UNIT 4   DATA COLLECTION

Objectives

After studying this unit you should be able to:

- Define the different types of data needed for marketing research m Evaluate the merits and demerits of each type of data

- Discuss the various sources of secondary and primary data

- Explain the basic methods of data collection

- Describe the sources of errors associated with secondary and primary data.

Structure

4.1   INTRODUCTION

In the last unit you have studied the research design alternatives available to the researcher. Whatever may be the type of research design chosen, it is necessary to collect accurate and reliable data in order to achieve the research objectives. In this unit, we shall discuss the different types of data used for marketing research problems and their role in the research process. We shall also learn about the major sources of secondary and primary data as well as the basic methods of collecting data. Whether data is collected by communication methods or through observation, errors of different types tend to be associated with different type and tool of data collection. We shall also discuss the important sources of error associated with data collection.

4.2   DATA AND THE RESEARCH PROCESS

Once the research objectives have been defined the next logical step is to identify the nature and type of information needed to achieve these objectives. Fig. 1 below represents the initial steps in the research process with emphasis upon the types of data sources and options available to the researcher, The research objectives define the type and extent of information needed to achieve the research objectives - the data needs are further clarified by the type of research design chosen as well as by the nature of research e.g. whether the problem at hand is that of exploratory or preliminary research or causal and conclusive research and so on.

All data sources available to the researcher can be classified as either secondary or primary. Secondary data are already published data collected for purposes other than the research problem at hand. Primary data are those generated specifically for the purposes ' of the research problem in question.

5

Generally, the first issue in the data collection process is to determine whether the information needed for the research problem has already been generated/collected. Only after the assessment of the data already available, additional data needs should be specified.

Figure 1 : Data and the Research Process
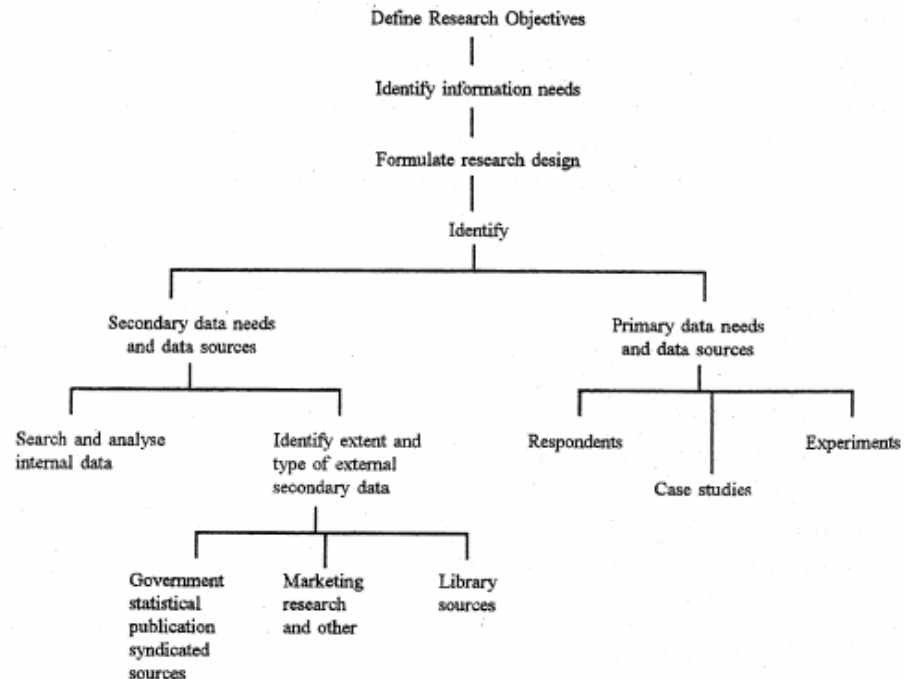
Types of Data

As you have already noted there are two general types of data - Primary and Secondary classified on the basis of purpose of collection or source. Primary data are those that are collected specifically for the research situation at hand. Conversely secondary data are already published data collected for purposes other than the specific research needs at hand. On the basis of location of sources, secondary data may again be classified as internal or external data.. The data originating within or available within the organisation as a byproduct of the MIS or the routine reporting system is called internal data of any given marketing research problem, initial data collected for purposes other than that specific problem could be termed internal secondary data.

Secondary data generated outside the organisation is termed external secondary data and can be collected from a multitude of sources like government publication, trade association publications, official reports, journals and periodicals and publication of marketing research agencies. Secondary data can &ISO be thought from research agencies though this is a fairly expensive preposition.

Secondary data may also be classified on the basis of whether it is periodic data or ' adhoc data. Periodic data characterises most statistics collected over fixed periods of time like the census data, data from statistical abstracts of trade and other sector's, price indices and so on. Adhoc data, on the other hand refers to the data obtained from a certain project report. Such data necessitates an external search to find the source from, which the data can be availed of

As the figure 1 shows the major sources of primary data are respondents, analogous case studies and experiments. The secondary data on the other hand can be obtained from official government publication, marketing research agencies and syndicated sources (which sell the data) and the library sources. Published individual research projects have not been dealt with separately as they are included the library source. Let us now examine the data options and their sources in detail.

Activity 1

Talk to the person responsible for marketing research function in your own organisation/ any other organisation that you are familiar with. Try to find out with respect to (a) a product introduction decision, and (b) a new sales incentive scheme introduction decision,what were the type of data utilized by the organisation, for these decision. What were the major sources of data and 'the methods used to collect the data.

a)   Product introduction decision

        Type of data

        ....
        ....
        ....

b)   Major data sources

        ....
        ....
        ....

c)   Methods used

        ....
        ....
        ....

b)   Sales incentive scheme introduction

        ....
        ....
        ....

4.3   SECONDARY DATA- NEED AND USAGE

Before discussing the sources of secondary data, let us first try to envisage what type of research objectives and information needs may necessitate use of secondary data. Some examples could be

a)   You want to estimate the total market potential of thermocol packaging cases in the country.

b)   Being a manufacturer of television sets, you want to develop a national method for establishing sales quota statewise.

c)   As a person making automobile batteries, you want to estimate the potential market for battery replacements as well as develop methods for determination of countrywise, statewise and districtwise sales quotas.

d)   You would like to predict the potential sales of paper and allied products to wholesaler and retailer, for the year 1991.

e)   You are required to select a city in Western Maharashtra to locate another distribution centre for the spare parts of your two wheeler products.

f)   You are interested in allocating your sales promotion budget in proportion to the potential markets, statewise in the northern pan of India.

The type of secondary data required to Meet some of the above research objectives would be population data, estimates of total television sales, statewise, number of households with televisions per capita income estimates, distribution of income statistics. Number of registrations, number of households without cars, value of box shipments by end use and so on. You have read about the various kind of research designs in the preceding unit. With the type of research design chosen need and use of secondary data differ.

In monitoring research, the researcher is trying to keep a track of a programme or phenomenon with respect to the variables under study (it may be the progress of pan

under study or a specific progress) and is keen to sense changes in the environment which could help or hamper the product programme under study. Most of the data for this type of monitoring would come from secondary sources and quite a substantial part of it would be internally generated. For example accounting data on sales, expenses and margins earned could provide significant clues to the progress of the product in the given market. Sales force reports could generate observations about customers and market trends that could be used to sound early warning bells to the executives.

Preliminary research also makes extensive use of secondary data, though data can be gathered afresh also, in small or informal primary studies. In this case secondary data is used to estimate parameters of a problem or a trend and the changes in its environment. Preliminary research, sinᵍe it explores nature and extent of the problem, does not generally merit the expense & delay of primary data collection. In this preliminary phase, usually descriptive data, both from internal and external sources can be used to understand and to define the nature of a research problem.

If you are undertaking exploratory research (for example if you are trying to explore the various possible combination of the communication mix which could yield the optimum mix) you will be able to make a relatively limited use of secondary data because this type of research is concerned with finding specific alternatives that should considered while arriving at a decision. Of course past experience of the company, its records or internal data could provide guidelines for action as could external secondary data, especially published material on consumer perception and reaction to different tools of the communication mix, media habits and so on.

Conclusive research, concerned as it is with establishing dependence or independence of the variables under study is mainly based on primary data gathered to suit the objectives of the research problem under study. For example, if you are interested in finding out whether a given sales promotion programme has a significant impact on the sales of 1neur new range of moulded luggage then you would have to rely primarily on fresh data generated specifically to establish the cause and effect relationship between the sales promotion programme and the luggage sold.

However, secondary data may be used for comparison purposes even in conclusive research. You might like to compare the industry trend in respect of promotional expenditures and sales with your own company position, or compare the position of the company statewise with that of its select competitors. Also secondary data is useful in framing a primary data collection programme, as well as in framing the sample or determining the sizes of the sub-samples.

Government Agencies and Official Publications: The largest single source of secondary data in macro terms are the' publications by the Union government. Marketing researches have relied on this source of data for estimating market potential and sales forecasts, determining distribution penetration and location of intermediate and final outlets, as well as for defining sales territories and routing schedules. Estimates of income and expenditure patterns become good starting points for estimation of paying capacity for different products and services. Estimates of literacy levels become effective inputs in planning promotional strategies. Over a period of time, the variety and depth of government data has increased manifold and it relevance to marketing research function has consequently enhanced. The following sources have special relevance for marketing research.

1) Population Statistics

Population statistics in India have been collected, every ten years since 1871-72. Besides forming a basis for most macro planning estimates, census data in respect of the marketing research function is the most important source of national population statistics and its basic characteristics especially demographic and economic. It provides factual bases for estimation of consumer demand for various goods and services, by furnishing

data on size of population and its distribution by age, sex, occupation and income levels. Census also furnishes a sampling frame for reliable sample designs. Estimates of population density, distribution, literacy levels, consumption patterns all become important inputs in distribution decisions, communication -decisions and overall marketing policy.

2) Statistical Abstract of India

Brought out annually by The Central Statistical Organisation, this publication contains the statistics of various section of the Indian economy for the preceding five years. Statewise, statistics for these economic variables are also furnished periodically.

3) Estimation of National Product, 'Savings and Capital formation: (White paper on National Income)

This is also an annual publication of the Central Statistical Organisation. It publishes annual estimates, of national income, savings and consumption, capital formation and expenditure as well as national and public sector accounts.

4) Monthly Statistics on the Production of Selected Industries

To bridge the gap between the census taken and data published, the C.S.O.publishes monthly statistics relating to production, installed capacity and stock positions in selected industries. More than 90 industries are included in the compilation of chosen statistics.

5) Basic Statistics Relating to Indian Economy

Published annually by the statistics and survey division of Planning Commission this publication contains basic indicators on various aspects of economy for the past few years based on time series.

6) India, Pocketbook of Economic Information

It is an annual publication of the Ministry of Finance. Giving particular emphasis on estimates of foreign financial and international comparisons, it includes statistics on the various aspects of the national economy.

Other important publications include India, a reference manual, Agricultural Situation in India, RBI Bulletin, Economic Survey, Bulletin of Food Statistics, Commercial Crop Statistics and Indian labour statistics (which includes transport and communication statistics, data on employment in shop and commercial establishments etc.).

7) Trade Statistics

The Department of Commercial Intelligence and Statistics compiles monthly statistics of commercial intelligence and statistics, publishes data on import and export of goods in terms of their quantity and value, classified as received from or sent to centres of consignment. In addition, this publication provides information on value of foreign trade, balance of trade, foreign trade with each country and currency area, foreign trade in groups of commodities with each country and currency area, foreign trade with selected countries etc. It therefore, furnishes a good source of data for assessment of international market trends and potential. The Reserve Bank of India also publishes statistics on imports and exports based on exchange control data.

In addition, official reports both published and unpublished from organisations like National Sample Survey Organisation, Directorate of Economics and Statistics, Labour Bureau etc, furnish data on the different aspects of our economy. Agencies like Cottage Industry Board, National Small Scale Industries Corporation, Tea Board, Handicraft Board etc, provide specific information in respect of their own product categories.

Other Sources of Secondary Data

Library: The library sources of marketing data include the whole gamut of publicly circulated material i.e, government documents and reports, books, periodicals, journals, individuals research project reports and trade association publications. As the categories

noted above are intended for public circulation , the library represents an easy , economic and efficient source of secondary data.

Some of the item inventioned above are the results of extensive original research, some a summary of the research of others and some presenting statements of opinions. In comparison to the official statistics described before, these items tend to include more discussion and less statistics which in turn may help you decide upon the suitability of a particular type of secondary data for your own purposes.

Research Agencies and Data Services: The growing demand for marketing data has brought forth several organisation which collect and sell standardized data. Also called syndicated sources, these agencies include the marketing research agencies as well as the data services which in addition to providing standardized data also undertake specific data collection research projects. In India, advertising agencies have also emerged as 'good sources of data on readership media habits, attitudinal research and other communication related areas.

The data that can be obtained from these syndicated agencies includes consumer data, retail data, wholesale data, industrial data, advertising evaluation data and, media and audience data.

Leading marketing research agencies like MARG and ORG regularly survey consumers attitude and opinions regarding consumption behaviour and a variety of contemporary issues relevant to marketing.

Activity 2

A) Study the demand estimation process for an existing product in your organisation or any other organisation that you are familiar with

   a) What were the sources of secondary data utilized?

   b) What specific use of secondary data was made in the study?

B) Repeat the exercise for the demand estimation for a new product that is yet to be launched.

   ……………………………………………………………………………………
   ……………………………………………………………………………………
   ……………………………………………………………………………………

4.5 ADVANTAGES AND LIMITATIONS OF SECONDARY DATA

Advantages of secondary data: The major advantage of secondary data is the economy of resources that it offers both in terms of money and time. Primary research involves selecting the sample frame, determining the sample size, choosing the tools of data collection, getting data collection instrument printed in case of field studies as well as editing tabulating and analysing the results, which turns out to be expensive and time consuming. Secondary data on the other hand can be collected by researcher from published or compiled research at very little cost and usually very speedily.

Another important advantage which characterises some secondary data services is that they provide access to information which would not ordinarily be obtainable by an individual organisation. The census of wholesale and retail establishments for example can require them to furnish details of sales, expenses and profit information which would be inaccessible to an inch dual researcher. Also, as information like this is collected in the usual course of events, the data is less prone to be biased which may be the case when the data is collected with a specific purpose in mind.

Limitations of secondary data: While economy and access to relatively unbiased information are important advantages, effective use of secondary data requires the overcoming of two major difficulties. These are the difficulties associated with fitness of data and accuracy of the data.

The data fitness problem: The problem here refers to the suitability or fitness of secondary data to the research problem at hand. As secondary data have been complied for other purposes, rarely are they completely pertinent to the information needs of the problem at hand. Even when they are relevant, the degree of fitness to a given research problem may be hampered by (a) Units of measurement (b) Class definition and (c) Recency.

Among the most common limitations of secondary data is the variation found in Me units of measurement used for a given variable. For example if you want to have a reliable estimate of the disposable income in order to base your market projections on it, you may discover that depending upon the source, income flay have been measured by individual, household or tax returns. In case you are required to use more than one of these sources the difference in the units of measurement would allow the information to be used only as rough approximation. Another common feature of secondary data is the variation in definition of classes that would be found in different data sources or even in different project reports of the same data source. For example definition of a class of literate population may have been defined as those who can read or write in any language, those who have had at least basic schooling, or those who have attained a certain educational certification, in different reports published by the same source or different sources.

Recency of the secondary data also becomes an important determinant of the use of secondary data. Marketing decisions can hardly be based on information that is dated and therefore no longer relevant. The use of census data which is an invaluable source of voluminous information is severely limited by the fact that not only do we have a decennial census, the reports of the census become available only a year or two after the census. Census information collected in 1981 would lose its relevance to most marketing decisions beyond 1985 or so, even though the data contains some of the most vital marketing data on income, consumption, expenditure savings and investment.

The accuracy problem: Once an appropriate source of secondary data has been identified, it is important to evaluate the accuracy of the data in relation to the research problem at hand. The following steps would help in evaluating the accuracy of the secondary data.

Identify the data source: Secondary data may be obtained from original source or from a secondary source which in turn has obtained them from somewhere else. As far as possible data must be collected from the original sources as you may then be able to evaluate the methodology of research that the original researcher has used. When secondary sources of secondary data are used, not only are details of methodology of original source non-available, errors originating at the secondary source may affect the accuracy of the original data. Moreover, secondary sources usually fail to keep up with the updated versions and adjustments reported by the original sources from time to time as more information becomes available to them. Secondary sources of data must therefore be used with caution and after careful scrutiny.

Examine the purpose for which data was published: Sources which publish to promote the interests of a particular group or for political commercial or social reasons must be treated with caution. Similarly data published to promote sales or to carry on a particular propaganda or to promote views of a particular interest group has limited use unless it is carefully interpreted in the light of purpose of publication. Data published anonymously or by an organisation which is on the defensive or under conditions which suggest a controversy or in a form which reveals a strained attempt at frankness or to controvert inferences from other data are generally suspect."

(I.E. Newmers and John H. Nyers 'Business Research: Text and Cases', New York, McGraw Hill, page-42.)

11

Assess the source in terms of the quality of data expected: In addition to identifying the organisation which collected the data, the ability of the organisation to procure the desired information must be assessed. For example, the department of internal revenue is much more likely to generate reliable information about income and expenditures simply because it has the authority to glean that kind of information.

Evaluation of the methodology of research used by the source in terms of research design, sampling plan data collection procedures and analysis procedures also furnishes important clues to the accuracy of secondary data. If the primary source is not willing to divulge the methodology of research one should be hesitant about using the data as unwillingness to share the details of methodology usually suggests inadequacies in the procedures used.,

4.6 SOURCES OF PRIMARY DATA

As noted earlier, though nearly all marketing research problems make some use of secondary data, both exploratory and conclusive research situations necessitate using a high proportion of primary data. The major sources of primary data include respondents, analogous case situations and research experiments.

Respondents: Respondents represent by far the most important source of primary marketing data. Marketing decisions are characterised by the fact that they always involve in one way or the other, prediction of the behaviour of the market participants- be it consumer, industrial users, marketing intermediaries or competitors. Decisions as diverse as product introduction, price or channel modification, determination of the advertising budget or reallocation of sales territories would require forecasting of the behaviour of one or more of the above groups. The study of respondents therefore characterises most marketing research situations.

The type of information that may be collected from respondents may include data on past behaviour, intentions of likely behaviour, extent of knowledge, attitudes and opinion and socio-economic characteristics and lifestyle data.

Past behaviour is frequently used as a predictor of future behaviour, acting on the premise that there is a relatively stable relationship between past and future behaviour patterns. Information or past behaviour of respondents is sought on past consumption habits, exposure and awareness of communication tools, experience of competitor's vis-a vis the company's products or services etc., in order to get an idea of the consumption of the product or service.

Intention of likely action are self prediction of planned action to be taken sometime in future. They represent a very frequently sought information from the respondent. However, as intention only represent concurrently valid statements, actual behaviour ma) get modified by changes over which the respondent has little or no control; for example- price changes or entry of a new attractive competitive brand. The utility of intentions data is far higher as a predictive tool, when intentions, given the conditions of purchase can be assigned probabilities.

Buyer decisions to purchase a product or services depend to a certain extent on what they know about the product or service. Extent of knowledge or awareness is therefore another type of information that is sought from respondents. As a marketer of colour television sets you would like to know what do the buyers know about your product's attributes like picture quality and sound fidelity, extra features available, price etc. vis-a-vis competitors products, if you think that these are the criteria which have important bearing or purchase decisions.

Attitudes and opinions represents predisposed mental state to act in a certain way. For' example, the statement by a respondent that he prefers fresh ground coffee to the instant" blends, gives a clue to his likely behaviour when faced with a decision to buy coffee. Attitude research in marketing has found the greatest use in product designs and

advertising. Both qualitative and quantitative techniques are used to collect information on attitudes and opinions. They have been discussed under the section *Basic Method of data collection'.

Socio-economic characteristics

Whenever in a consumer research situation there is a basis for the belief that some socio-economic characteristics like income, education occupational status etc. are associated with the purchase of a product or service, information is sought on these characteristics. Study on these types of variables may help defining the target market more precisely and arriving at more accurate pricing and promotion decisions.

Life styles are increasingly being used to segment markets reposition products and targeting positioning strategies wore closely to the consumer profile. Life style profiles are developed though systematic study of values attitudes, opinions and interests, as well as by income, all of which can be elicited from the respondents through communication methods or a combination of communication and observation methods. You are already familiar with the information needed and methodology used to develop life style profiles through your exposure to Unit 2 of this course.

Primary data from Analogous situations - Case Study: Evolved from the behavioural sciences, case study or case history is in extensive use in marketing research today. . Using situations analogous to or relevant to the problem situation, an in depth investigation is carried out to thoroughly study the case situation. The emphasis of the study is on identifying key variables, defining the nature of their relationships and possibly defining the problem/opportunity in order to suggest alternative courses of action for the decision situation. The source is typically relevant in situations where multiple variables interact to produce the problem or the opportunity. Examples would be research problems involving, a) study of changes in sales performance with the entry of a new competitor, b) contrasting performance levels in different markets, c) transitional stages like study of sales territories where the company is altering its distribution channels from indirect to direct sales.

Simulation: A marketing simulation is an incomplete representation of -the marketing system or some aspect of the system. It is a relatively new data source and is largely computer based. It involves the study of dynamics of the marketing system by manipulating independent variables like marketing mix or situational factors and observing their influence on the dependent variables. It would therefore need characteristics of the phenomenon and the relationships between variables to be represented as data inputs.

The researcher creating a marketing simulation is required to conceptualized and record the structural component and assign probabilities to represent the behaviour of the components. Depending upon the phenomenon being studied these components could be buyers, retailers, households etc. while the variables which affect how these components would behave could be price charges, advertising expenditure levels, sales promotion programmes, quality of the product and so on.

The constants in respect to phenomenon under study are represented as parameters of the study which can be manipulated by the researcher in order to experiment and explore alternative marketing strategies. The behaviour of the components results in numerical output. An example could be, suppose you want to study the consumer response to varying advertising expenditures. in terms of sales. The components here would be consumers, the variable is a expenditure and sales is the dependent variable. The researcher would assign probabilities of consumers reacting to a particular level of advertising expenditure in terms of buying the product and by varying the different expenditure levels, would get the output in terms of sales expected at these levels. The constants could be the size of the advertising budget.

Simulation is a complicated process and its success depends how well has the actual situation (reality) been represented in the simulation (abstraction). It requires the skill of a specialist and is therefore limited in application as far as the routine marketing research situations are concerned.

13

Experimentation

Experimentation also represents a fairly rich source of primary data and is used mostly to study cause and effect relationships among marketing variables. Marketing experiments are conducted almost like other scientific experiments. One or more independent variables are consciously manipulated or controlled and their impact on indendent variables is studied. All the same effort is made to control those variables which may hamper the ability to interpret valid causality. Various experimental designs have already been discussed in the preceding unit under the subtitle of experimental research design.

Advantages of Primary data

Primary sources usually provide more detailed information than the secondary sources. This is partly because methods of data collection and the tools used can be tailored more precisely to the informational needs of the researcher. This also contributes to the flexibility of analysis for the research purpose at hand.

Terms and units can be more precisely defined and the researcher can choose the appropriate unit for this purpose with a much greater degree of fit since he has greater degree of flexibility in choosing the appropriate unit. In case of secondary data, as the data have been collected for some other purpose, the researcher may not have much of control over the choice of units used in the secondary data source.

The user in case of primary data can judge the degree of confidence that he may place on the data because he has an accurate idea of the tools and methodology used and their limitations.

In addition to the above general advantages, there are other merits which characterise the different methods of primary data collection. We shall study them under the section concerned with tools of data collection. The major limitations of primary data relates itself to costs in terms of time and money involved, availability of skilled investigators or interviewers and the errors that may creep in various stages i.e. sample selection, sample size selection, tools chosen etc. As the errors related to primary data collection require greater attention they have been explained separately in Section 4.8.

4.7   BASIC METHODS OF DATA COLLECTION

When you, in your everyday life seek information on something from people, the only ways of getting it are either asking for it or observing the phenomenon on which you want information. Formal research has merely formalised and systematized these two alternative methods of data collection i.e. communication and observation. Within each category you will find several tools that can be used to affect communication and/or observation. The fact that distinguishes communication methods is that there is a dialogue with the respondent either spoken or written, while the observation methods involve keeping a track of or monitoring the behaviour of the subject matter of the research, without interfering with the same.

Communication methods

Interviews: Interviews in marketing research are by far the most common method of data collection. Interviews may be:

a)	Structured and Direct	Involving the use of a structured formal questionnaire as well as an interviewer (e.g. surveys using questionnaires).
b)	Unstructured and Direct	Not involving the use of a predecided questionnaire, only the interviewer (e.g. personal interviews)

c)	Structured and indirect	Where the responded to a non-personal ambigous situation which is later interpreted (e.g. projective techniques like word association)
d)	Unstructured and indirect	where the respondent is asked to respond to non-personal ambigous situation but where the interviewer has considerable degree of freedom in modifying/altering the situation.

a) Structured and Direct (Surveys using Questionnaires)

Among the communication methods in use today, surveys involving structured questionnaires are the most extensively used. If you want to know the consumption patterns of a particular product category, reasons for brand choice, the relative influence of different media on a person's decisions, the natural course would be to ask the people themselves. In order to get comparable information from repondents, formal nondisguised questions are predesigned in the form of a questionnaire. Administrating the questionnaire may involve participation of an interviewer, who is instructed to ask the questions in the order given on the form and to ask only those questions. The questionnaire method, with or without the participation of an interviewer is such a prevalent and versatile method of data collection that it was felt that various stages of questionnaire planning and execution warranted a detailed explanation. Unit 15 of this block is therefore devoted entirely to questionnaire planning and execution. You will get a detailed exposure to this tool of data collection in that unit.

The basic advantages of using questionnaires are versatility and economy. The questionnaire method is versatile enough to be tailored to the needs of most marketing research situations. In addition variables like knowledge, opinions, intentions, motivations and personal habits which do not lend themselves to observation, can only be elicited through questioning. It is also the only method to get information on past events for which records have not been maintained.

Relatively speaking, questioning turns out to be speedier and less costly then observation. Interviewers using questionnaires have a higher degree of control over information gathering activities than do observers, who have to wait for the respondent to perform the action under study. Moreover the lag time between one interview and another is controllable by the interviewer while the wait time between two successive observations cannot be controlled by the observer.

The questionnaire method is however subject to certain limitations also. Important among them are inability of the respondent to furnish information. Even though they may be willing to share information 'about themselves, many respondents are actually unable to give accurate, information to the questions asked. For example, it might be difficult for you, at a point of time to give the exact reason of choosing a brand of soap over another unless you have carefully analysed these reasons beforehand. Therefore, on questions of buyer motivations quite often inaccurate information is given which interferes with the reliability of data collected.

Unwillingness of respondent to provide information

The questionnaire aided interviewer is usually an imposition on the time of the respondent, who is required to answer questions on his personal habits, values, socio-economic characteristics and so on to an unknown interviewer. In some cases, respondents refuse to spend the time needed to answer the questions and in quite a few cases, refuse to answer question in relation to their income or on very personal subjects. These gaps in the information collection may affect the overall research process.

Influence of the questionning process

As the situation in which *the* questionnaire asks respondents to relate, routine actions is an artificial one, they may furnish information which is at variance with the truth.

Quite often, respondentary to give answers that they think will please the interviewer. At times, if the true answer to a question shows them in a bad light or is damaging to

the ego, respondents tend to give 'doctored' or manufactured responses. Also some respondents try at the whole process very lightly and try to amuse shock or astonish the interviewer/reader through their responses.

b) Unstructured direct interviews

This method does not involve the use of a formally structured questionnaire. The interviewer is only given general instructions on the type of information sought. He has therefore a considerable degree of freedom in choosing the questions, as well as the wording and the order which may seem most appropriate for a given interview. Generally used in exploratory research studies these interviews are useful in obtaining a clearer understanding of the problem and determining the areas to be investigated.

This technique as also applied for obtaining information on motives for respondents action. When so used, it usually takes the form of a depth interview. Exploring the motive behind a particular purchase the interviewer may continue asking probing questions like "could you explain what you meant by this statement ? Why do you feel like that? Do you have any other reason"? Until he has obtained all the information he needs. The constraints in this case are requirements of the problem, time limitations and the willingness as well as ability of respondents to put their motives into words.

The advantage of unstructured direct interview is that since it is free of the restrictions imposed by formalistic questions, the interview may be conducted in a conversational casual mode, which is more conducive to information gathering. Moreover the vocabulary can be adjusted to ensure rapport and understanding. The flexibility and informality of the direct unstructured interview often provides access to information not usually obtained in structured questionnaire. The possibility that respondents will knowingly or unknowingly furnish wrong responses is however still there.

The success to the unstructured interview depends upon the skill of the interviewer in formulating and asking questions. The use of highly skilled and competent interviewers adds to the cost per interview. Another factor contributing to costs is - that unstructured interviews are much longer than structured ones, and the varied order and form of questions makes editing and tabulating much more complicated.

c) Structured and Unstructured indirect interviews

In order to overcome the nonresponse or inaccurate response problem associated with direct questionnaires and unstructured interviews, some techniques have been developed to elicit information in an indirect fashion. Applying the projective techniques developed in the field of psychology, the respondents are given to describe a non personal ambigous situation. The premise is that the description by the respondents would involve a projection of their own personalities, values, motives,, needs and desires. The most common 'among these techniques are word association tests, sentence completion tests and thermatic appreception (interpretation of pictorial representation). These techniques have been mostly used in the study of buying motivations and consumption patterns of consumer products which are very similar in performance, price and quality, so that the consumer has little basis of differentiating them except brand and company image. Examples are soaps and detergents, food products, cigarettes etc.

Since predesigned set" of words, statements or pictoral representation are used to elicit respondents descriptions, all indirect interviews involve some structuring. The degree however varies from highly structured ones to relatively unstructured interviews where the interviewer has a high degree of latitude in questioning the respondent to get fuller responses. Some of the commonly used indirect interview techniques are described below to help you to understand these tools of data collection.

Word association tests

The word association tests involves the presentation of a series of stimulus words to a respondent who is asked to quickly supply the word that first conies to his mind after hearing the stimulus word. Presumably the respondent would give the word that he most closely associates- with the stimulus used. For example given the name of a particular brand of instant coffee you may respond in terms of your feelings regarding its flavor,

its evaluation vis-a-vis ground coffee, its costliness, its convenience depending upon the dominant association that you have for this product. It helps in identifying respondent perceptions about products and services, new packaging designs, advertising compaigns and so on.

Sentence completion tests

Developed as an extension of the word association test the sentence completion test involves providing an opening phrase of the sentences to the respondent who is then asked to complete it quickly. Trying to identify the motivation of car buyers a marketing research study used the sentence completion test. The phrases provided were:

People who drive a Maruti ……………………………………………………

Professionals usually drive ……………………………………………………..

Most of the new cars ……………………………………………………..

When I drive fast ……………………………………………………

Family men prefer ……………………………………………………

The responses given .by men and women were found to differ. Fuel economy, maneverability and stylish looks were interpreted to be the more important motives for men while safety was for man important to women.

Thematic appreciation test (Picture interpretation tests)

The test involves showing one or more pictures/cartoons to the respondent and asking them to describe the situation or to assume 'the role of one of the people shown in the picture/cartoon. Usually the picture shows one or more persons in an ambiguous situation and it is expected that the respondent will decipher the situation in accordance with his own personality.

In marketing research situations, use of cartoon involving people and the product under study is frequently made. The descriptions provided by the respondents are then interpreted by skilled interpreters to understand their price and quality association with the product, its valuation in the mind vis-a-vis competing products and soon.

Another variation of this projective technique is the story completion test when the respondents are given the beginning of ¸a situational narrative and are asked to complete it.

Focus groups

Focus groups are the most widely used type of indirect interviews. Here a group of people jointly participate in an unstructured indirect interview. A purposively selected group of 8-12 ¸people who have a common background on similar user experience relating to the problem under study is brought together in a group discussion exercise. The interviewer, who plays the role of a moderator attempts to focus the discussion on the problem areas in a nondirected, relaxed way. The purpose of this exercise is to provide interaction and involvement among the group members doing the interview, so that a spontaneous discussion with expressions of attitudes opinions values user experience and intention of further product use takes place.

Focus groups characterize preliminary and exploratory studies rather than casual research and are used to generate information to facilitate the formulation of hypothesis or settling of premises for further research. it finds application in studies involving examination of new product concept, finalizing creative concepts ' for 'advertising generation of information for modification of existing products as well as determining improved ways of distributing products.

Media used in direct and indirect interviews

The interviews may use the personal medium, involving the use of an interviewer or use mail as a medium. Mail surveys are very common and offer the advantage of a lower costs. However, the incidence of nonresponse in our country is so high that it has severely limited the use of this economical medium.

17

Another medium though not used frequently in India, is the telephone which lends itself to both structured and unstructured interviews. In countries where telephone interviews are exclusively used they are found to be a. speedy and inexpensive medium for conducting interviews.

Let us now look at the other category of data collection methods.

Observation method

Observation methods represent another class of data collection methods where without interacting with the respondents or the subjects under study, observations are made and recorded. Sometimes it is less costly and more accurate, if instead of asking the respondents their behaviour is carefully observed. Though past behaviour cannot be observed, the result of such behaviour can be. For example, instead of asking what brand of television do you own '? Is it black or white ? It is remote operated '? It would be simpler to look at the set and observe its usage.

Observation may be used as the only means of data collection or in combination with other means. It is quite commonly used in," exploratory and monitoring researcher designs. For example, if the newly opened Green Supermarket in Delhi wanted to know if its prices of commodities were competitive, the only way to get the information is to regularly monitor, through observation, the prices of these commodities in other supermarkets and comparable stores.

In certain circumstances, though alternative modes of data collection are available, observation may actually be the preferable method of information collection, because of cost or accuracy consideration. In the section dealing with limitation of survey we have discussed inability or unwillingness of respondents to give information accurately. Observation may become the more suitable mode of data collection in such a situation. Studies on brand buying behaviour have shown that purchase/possession of the reputed or prestige brand is overstated while availing of discounts is understated by the respondents. In such cases observations of buying patterns may give better results.

Observation may be carried out manually or mechanically. In countries like UK, USA and Canada devices like the hidden motion picture cameras are used to observe consumer behaviour in the natural situation. The audiometer is another gadget, (described as the electromechanical equivalent of the respondent diary) used for recording day to day behaviour. It is installed in the televisions of a. selected panel of people and automatically records the times the set in turned on or as well as the stations tuned in. The recordings collected are used to get total average and share of audience for the different channels and commercials. These devices permit a complete and accurate recording: In India, some jewellery stores make use of short circuit cameras.

Manually, the observation method involves the use of trained observer who make and record direct observations of respondent behaviour. Information pertaining to the following questions may be collected by direct observations:

a) Who actually is the buyer for the product in question ?

b) To what extent do they seem to have predetermined idea of buying this product?

c) Did they appear to be influenced by the person accompanying them?

d) How many of them checked the price and/or compared it with that of competitors?

e) How many of them studied the package before buying?

Direct observation discloses what was purchased rather than why it was purchased. As motives cannot be observed they must be inferred and these inferences can seldom be substantiated. Another limitation of manual observation is the observers bias - observer may selectively record actions that 'seem meaningful to them and ignore others. This has implications for the objectivity of the research.

4.8 SOURCES OF ERROR IN PRIMARY DATA COLLECTION

You are familiar with the sources or error associated with secondary data. The primary data collection methods are also subject to three important types of errors. These are sampling error, non-response error and response error.

Sampling error as the name implies is inherent in the procedure of sample chosen and results in the sample becoming non-representative of the population. For example, in order to study the patterns of cigarette consumption among Indian males if you chose a. sample of college student in a metropolitan city, this sample would not representative of the population of males in India. The study that you conduct on this sample, no matter which tool of data collection you use, would not be valid because it suffers from sampling error. The range of sampling error however can be controlled by changing the characteristics of sample drawn. Moreover, the extent of the sampling error can be measured if we take a. probability sample. More about sampling error has been discussed in the next unit on sampling.

A non-response error occurs when a. unit (unit here may be an individual, a family or an establishment) included in the sample, cannot or has not been reached. For example, in a sample of housewives from a particular city area, if a number of them happen to be away everytime the interviewer chooses to come, non-response is likely to occur. Incidence of non-response error as already noted is very high in mail interviews as respondents simply ignore the questionnaire received by them..

In most direct structured interviews i.e. surveys involving use of questionnaires, non-response bias is a sizeable error. It may affect completeness as well as objectivity in data collection as families who cannot be reached after certain attempts during the day may be significantly different from those which can be easily contacted. The non-response error is a serious matter because the direction of the error is generally unknown. One can assume that the non response respondents would each have responded in a given way and therefore can determine the maximum error due to non-response but it is difficult to measure the magnitude of the error. One simple way of minimising this error would be to fix up an appointment before the interview but specially in a country like ours where a large number of respondents do not have access to the telephone, this may not be very practicable.

Response error occurs when the value of the reported variable differs from the actual value of that variable. We world here include errors of both communication and observation. We have already talked about two reasons for response error i.e. inability of the respondent to give accurate information or their unwillingness to give accurate information because of time factor, prestige factor and invasion of primary factor. Let us now discuss the sources of response error related to the investigator and the tools used by him.

Inaccurate information due to the investigator

The most common cause of this type of inaccuracy is cheating by the interviewer. There are a number of way in which interviewers deliberately obtain inaccurate information and supply it. If the questionnaire, happens to contain a question that the investigator finds embarrassing to ask, he may decide to supply his own answer or supply an inference on what the respondents answer would have been. In extreme cases reports of interviewees' without ever having contacted the interviews have been discovered to be submitted. Another in between situation that is found to exist is that interviewers get their own friends and associates to fill in the questionnaire or respond to a direct interview amid list the responses in the names of the people listed in the sample, thus vitating the entire sampling exercise.

Experienced marketing research agencies feel that like other petty forms of cheating, interviewer cheating can only be controlled to lower its incidence, it can never be eliminated completely. Care in selection, training and supervision of interviewers can

and does help in controlling the incidence of cheating. In addition, certain control procedures like cross checking of small samples of respondents and use to cheater question which disclose the fabricated answers with a fairly high success rate are employed to minimise the incidence of interviewer generated inaccuracy.

Ambiguity

Ambiguity which is defined as errors made in interpreting behaviour or words spoken or written is source of error which occurs in both, communication and observation methods of data collection. All languages are capable of being ambiguous as the person transmitting information and the person receiving them are two different people and the interpretation of the question/behaviour may differ from one person to another.

4.9 SUMMARY

In this unit we have discussed the main types of data used for marketing research in the context of the different research designs that may be employed. The sources of secondary and primary data have been discussed. The two basic methods of data collection i.e. through communication and through observation have. been explained alongwith their various modification. The unit ends with a discussion of the sources of error in data collection.

4.10 SELF-ASSESSMENT QUESTIONS

1) One of the important reasons for the use of surveys is that they can obtain sound information on what people actions in the future will be.

   Do you agree? Give reasons for your answer.

2) Discuss the main sources of primary and secondary data.

3) What is the type of data available from official publications?

4) What are the limitations associated with the use of secondary data?

5) What are the tools of collecting data from respondents?

6) Discuss the important sources of error in both secondary and primary data.