two fields:

1. To collect information from the respondent who are scattered in vast area.
2. To achieve success in collecting reliable and dependable data.

FORMS OF QUESTIONNAIRE

STRUCTURED VS NON-STRUCTURED:

The structured contains definite, concrete and direct questions, where as non-structured may consist of partially completed questions or statements. A non-structured questionnaire is often used as the interview guide who is non-directive. The interviewer possesses only a blue print of the enquiries and he is largely free to arrange the form or statements of the questions

CLOSED FORM VS OPEN FORM:

The questions that call for short check responses are known as restricted or closed form type. They provide for making a yes or no, a short response or checking an item out of a list of given responses. He has simply to select a response out of supplied responses and has not to frame his response in own way.

It is easy to fill out, takes less time, keeps the respondent on the subject, is relatively more objective, more acceptable and convenient to respondent and is easy to tabulate and analyze.

The open form or open end type questionnaire calls for a free response in the respondents own words. The respondent frames and supplies his own reasons. No clues are provided. It probably provides for greater depth of response. The subject reveals his mind, gives his frame of reference and possibly the reasons for his response. This type

of item is sometimes is difficult to interpret, tabulate and summaries in the research report.

THE MIXED QUESTIONNARIE:

The mixed questionnaire consists of both close and open type questionnaire. For social research this method is very useful. Many questionnaires include both open and closed type item each type has its specific merits and limitations.

FACT AND OPANION QUESTIONNARIES:

Questionnaire which requires certain information of facts from the respondent without any reference to his opinion or attitude about them is called fact questionnaire. The questionnaire of opinion, is one in which the informant's opinion, attitude or preference regarding some phenomenon is sought.

PICTORIAL QUESTIONNMARIES:

In the pictorial questionnaire, pictures are used to promote interest in answering questions. It is extensively used in studies of social attitudes and prejudices in children or illiterate persons. This questionnaire may be very useful for collecting data in a developing country like India, especially from the rural area masses that are mostly illiterate and less knowledgeable. The serious notation of this questionnaire is that it is lengthy in form. Also it is highly expensive.

ESSENTIALS OF A GOOD QUESTIONNARIES

1. To be successful, questionnaire should be comparatively short and simple i.e. the size of the questionnaire should be kept to the minimum.
2. Questions should be proceeding in logical sequence moving from easy to more difficult question.
3. The language should be simple, use wards of precise (not flexible) meaning, useful to insert a transitional phrase, at the end use an expression of appreciation for the efforts pup broth by the respondent.
4. Technical terms and vague expression capable of different interpretation should be avoided in a questionnaire.

MERITS OF THE QUESTIONNAIRES

1. Economical:

Since the questionnaire approach makes it possible to cover at the same time a large number of people over a large territory. It is decidedly more economical in terms of money, time and energy.

2. Time saving:

Besides saving money, the mailed questionnaire also saves time. Simultaneously hundreds of persons are approached through it where as if they are to be interviewed it may take a very long time.

3. Less pressure on the respondent:

The method places less pressure on the subject for immediate response and gives more time to the respondent for properly answering questions.

4. Ensures anonymity:

As the respondents are not required to indicate their names on the questionnaire they feel free to express their views and opinions.

5. Suitable in certain type of response:

The information about certain problems can be best obtained through questionnaire method. For example the research about sexual habits marital relations dreams etc.

LIMITATIONS OF QUESTIONNAIRES

1. The mailed questionnaire can be used only for educated people. This restricts the number of respondents.
2. The return rate of questionnaires is low. The common return rate is 30-40 %.
3. Sometimes different respondents interpret questions differently. The misunderstanding cannot be corrected.
4. The mailing address may not be correct which may omit some eligible respondents. Thus the sample selected may at a time is described as biased.
5. Questionnaire does not provide an opportunity to collect additional information while they are being completed.
6. Many questions remain unanswered. The partial response affects the analysis.

7. The respondent can consult other persons before filling in the questionnaire. The response there fore cannot be viewed as his opinion.

d. COLLECTION OF DATA THROUGH SCHEDULES

This method of data collection is very much like the collection of data through questionnaire with little difference which lies in the fact that schedules (Proforma containing a set of question) are being filled in by the enumerators who are specially appointed for the purpose. These emulators along with schedules go to respondents, put to them the question from the Performa in the order the questions are listed and record the replies in space meant for the same in the Performa. Enumerators explain the aims and objects of the investigation and also remove the difficulties which any respondent may feel in understanding the implications of a particular question or the definition or concept of difficult terms.

PURPOSE OF SCHEDULE

THE MAIN PURPOSE OF SCHEDULE IS THREE-FOLD:

1. To provide a standardized tool for observation or interview in order to attain objectivity.
2. To act as memory tickler i.e. the schedule keeps the memory of the interviewer.
3. To facilitate the work of tabulation and analysis.

MERITES

1. This method of data collection is very useful in extensive enquiries.
2. It leads fairly reliable results.

DEMERITES

1. This is very expensive method.
2. This method is relatively more time consuming method.

LOCAL REPORTS

In this method the data are not formally collected by the enumerators but by the local respondents or agents in their own fashion and to their own likings. Obviously such data can not be very reliable and as such this method is used in those cases where the

purpose of investigation can be served with rough estimates only and where a high degree of precision is not advantage of being least expensive and it also saves the botheration usually associated with the statistical investigation of other types.

PANTRY AUDITES

Pantry audit technique is used to estimate consumption of the basket of goods at the consumer level. In this type of audit the investigator collects an inventory of types, quantities and price of commodities consumed. Thus in a pantry audit data are recorded from the examination of the certain products and certain brand the assumption being that the contents of the pantry accurately portray consumers' preferences. A pantry audit may or may not be set up as a panel operation since a single visit is often considered sufficient to yield an accurate picture of consumer's preference. An important limitation of pantry audit approach is that at times it may not be possible to identify consumer's preferences from the audit data alone particularly when promotion devices produce a marketed rise in sales.

CONSUMER PANALS

A extension of the pantry audit approach on a regular basis is known as consumer panel where a set of consumer are arranged to come to an understanding to maintain detailed daily records of their made available to investigator on demands. In other wards a consumer panel is essentially a sample of consumer who is interviewed repeat over a period of time mostly consumer panels are of two types viz the transitory consumer panel and containing consumer panel. A transitory consumer panel is set up to measure the effect of a particular phenomenon usually such panel is conducted on a before and after basis. Initial interviews are conducted before the phenomenon takes place to record the attitude of the consumer.

SOURCES OF SECONDARY DATA

Secondary data means data that are already available i.e., they refer to the data which have already been collected and analysed by someone else. When the researcher utilizes

secondary data, then he has to look in to various sources from where he can obtain them. In this case he is certainly not confronted with the problem that is usually associated with the collection of original data. Secondary data may either be published data or unpublished data.

Sources of published data are:

- a. Official publication of central, state and local governments.
- b. Official publication of the foreign government or international bodies like United Nations organization and its subsidiary bodies.
- c. Reports and publications of the trade associations, chambers of commerce, banks, co-operative societies, stock exchanges and trade unions etc.
- d. Technical trade journals like economics, Indian journal of economics, commerce, capital etc. and books and newspapers.
- e. Report submitted by economists. Research scholars, university bureaus and various other educational associations etc.

The sources of unpublished data are many; they may be found in diaries, latter, unpublished biographic and autobiographic and also may be available with scholar and research workers, trade association, labour bureaus and other public/private individuals and organizations.

PRECAUTIONS IN THE USE OF SECONDARY DATA

The secondary data must be used with caution. It is usually very difficult to verify such data and to edit them to find out in consistencies. In the words of Bowley, “it is never safe to take published statics at their face value without knowing their meanings and limitations and it is always necessary to criticize arguments that can be based on them”. By way of caution the researcher before using secondary data must see that they possess following characteristics:

1. **Reliability:** The reliability can be tested by finding out such things about the said data:
 - (a) Who collected the data?
 - (b) What were the sources of data?
 - (c) Were they collected by using proper methods?
 - (d) At what time were they collected?

- (e) Was there any bias of the compiler?
 - (f) What level of accuracy was desired? Was it achieved?
2. **Suitability of data:** The data that is suitable for one enquiry may not necessarily be found suitable in another enquiry. Hence, if the available data are found to be unsuitable, they should not be used by the researcher.
 3. **Adequacy of data:** if the level of accuracy achieved in data is found inadequate for the purpose of present enquiry, they will be considered as inadequate and should not be used by the researcher. The data will also be considered inadequate if they are related to an area which may be either narrow or wider than the area of present enquiry.

CONCLUSION

The data which is the basic raw material for a researcher must be collected in a systematic and scientific way. Each method of data collection has its uses and none is superior in all situations. So it is necessary that the researcher must judiciously select the method for his study. Before the collection of data the researcher should be aware about the nature, scope, objective of enquiry, availability of funds and time factor. These factors affect the research work to a large extent. While using secondary data one must make a minute scrutiny because it is just possible that the secondary data may be unsuitable or may be inadequate in the context of the problem which the researcher wants to study.