Chapter Three: Survey Methods

When you completed this chapter, you will be able to:

- ✓ Recognise different types of data;
- ✓ decide how the information is to be collected;
- ✓ identify the application area of various survey methods;
- ✓ identify the advantages and disadvantages of using Postal Questionnaire; and
- ✓ recognise the important points in designing questionnaire.

Reference(s): Owen Chapter 16, Newbold Chapter 18
Exercise(s): Seminar 5

Once the sample is selected, the attribute of interest must be measured, i.e. *data must be collected*. Before studying the methods of collecting data, we look at the classification of data first.

Primary and Secondary Data

Primary Data: Data that are used for the ***specified purpose for which they are collected***. They will contain no unknown quantities in respect of method of collection, accuracy of measurements or which members of the population were investigated. Sources of primary data are either censuses or samples.

Secondary Data: *Data that are being used for some purpose other than that for which they were originally collected*. Summaries and analysis of such data are sometimes referred to as secondary statistics.

Internal and External Data

Internal Data: Data are generated from the activities ***within*** a firm. They may be taken from a firm's order book, inventory, payroll, personnel or accounting records.

External Data: Data are obtained from sources ***outside*** the firm. They may come from records of state, local government, other business organisations, etc.

Existing Sources of External Data

1. Government Publication is a prime resources.
   e.g. monthly digest of statistics, financial statistics

2. Annual report of business organisation.

3. Many universities regularly provide information on local and national economic activity.

4. Newspaper, Magazines, and any other possible sources.

External Data are sometimes classified as *primary*, meaning that the organisation gathering the data also publishes or releases them, or as *secondary*, meaning that the data are published by an organisation other than the one that gathered them.

Methods of primary data collection

We now try to answer the first question by outlining the various methods of data collection. In some cases it may be necessary to use a *combination of data collection methods*. The following are various methods used to collect data.

1. Direct Observation

The scope of direct observation is limited. However, it is to observe a phenomenon with your own eyes. This method is concerned with ***what people do rather than why they do it***.

Observation can provide *accurate data, free of biases introduced by interviewers*. This method is particularly useful in ***continuously collecting data*** about routine consumer behaviour.

e.g.1 Suppose you wish to know what items a housewife had taken from a shelf and *considered purchasing, but not actually purchased* in a supermarket.

e.g.2 Suppose that a local authority was interested in how many people used a particular road, it could have people to *observe the number of cars passing a definite spot*.

Advantages:

1. the ***actual actions*** or habits of person are observed.

2. the method is applicable when it is undesirable for ***people to know an experiment is taking place*** since their actions would change and the experiment would be spoiled.

3. observation provides one of the most reliable methods of data collection especially under the circumstances where other methods may not be feasible.

Disadvantages:

1. result of observation *depend on* ***the skill*** of the observer;
2. ***opinions and attitudes*** *cannot be obtained by observation;*
3. forms of behaviour such as the frequency of a person's purchase, church going, smoking and crossing roads cannot be obtained by ***'one-time'*** observation, but by a continuous and lengthy period of observation;
4. it would be ***expensive*** *to tie up personnel* in such tasks.

A balance may have to be drawn between the accuracy of the method and the cost of a less accurate but cheaper alternative.

2. Interview

A very similar technique to the method of direct observation.

In this case, the people have to be ***interviewed***.

Personal Interview:

Advantages:

1. It discovers not only what people are, do or *think*, but also *why* may be so. This method generates very rich data sources, *both* ***quantitatively and qualitatively***.
2. It normally achieves a ***high response rate*** because people usually respond when confronted in person. It would be appropriate where we were engaged in any form of *attitude study*, such as the attitude of customers towards the services they obtained from bank cashiers.
3. Well-trained interviewers may assess the person being interviewed in terms of age and social class, and even sometimes ***assess the accuracy of the information*** given. Moreover, help may be given to those respondents who are unable to understand the questions.

Disadvantages:

1. It is probably the ***most expensive*** *data collecting methods*.
2. Interviewers must also be ***well trained***: it is of serious concern if an interviewer deviates from required protocol and introduces bias in results.
3. People may not like to give *embarrassing* information in a face-to-face interview.
4. Some types of people are more difficult to locate and interview than others, e.g. travellers.

Telephone Interview:

Advantages:

1. A great advantage of telephone interviews is their *speed* and *relative economy* compared with personal interview, particularly when ***only a limited amount information is required***.
2. In recent years, the method of *computer-assisted telephone interviewing* is becoming popular among Hong Kong's large research companies. This trend definitely increases *data input accuracy and saves on labour costs.*

Disadvantages:

1. Asking questions by telephone is not usually a good method, because ***refusing to answer questions is easier*** than in a face-to-face situation.

2. ***Time may be wasted*** in phoning people who are not in home.

Conclusions:

Several people can conduct the interviews and provided they are all working to the basic structure, the answers could be aggregated.

The *quality of data* collected using personal and telephone interviews *depend on individuals* who are *willing* and *able* to verbalise their behaviour or attitudes. In certain circumstances people may not do this, it is not because they are unwilling to do so, but simply because they *do not record* the data required, and are therefore *unable to report* it.

3. Experiments

Suppose that a company intends to introduce a fire fighting equipment using a new technique, they are unable to speak to anyone who has used it in the past. It may be that the only way they can gather data will be to carry out some experiments using the new fire fighting equipment. A set of experiments could be designed and carried out with the results being carefully logged.

4. Postal Questionnaires

One of the most common methods for generating raw data is the questionnaire. A form of questions is designed and sent to people for their completion. The respondents complete the forms and return them.

Advantages :

1. The advantages of using questionnaire are the ***speed*** and the ***cost***. The questionnaires can be sent around the country or indeed the world very quickly, and in low cost.

2. There is no ***interviewer bias***.

3. The respondent has ***enough time to consult*** any necessary information.

Disadvantages :

1. The design of questionnaires requires ***great care***. They differ from personal or telephone interviews, where respondents have a chance to clarify questions. A poorly designed questionnaire can cause *administrative problems* and *incorrect deductions* during statistical analysis of data. The accuracy of a questionnaire design may be checked in pilot studies.

2. The drawback is the ***poor response rate***. A very low percentage of questionnaires are ever returned, and many of the returned questionnaires are very often ***incomplete*** **or** ***wrongly completed*** forms. A respond rate of 20% is considered quite good in the average survey.

3. ***Spontaneous*** answers cannot be collected.

4. The "***wrong***" person may complete the questionnaire.

Questionnaire is still a very popular method of generating data, and the advantages probably outweigh the disadvantages.

Pointers for developing questionnaires

1. Brevity

The questionnaire should be as brief as possible. Nobody likes to be faced with a ten-page questionnaire. The respond rate is low enough without giving a further disincentive.

2. Simplicity

A complicated form may well conceal the real point of the questionnaires. There is a mistaken idea that, in the design of questionnaires, one has to use extremely complicated language. It is not necessary to use four or five words when one would suffice.

3. Ambiguity

The respondent must be in *no doubt* as to what a question means.

A question such as " Have you ever been involved in an accident in the past?", is open to a number of different interpretations. The definition of accident is not given and so to how far back the respondent will be unsure as to whether it is intended to embrace thefts, fires, accidents, or windscreen breakage.

4. Leading Questions

It is unwise to lead the respondent to provide a certain response to a question you have posed.

This would occur where, you frame a question such as, " Responsible jewellers always use the machine guards, do you use guards?". This would therefore reflect the respondent's perception of what you think is right, rather than what he really does.

5. Personal Questions

It is always wise to avoid the use of personal questions unless they are absolutely necessary. Most people need very little excuse for not completing a questionnaire and if you were to ask a question that they considered too personal, this might be enough for them to reject the entire form.

6. Nature of the response

The questions should always be designed with the eventual analysis in mind. Most questionnaires will be analysed by computer and the questions will have to reflect this. Computers operate best in term of *numbers*.

e.g. 3

Your Living District : _____ {1 : HK Island, 2 : Kowloon, 3 : N.T., 4 : others}

The respondents would be presented with a number of options, leaving him to select the one that matches his circumstance. The final analysis would be carried out by the computer. Not only is the analysis of the responses to this particular question made simple, it is also possible to carry out other forms of analysis. For example, we could ask the computer to select the people who are living in HK Island, and then the computer provides other information of these people.

All of this would be possible by the ***coding of answers***. Apart from this, an additional advantage of ***ticking boxes*** is that people find it easier to do than to make up words and sentences.

Important Points to Note

Whatever survey method you intend to use, some guidelines are useful for you to gather the data :

1. Decide your ***objectives***, your ***target interviewee*** and ***questions***. Remember that you may not have the second chance to meet the interviewee again and ask him the questions you have missed.

2. For statistical purpose, try to use ***closed-ended answers*** or ***multiple choice***. For example, if you want to know the age of the respondent, give them the choices:
{1 : below 20, 2 : 21-30, 3 : 31-40, 4 : 41-50, 5 : above 50}

3. Try to collect personal data ***at the end of your questionnaire***. People are sensitive to give you their personal particulars, if you ask them their personal information at the beginning of the interview (or questionnaire), they will be reluctant to do the whole survey.

4. Use the sentence " *All information is for statistical purpose only. The information you have given will be kept confidentially ... etc.*" to explain your purpose.
(of course, you really have to keep others' personal data confidential)

5. People are lazy to ***think*** and ***write***, you have to find out ***all answers as possible*** (or the most common ones), set them out, and give some space, at the end, for them to fill in answers that are not listed.

e.g. 4 *What is the brand of your TV?*

Panasonic ____, *Sharp* ____, *Sony* ____, *Toshiba* ____, *Hitachi* ____,
Fortress ____, *Philips* ____, *JVC* ____, *Sanyo* ____,
Others (please specify) _____.

6. Remember to show your ***identify*** before you talk to your target interviewees.
(your student ID, your staff ID card, or etc)

Steps in a Sampling Study

1. *What information is required?*

   The answer to this question provides both the motivation and the starting point for the study. If the necessary information is either already available or impossible to obtain, there is no point in carrying out the survey.

2. *What is the relevant population, and is a listing of it available?*

   In many practical studies, the real population of interest may be very difficult or even impossible to define. For example, many publication invite the opinions of their readers on particular questions. It would be very dangerous to generalise their response to a wider population, and this readership is likely to be unrepresentative of the public at large.

3. *How should the sample members be selected?*

   There is no unique way to go about providing the "best" sampling scheme. The appropriate choice will generally depend on the problem at hand and on the resources of the investigator.

4. *How should information be obtained from the sample members?*

   This is an extremely question. First the investigator will want to obtain answers from as high a proportion as possible of the sample members. If the number of not responding is high, it will be difficult to be sure that those who do respond are representative of the population at large.

   For the low responding of postal questionnaires, many researchers attempt to improve the respond rate by including a ***covering letter***, explaining the purpose of the study, and the inclusion of a ***post-paid envelope for returning*** the questionnaire is generally worthwhile, and sometimes ***gift*** may be promised for returning questionnaires.

   The second point is to collect answers that are as accurate and as honest as possible. There is an art in designing questions whether the survey methods used. It is important that the questions are phrased as clearly and unambiguously as possible, so that the interviewee can understand what is being asked.

5. How should sample information be used to make inferences about the population?

6. What conclusions can be drawn about the population?

Sampling and Non-Sampling Error

When a sample is taken from a population, we will not able to know precisely the value of any population parameter, such as mean. Any point estimation will inevitably be in error. The error, that resulting from the fact that information is available on only a subset of all the population members, is called *Sampling Error*.

In practical analyses, some errors, they are ***unconnected*** with the kind of sampling procedure *used*, could just as well arise if a complete census of the population were taken. These are referred to as *Non-Sampling Error*.

Reasons for Non-Sampling Error

1. The population sampled is not the ***relevant*** one.

   If one wants to make inference about a population, it is important to sample ***from that population*** but not ***from some subgroup*** of it.

   For example, a study about the school life of the Hong Kong Institute of Vocational Education students is carried out, and the investigators have a well-designed questionnaire, but they only select their samples within the Institute Library. The result, most likely, would be saying that too much coursework, the course is too difficult, no spare time, and not enough time for exercising. *The result is* ***biased***.

2. Survey subjects may give ***inaccurate*** or ***dishonest*** answers.

   A problem often faced in market research is that some of the questions we would like to ask are so *difficult* to understand, or so *sensitive* that many subjects will either refuse to reply or you will give ***dishonest answers***.

   For example, a plant manager wants to assess the annual losses to the company caused by employee thefts. In principle, a random sample of employee could be selected and sample members asked " *What have you stolen from this plant in the past 12 months?*" That is clearly not a reliable means of obtaining the required information!

3. Another possibility arises through ***non-response***.

   Non-response can induce additional sampling and non-sampling errors. The sampling error arises because the achieved sample size will be *smaller than* the intended will.

Non-sampling error possibly occurs because, in effect, the population being sampled is not the population of interest. If it is suspected that non-response bias is likely to be troublesome, three possibilities are open:

1. the investigator can solicit information through a mechanism known to achieve a relatively high response rate.

2. characteristics of respondents and non-respondents can be compared, in such matters as age, sex, and race, to see if there are obvious differences between two groups.

3. an attempt can be made to contact non-respondents, some of who may well be prepared to provide answers to a few key questions.

There is no general procedure for identifying and analysing non-sampling errors. But they could be important. The main prescription is that the investigators take care in such matters as identifying the relevant population, designing the questionnaire, and dealing with non-response in order to minimise their significance.