CHAPTER FIVE

THE SOURCES AND COLLECTION OF DATA

Introduction

Data is the raw material of problem solving and decision making. Effective data collection is pivotal in the research process. Data provide the basis for beginning to address and investigate research and consultancy problems. However data is differ from information. Ultimately, the researcher is interested in information rather than data. However information stems from raw data and as such data is essential to the problem solving process. The researcher must know not only what data is required, but also the **principal methods, approaches and techniques** for collecting data such that the most appropriate data collection techniques can be used. What then are the principal types of data, how may we classify data and data collection methods, and what are the principle ways and techniques of collecting data as part of a research exercise?

5.1 TYPES AND SOURCES OF DATA

Data comes in a wide variety of shapes and forms. When it comes to types of data we can distinguish between the following major categories:

5.1.1 PRIMARY VERSUS SECONDARY DATA

One of the most fundamental distinctions between types or categories of data is that of primary versus secondary data.

PRIMARY DATA: does not actually exist until and unless it is generated through the research process as part of the consultancy or dissertation or project. As we shall see, primary data is closely related to, and has implications for, the methods and techniques of data collection. For example primary data will often be collected through techniques such as experimentation, interviewing, observation and surveys.

SECONDARY DATA: on the other hand, secondary data is information which already exists in some form or other but which was not primarily collected, at least initially, for the purpose of the consultancy exercise at hand. In fact secondary data is often the start point for data collection in as much as it is the first type of data to be collected. Because of this and the importance and potential value of secondary data, we shall consider this type of data and its collection.

We shall also see that secondary data comprise both internal and external data sources encompassing internal company information such as databases, reports, and company analysis and so on, and external data sources such as published reports, government surveys, competitor information and increasingly, internet and web based sources of information.

5.1. 2 SOURCES OF DATA

There are two sources of data: these are primary sources and secondary sources.

PRIMARY SOURCES: it is source of data that provide firsthand information for the use of immediate purpose and collected from primary sources are new data which had not existed before and for which the researcher received full credit. More accurate and costive.

SECONDARY SOURCES: individual or agencies which provide data originally collected for other purpose by them or others and usually they are published or unpublished materials, records, reports, magazines, market reports, etc. data which is not originated by the investigator himself but which he get from some one's records. Here compared to primary data, which is costly but accurate and more reliable, a secondary data is less costly and less accurate.

5.2 COLLECTION OF PRIMARY DATA

We can obtain primary data either through observation or through direct communication with respondents in one form or another. This means that there are several methods of collecting primary data. The most common are:

i. Observation method
ii. Interview method
iii. Questionnaire method
iv. Schedule method

I, Observation Method: is the most commonly used method especially in studies relating to behavioral sciences (Psychology; Sociology, Anthropology, etc). Under the observation method, the information is collected by way of investigator's own direct observation without asking from the respondent. The main advantage of this method is that subject bias is eliminated, if observation is done accurately. Secondly the information obtained under this method relates to what is currently happening; it is not complicated by either past behavior or future intention or attitudes. Thirdly' this method is independent of respondents' willingness to respond and as such is relatively less demanding of active cooperation on the part of respondents as happens to be the case in the interview or the questionnaire method. This method is particularly suitable in studies which deal with subjects (i.e., respondents) who are not capable of giving verbal reports of their feeling for one reason or the other.

However, observation method has various limitations.

- It is expensive method
- The information provided by this method is very limited
- Sometimes unforeseen factors may interfere with the observational task

While using this method the researcher should keep in mind things like: what should be observed? How the observation should be recorded? Or how the accuracy of the observation can be ensured?

II, Interview Method: this method of collecting data involves presentation of oral- verbal questions and reply in terms of oral-verbal responses. This method can be used through personal interview and if possible through telephone interviews.

Personal interview: it is an oral questioning of respondent either individually or in group (focus group). Personal interview method requires a person known as the interviewer asking questions generally in a face to face contact to the person or persons. The method of collecting information through personal interview is usually carried out in a structured way. As such we call the interviews as structured interviews. Such interview involve the use of a set predetermined question and of highly standardized techniques of recording. Thus the interview in a structured interview follows a rigid procedure laid down, asking questions in a form and order prescribed. As against it, the unstructured interviews are characterized by a flexibility of approach to questioning.

Unstructured interview do not follow a system of pre-determined questions and standardized techniques of recording information. In a non-structured interview, the interviewer is allowed much greater freedom to ask, in case of need supplementary questions or at time he may omit certain questions if the situation so requires. He may change the sequence of questions. He has relatively greater freedom while recording the response to include some aspect and exclude others. But this sort of flexibility results in lack of comparability one interview with another and the analysis of unstructured response become much more difficult and time consuming than that of structured response obtained from structured interviews. Unstructured interview also demand deep knowledge and greater skill on the part of the interviewer.

Advantage of personal interview

a) More information and that too great depth can be obtained.
b) Interview by his/her own skill can overcome the resistance, if any, of the respondents.
c) There is greater flexibility under this method as the opportunity to restructure question is always there, especially in the case of unstructured interviews.
d) The language of the interview can be adapted to the ability or education level of the person interview and as such misinterpretations concerning questions can be avoided.
e) It offers a lot of flexibility in allowing the interview to explain questions, to probe (investigate) more deeply in to the answers provided.
f) It is useful in situations where great depth study is required
g) In face- to- face interview, the interviewer can see and assess the respondents' non-verbal behaviors.

Weakness of personal interview

a) It tend to be relatively expensive and time consuming especially when large and widely spread geographical sample is taken and hence not ideal to large group of informants.
b) There remains the possibility of bias of interviewer as well as that of the respondent.

c) Certain type of respondents (officals, VIPs etc) may not be easily approachable under this method and to that extent the data may prove inadequate.

Precondition and basic tenets of interviewing

For successful implementation of the interview method, interviewers should be carefully selected, trained and briefed. They should be honest, sincere, hardworking, and impartial must possess the technical competence and necessary practical competence. The interviews approach must be friendly, polite, conversational and unbiased.

The Telephone interview: this method of collecting information consists in contacting respondents through telephone itself. It is similar to the personal interview, but uses telephone instead of personal interaction. It is not a very widely used method, but plays important part in industrial surveys, particularly in developed nations.

Advantages:

i). It is more flexible comparison to mailing method
ii). It is faster than other methods, i.e., quick way of obtaining information
iii). It is cheaper than personal interviewing method; here the cost per response is relatively low
iv). No field staff is required etc.

Weakness of Telephone interview:

i). Little time is given to respondents to respond
ii). Surveys are restricted to respondents who have telephone facilities
iii). Extensive geographical coverage may get restricted by cost considerations.
iv). It is not suitable for intensive surveys where comprehensive answers are required to various questions
v). Questions have to be short and to the point; probes are difficult to handle.

III. Collection of Data through Questionnaire Method

This method of data collection is quite popular, particularly of big inquiries. In this method a questionnaires is sent (usually by post) to the persons concerned with a request to answer the question and return the questionnaires. Questionnaires are groups or sequence of questions designed to collect data up to a subject/ sample. This questionnaire is mailed to respondents who are expected to read and understand the questions and write down the reply in the space provided in the questionnaire itself. The respondents have to answer the questions on their own.

The advantages of this method are:

- ✓ There is low cost even when the population is large and widely spread
- ✓ It is free from the bias of the interviewer; answer are in respondents own words
- ✓ Respondents have adequate time to give well thought out answers.

- ✓ Respondents, who are not easily approachable, can also be reached conveniently.
- ✓ Large sample can be made use of it and thus the results can be made more dependable and reliable.

The Disadvantages of this method are:

- ✓ Low rate of return of the study filled in questionnaires
- ✓ It can be used only when respondents are educated and cooperating
- ✓ The control over questionnaire may be lost once it is sent,
- ✓ There is inflexibility because of the difficulty of amending the approach once the questionnaires have been dispatched.
- ✓ There is also the possibility of ambiguous replies or omissions of replies.
- ✓ This method is the slowest of all.

Iv, Schedule Method

This method of data collection is very much like collection of data through questionnaire, with little difference which lies in the fact that schedules (document containing a set of questions) are being filled by the enumerators / research workers/assistant/ data collectors who are especially appointed for the purpose. These enumerators along with schedule go to respondents, put to them the questions from the document in the order the questions are listed and record the replies in the space provided.

5.4 COLLECTION OF SECONDARY DATA

Secondary data means data that are already available i.e., they refer to the data which have already been collected and analyzed by someone else. When the researcher utilizes secondary data, then he has to look into various sources from where he/she can obtain them.

5.4.1 WHAT TO LOOK FOR AS SECONDARY DATA AND WHERE TO FIND IT

5.4.1.1 What is secondary data?

Secondary data is data, which the researcher did not collect for themselves directly from respondents or subjects. This means that secondary data was not collected with the researcher's purpose and objectives in mind. It may have been collected:

- ❖ By other researchers, perhaps in the process of academic studies(could be available in journal articles, or published doctoral thesis or conference proceeding) or
- ❖ In the process of normal operations(e.g an organization's 'grey' material information it publishes internally such as sales figures, information about product launches, company minutes etc, or an individual's personal diary or learning log)

- By institutions, whose job is to collect data (e.g government or regional offices of statistics and information, international bodies whose purpose is information collection e.g academic, media and professional bodies set up for the purpose of collecting information and data directly and from these government or international body).

For many business research studies, especially qualitative ones, it will be difficult to find exactly the kind of data needed, since it is unlikely you would be doing the research were it not for the fact that it hadn't been done before. So most studies will need to design collection methods for primary data. However, there is a vast amount of secondary data out there, much of it surprising accessible over the web, which may save us considerable time or give us a useful benchmark or context in which to set up our research design or a way of triangulating our results.

Where your research relates to a national or international level of operation, it is likely that national and international statistics will form part or all of your study since these studies take time and money to achieve. Much of this kind of data. E.g census data will be available free over the web or free from government offices. Secondary data may be documentary data implies two things- first of all you know how to find it and second you have permission to use it.

5.4.2 THE CONTRIBUTION OF SECONDARY DATA TO BUSINESS RESEARCH

A. Broad Data Groupings Available

Survey secondary data will usually have been analyzed for its original purpose and could be a national periodic compulsory census. A regular e.g annual survey or a one-off survey. You should be aware that this may not be raw data, i.e some filtering and data decisions will have had to be made (e.g coding of non- response, grouping of data etc).

B. Contextual Background

Much research will require an awareness of industry, national or sector context (for example if you are conducting primary research in healthcare organization, it will be useful to set the context for this by comparing national or international healthcare statistics, or you may be reviewing you local area's labour force want to see how this relates to your country's or other country's labour force statistics).

C. Quick and Cheap Data

Secondary data is often cost free and especially if it can be gained over the web or from your local/university library, this should be quick to find. Since you are not in touch with that data respondents, you do not overload them with subsequent research questionnaires and you, as the researcher, cannot affect their responses (which may be the case in primary research).

D. longitudinal and cross-cultural data sources

Much national and international data is collected on periodic basis over time, so allows longitudinal research studies- not normally possible through primary research in view of cost and time constraints. Similarly cross cultural studies can use large survey data, when conducting this primary research is particularly complex.

5.4.3 The disadvantages of secondary data in business research

i. Difference of purpose- Because the original researcher had a different purpose and constraints from your current project, there may be inconsistencies or elements of the research which are not compatible with your own. This could include currency, terminology, samples, market changes, boundary changes, new discoveries or technology since the research was carried out etc.

ii. Cost or access constraint- for secondary data this problem arises mostly when searching for corporate data, either directly from companies or via agencies for market research data.

iii. Aggregation and presentation of data-other researchers working for other research purpose, will often aggregate data in a way, which is not useful for your own research; the presentation of the data will depend on the purpose of the original research too.

IV. Data quality- from this study you will be getting an idea of the attention to detail and careful planning and thought which goes into good quality research. This is one thing which is difficult to check when using secondary data. Instead the best we can usually do is to ensure the credibility and professionalism of the source institution, rather than the data. Be particularly careful when using secondary data from internet sources, where organizations are not known to you, as anyone can put up false data on the web without challenge. This also lead us to be particularly cautions about references sources. A further step in assessing data quality will be to critically evaluate the research methods used to collect the secondary data it is often reasonable to contact the data sources to establish their methods, if you considering using their publicly available data. Government sources usually publish detailed technical background alongside the data to enable you to interpret the data appropriately.

V. Measurement validity- here we cannot expect secondary data to be some kind of 'truth'. The data will reflect the purpose and pre- conceptions, of original researchers.

vi. Data coverage- does the secondary data cover the exact population in which you are interested? Are there any unwanted exclusions or inclusions, which may affect the way you use this data?

vii. Data use- unless you are simply going to quote from the secondary data as background information, you are likely to want to download or enter secondary data to run statistical analysis of some kind. Some data will be in a format ready for this and some may not.