

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

Module 3

Part 1

METHODS AND TECHNIQUES OF DATA COLLECTION

Si no	Title of the topic	Page no
3.1.1	Introduction	2
3.1.2	Types of Data	2
3.1.3	Distinction between Primary Data and Secondary Data	7
3.1.4	Selection of appropriate method for data collection	11
3.1.5	Ethical considerations in Collection of Data	13
3.1.6	Precautions in Data Collection	14
3.1.7	Collection of secondary data	15
3.1.8	Data Collection Procedure for Primary Data	17

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

3.1.1 Introduction

Once the researcher has decided the ‘Research Design’, the next job is of data collection. Statistical investigation requires systematic collection of data so that all relevant groups are represented in the data.

For example, to determine the potential market for new product, a researcher might study 500 consumers in a certain geographical area. It must be ascertained that the group contains peoples representing variables such as income level, race, education and neighbourhood. The quality of data will greatly affect the conclusions and hence, utmost importance must be given to this process and every possible precaution should be taken to ensure accuracy while gathering and collecting data.

Depending upon the sources utilized whether the data has come from actual observations or from records that are kept for normal purposes, the statistical data can be classified into two categories.

3.1.2 Types of data

There are two major approaches for gathering information about a situation, problem or phenomenon. Sometimes, information required is already available and need only be extracted. However, there are times when the information must be collected. Based upon these broad approaches to information gathering, data are categorized as:

- Primary data.
- Secondary data;

In a nutshell, primary data and secondary data both have their advantages and disadvantages. Therefore, when carrying out research, it is left for the researcher to weigh these factors and choose the better one.

It is therefore important for one to study the similarities and differences between these data types so as to make proper decisions when choosing a better data type for research

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

work.

What is Primary Data?

Primary data is the kind of data that is collected directly from the data source without going through any existing sources. It is mostly collected specially for a research project and may be shared publicly to be used for other research. Primary data is often reliable, authentic, and objective in as much as it was collected with the purpose of addressing a particular research problem. It is noteworthy that primary data is not commonly collected because of the high cost of implementation.

Primary data is one which is collected by the investigator himself for the purpose of a specific inquiry or study. Such data is original in character and is generated by surveys conducted by individual or research institutions.

Pros	Cons
<p>(a) Primary data is specific to the needs of the researcher at the moment of data collection. The researcher is able to control the kind of data that is being collected.</p> <p>(b) It is accurate compared to secondary data. The data is not subjected to personal bias and as such the authenticity can be trusted.</p> <p>(c) The researcher exhibit</p>	<p>(a) Primary data is very expensive compared to secondary data. Therefore, it might be difficult to collect primary data.</p> <p>(b) It is time-consuming.</p> <p>(c) It may not be feasible to collect primary data in some cases due to its complexity and required commitment.</p> <p>(d) Unwillingness of respondents to provide information: This requires salesmanship on the</p>

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

<p>ownership of the data collected through primary research. He or she may choose to make it available publicly, patent it, or even sell it.</p> <p>(d) Primary data is usually up to date because it collects data in real-time and does not collect data from old sources.</p> <p>(e) The researcher has full control over the data collected through primary research. He can decide which design, method, and data analysis techniques to be used.</p>	<p>part of the interviewer. The interviewer may assure that the information will be kept secret or some present may be given.</p> <p>(e) Inability of the respondents to provide information: This may be due to</p> <ul style="list-style-type: none">i) Lack of Knowledgeii) Lapse of memoryiii) Inability to identify their motives and prove reason why for their actions.iv) Human Biases of the respondent etc.
--	--

What is Secondary Data?

Secondary data is the data that has been collected in the past by someone else but made available for others to use. They are usually once primary data but become secondary when used by a third party.

Secondary data are usually easily accessible to researchers and individuals because they are mostly shared publicly. This, however, means that the data are usually

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

general and not tailored specifically to meet the researcher's needs as primary data does.

For example, when conducting a research thesis, researchers need to consult past works done in this field and add findings to the literature review. Some other things like definitions and theorems are secondary data that are added to the thesis to be properly referenced and cited accordingly.

Some common sources of secondary data include trade publications, government statistics, journals, etc. In most cases, these sources cannot be trusted as authentic.

Pros	Cons
<p>(a) Secondary data is easily accessible compared to primary data. Secondary data is available on different platforms that can be accessed by the researcher.</p> <p>(b) Secondary data is very affordable. It requires little to no cost to acquire them because they are sometimes given out for free.</p> <p>(c) The time spent on collecting secondary data is usually very little compared to that of primary data.</p> <p>(d) Secondary data makes it</p>	<p>(a) Secondary data may not be authentic and reliable. A researcher may need to further verify the data collected from the available sources.</p> <p>(b) Researchers may have to deal with irrelevant data before finally finding the required data.</p> <p>(c) Some of the data is exaggerated due to the personal bias of the data source.</p> <p>(d) Secondary data sources are sometimes outdated with no new data to replace the old ones.</p>

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

possible to carry out longitudinal studies without having to wait for a long time to draw conclusions. (e) It helps to generate new insights into existing primary data.	
---	--

Similarities between Primary & Secondary Data

Contains Same Content:

Secondary data was once primary data when it was newly collected by the first researcher. The content of the data collected does not change and therefore has the same content with primary data.

It doesn't matter if it was further visualized in the secondary form, the content does not change. A common example of these are definitions, theorems, and postulates that were made years ago but still remain the same.

Uses

Primary data and secondary data are both used in research and statistics. They can be used to carry out the same kind of research in these fields depending on data availability. This is because secondary data and primary data have the same content. The only difference is the method by which they are collected.

Since the method of collection does not directly affect the uses of data, they can be used to perform similar research. For example, whether collected directly or from an

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

existing database, the demography of a particular target market can be used to inform similar business decisions.

3.1.3 Distinction between primary data and secondary data

1. Primary data is Original in character. Secondary data is Not original
2. Collection of data is expensive in the case of primary data. Collection of secondary data is less expensive.
3. Primary data is in the shape of raw materials. Secondary data is the shape of finished products.
4. Primary data is adequate and suitable. Secondary data need not be ample and apposite.

Data types

The type of data provided by primary data is real-time, while the data provided by secondary data is stale. Researchers are able to have access to the most recent data when conducting primary research, which may not be the case for secondary data.

Secondary data have to depend on primary data that has been collected in the past to perform research. In some cases, the researcher may be lucky that the data is collected close to the time that he or she is conducting research.

Therefore, reducing the amount of difference between the secondary data being used and the recent data.

Process

Researchers are usually very involved in the primary data collection process, while secondary data is quick and easy to collect. This is due to the fact that primary research is mostly longitudinal.

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

Therefore, researchers have to spend a long time performing research, recording information, and analyzing the data. This data can be collected and analyzed within a few hours when conducting secondary research.

 **Availability**

Primary data is available in crude form while secondary data is available in a refined form. That is, secondary data is usually made available to the public in a simple form for a layman to understand while primary data are usually raw and will have to be simplified by the researcher.

Secondary data are this way because they have previously been broken down by researchers who collected the primary data afresh.

 **Data Collection Tools**

Primary data can be collected using surveys and questionnaires while secondary data are collected using the library, books, etc. The different ones between these data collection tools are glaring and can it be interchangeably used.

 **Sources**

Primary data sources include; Surveys, observations, experiments, questionnaires, focus groups, interviews, etc., while secondary data sources include; books, journals, articles, web pages, blogs, etc. These sources vary explicitly and there is no intersection between the primary and secondary data sources.

Primary data sources are sources that require a deep commitment from researchers and require interaction with the subject of study. Secondary data, on the other hand, do not require interaction with the subject of study before it can be collected.

In most cases, secondary researchers do not have any interaction with the subject of research.

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

 **Specific**

Primary data is always specific to the researcher's needs, while secondary data may or may not be specific to the researcher's need. It depends solely on the kind of data the researcher was able to lay hands on.

 **Advantage**

Some common advantages of primary data are its authenticity, specific nature, and up to date information while secondary data is very cheap and not time-consuming.

Primary data is very reliable because it is usually objective and collected directly from the original source. It also gives up to date information about a research topic compared to secondary data.

Secondary data, on the other hand, is not expensive making it easy for people to conduct secondary research. It doesn't take so much time and most of the secondary data sources can be accessed for free.

 **Disadvantage**

The disadvantage of primary data is the cost and time spent on data collection while secondary data may be outdated or irrelevant. Primary data incur so much cost and takes time because of the processes involved in carrying out primary research.

For example, when physically interviewing research subjects, one may need one or more professionals, including the interviewees, videographers who will make a record of the interview in some cases and the people involved in preparing for the interview. Apart from the time required, the cost of doing this may be relatively high.

Secondary data may be outdated and irrelevant. In fact, researchers have to surf through irrelevant data before finally having access to the data relevant to the research purpose.

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

✚ **Accuracy and Reliability**

Primary data is more accurate and reliable while secondary data is relatively less reliable and accurate. This is mainly because the secondary data sources are not regulated and are subject to personal bias.

✚ **Cost-effectiveness**

Primary data is very expensive while secondary data is economical. When working on a low budget, it is better for researchers to work with secondary data, then analyze it to uncover new trends.

In fact, a researcher might work with both primary data and secondary data for one research. This is usually very advisable in cases whereby the available secondary data does not fully meet the research needs.

✚ **Collection Time**

The time required to collect primary data is usually long while that required to collect secondary data is usually short. The primary data collection process is sometimes longitudinal in nature.

Therefore, researchers may need to observe the research subject for some time while taking down important data. For example, when observing the behavior of a group of people or particular species, researchers have to observe them for a while.

✚ Secondary data can, however, be collected in a matter of minutes and analyzed to dead conclusions—taking a shorter time when compared to primary data. In some rare cases, especially when collecting little data, secondary data may take a longer time because of difficulty consulting different data sources to find the right data.

3.1.4 Selection of appropriate method for data collection

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

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

1. Nature, scope and object of enquiry: This constitutes the most important factor affecting the choice of a particular method. The method selected should be such that it suits the type of enquiry that is to be conducted by the researcher. This factor is also important in deciding whether the data already available (secondary data) are to be used or the data not yet available (primary data) are to be collected.

2. Availability of funds: Availability of funds for the research project determines to a large extent the method to be used for the collection of data. When funds at the disposal of the researcher are very limited, he will have to select a comparatively cheaper method which may not be as efficient and effective as some other costly method. Finance, in fact, is a big constraint in practice and the researcher has to act within this limitation.

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

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

But one must always remember that each method of data collection has its uses and none is superior in all situations. For instance, telephone interview method may be considered appropriate (assuming telephone population) if funds are restricted, time is also

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

restricted and the data is to be collected in respect of few items with or without a certain degree of precision. In case funds permit and more information is desired, personal interview method may be said to be relatively better. In case time is ample, funds are limited and much information is to be gathered with no precision, then mail-questionnaire method can be regarded more reasonable. When funds are ample, time is also ample and much information with no precision is to be collected, then either personal interview or the mail-questionnaire or the joint use of these two methods may be taken as an appropriate method of collecting data. Where a wide geographic area is to be covered, the use of mail-questionnaires supplemented by personal interviews will yield more reliable results per rupee spent than either method alone. The secondary data may be used in case the researcher finds them reliable, adequate and appropriate for his research. While studying motivating influences in market researches or studying people's attitudes in psychological/social surveys, we can resort to the use of one or more of the projective techniques stated earlier. Such techniques are of immense value in case the reason is obtainable from the respondent who knows the reason but does not want to admit it or the reason relates to some underlying psychological attitude and the respondent is not aware of it. But when the respondent knows the reason and can tell the same if asked, than a non-projective questionnaire, using direct questions, may yield satisfactory results even in case of attitude surveys. Since projective techniques are as yet in an early stage of development and with the validity of many of them remaining an open question, it is usually considered better to rely on the straight forward statistical methods with only supplementary use of projective techniques. Nevertheless, in pre-testing and in searching for hypotheses they can be highly valuable.

Thus, the most desirable approach with regard to the selection of the method depends on the nature of the particular problem and on the time and resources (money and

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

personnel) available along with the desired degree of accuracy. But, over and above all this, much depends upon the ability and experience of the researcher.

3.1.5 Ethical considerations in Collection of Data

Any researcher who involves human sample subjects in his research has certain responsibilities towards them. Since the activities of the sample subjects are often closely associated with data collection process, it is appropriate to consider ethical considerations here.

The following points have to be considered in process of data collection:

1. The researcher must protect the dignity and welfare of human sample subjects.
2. The human sample subjects freedom to decline participation must be respected, and the confidentiality of research data must be maintained.
3. The researcher must guard against violation or invasion of privacy.
4. The responsibility for maintaining ethical standard remains with the individual researcher and the principal investigator or supervisor is also responsible for actions of his scholars.

Any researcher anticipating “the use of human sample subjects should consult on ‘ethics’ statements such as those mentioned above.

A researcher should not mention the name of subjects anywhere in the report. If possible name of institutions where sample subjects have selected for data collection should not be mentioned even in the appendix. The code number should be used for this purpose. As a general rule, he must respect the human sample subjects selected in his specific research study.

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

3.1.6 Precautions in Data Collection

In the data collection the following precautions should be observed:

1. The data must be relevant to the research problem.
2. It should be collected through formal or standardized research tools.
3. The data should be such as these can be subjected to statistical treatment easily.
4. The data should have minimum measurement error.
5. The data must be tenable for the verification of the hypotheses.
6. The data should be such as parameters of the population may be estimated for inferential purpose.
7. The data should be complete in itself and also comprehensive in nature.
8. The data should be collected through objective procedure.
9. The data should be accurate and precise.
10. The data should be reliable and valid.
11. The data should be such that these can be presented and interpreted easily.
12. The scoring procedure of the research tool should be easy and objective.

3.1.7 Collection of secondary data

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

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 can obtain them. In this case he is certainly not confronted with the problems that are usually associated with the collection of original data. Secondary data may either be published data or unpublished data. Usually published data are available in:

- ✚ Various publications of the central, state and local governments;
- ✚ Various publications of foreign governments or of international bodies and their subsidiary organisations;
- ✚ Technical and trade journals;
- ✚ Books, magazines and newspapers;
- ✚ Reports and publications of various associations connected with business and industry, banks, stock exchanges, etc.;
- ✚ Reports prepared by research
- ✚ Scholars, universities, economists, etc. in different fields; and (g) public records and statistics, historical
- ✚ Documents, and other sources of published information. The sources of unpublished data are many;
- ✚ They may be found in diaries, letters, unpublished biographies and autobiographies and also may be
- ✚ Available with scholars and research workers, trade associations, labour bureaus and other public/
- ✚ Private individuals and organizations.

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

Researcher must be very careful in using secondary data. He must make a minute scrutiny because it is just possible that the secondary data may be unsuitable or may be inadequate in the

context of the problem which the researcher wants to study. In this connection Dr. A.L. Bowley very aptly observes that it is never safe to take published statistics at their face value without knowing their meaning and limitations and it is always necessary to criticize arguments that can be based on them.

By way of caution, the researcher, before using secondary data, must see that they possess following characteristics:

1. Reliability of data: The reliability can be tested by finding out such things about the said data:

- (a) Who collected the data?
- (b) What were the sources of data?
- (c) Were they collected by using proper methods?
- (d) At what time were they collected?
- (e) Was there any bias of the compiler?
- (f) What level of accuracy was desired?
- (g) Was it achieved?

2. Suitability of data: The data that are suitable for one enquiry may not necessarily be found suitable in another enquiry. Hence, if the available data are found to be unsuitable, they should not be used by the researcher. In this context, the researcher

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

must very carefully scrutinize the definition of various terms and units of collection used at the time of collecting the data from the primary source originally. Similarly, the object, scope and nature of the original enquiry must also be studied. If the researcher finds differences in these, the data will remain unsuitable for the present enquiry and should not be used.

3. Adequacy of data: If the level of accuracy achieved in data is found inadequate for the purpose of the present enquiry, they will be considered as inadequate and should not be used by the researcher.

The data will also be considered inadequate, if they are related to an area which may be either narrower or wider than the area of the present enquiry.

From all this we can say that it is very risky to use the already available data. The already available data should be used by the researcher only when he finds them reliable, suitable and adequate. But he should not blindly discard the use of such data if they are readily available from authentic sources and are also suitable and adequate for in that case it will not be economical to spend time and energy in field surveys for collecting information. At times, there may be wealth of usable information in the already available data which must be used by an intelligent researcher but with due precaution.

3.1.8 Data collection procedure for primary data

The various steps involved are:

- Planning the study
- Modes of data collection

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

- Sample selection
- Editing the primary data

a. Planning the Study:

Since the quality of results gained from statistical data depends upon the quality of information collected, it is important that a sound investigative process be established to ensure that the data is highly representative and unbiased. This requires a high degree of skill and also certain precautionary measures may have to be taken.

b. Modes of primary data collection:

Primary Data Collection Methods

Primary data collection methods are different ways in which primary data can be collected. It explains the tools used in collecting primary data, some of which are highlighted below:

 **Observation**

Observation method is mostly used in studies related to behavioral science. The researcher uses observation as a scientific tool and method of data collection. Observation as a data collection tool is usually systematically planned and subjected to checks and controls.

There are different approaches to the observation method—structured or unstructured, controlled or uncontrolled, and participant, non-participant, or disguised approach.

The structured and unstructured approach is characterized by careful definition of subjects of observation, style of observer, conditions, and selection of data. An observation process that satisfies this is said to be structured and vice versa.

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

A controlled and uncontrolled approach signifies whether the research took place in a natural setting or according to some pre-arranged plans. If an observation is done in a natural setting, it is uncontrolled but becomes controlled if done in a laboratory.

Before employing a new teacher, academic institutions sometimes ask for a sample teaching class to test the teacher's ability. The evaluator joins the class and observes the teaching, making him or her participant.

The evaluation may also decide to observe from outside the class, becoming a non-participant. An evaluator may also be asked to stay in class and disguise as a student, in order to carry out a disguised observation.

Pros	Cons
The data is usually objective.	Data is not affected by past or future events. The information is limited. It is expensive

 **Interviews:**

Interview is a method of data collection that involves two groups of people, where the first group is the interviewer the researcher(s) asking questions and collecting data and the interviewee (the subject or respondent that is being asked questions). The questions and responses during an interview may be oral or verbal as the case may be.

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

Interviews can be carried out in 2 ways, namely; in-person interviews and telephonic interviews. An in-person interview requires an interviewer or a group of interviewers to ask questions from the interviewee in a face to face fashion.

It can be direct or indirect, structured or structure, focused or unfocused, etc. Some of the tools used in carrying out in-person interviews include a notepad or recording device to take note of the conversation—very important due to human forgetful nature.

Telephonic interviews, on the other hand, are carried out over the phone through ordinary voice call or video calls. The 2 parties involved may decide to use video calls like Skype to carry out interviews.

A mobile phone, Laptop, Tablet or desktop computer with an internet connection is required for this.

Pros	Cons
In-depth information can be collected. Non-response and response bias can detected. The samples can be controlled.	It is more time-consuming. It is expensive. The interviewer may be biased.

 **Surveys & Questionnaires**

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

Surveys and questionnaires are 2 similar tools used in collecting primary data. They are a group of questions typed or written down and sent to the sample of study to give responses.

After giving the required responses, the survey is given back to the researcher to record. It is advisable to conduct a pilot study where the questionnaires are filled by experts and meant to assess the weakness of the questions or techniques used.

There are 2 main types of surveys used for data collection, namely; online and offline surveys. Online surveys are carried out using internet-enabled devices like mobile phones, PCs, Tablets, etc.

They can be shared with respondents through email, websites, or social media. Offline surveys, on the other hand, do not require an internet connection for it to be carried out.

The most common type of offline survey is paper-based surveys.

Pros	Cons
Respondents have adequate time to give responses. It is free from the bias of the interviewer. They are cheaper compared to interviews	A high rate of non-response bias. It is inflexible and can't be changed once sent. It is a slow process.

Conclusion

When performing research, it is important to consider the available data options so as to ensure that the right type of data is used to arrive at a feasibility conclusion. A good

Module 3 Part 1 Methods of data collection – Types of Data
Dr Josheena Jose, Assistant Professor, PG Department of Commerce
Christ College (Autonomous) Irinjalakuda

understanding of the different data types, similarities, and differences are however required to do this.

Primary data and secondary data both have applications in business and research. They may, however, differ from each other in the way in which they are collected, used, and analyzed. Primary sources provide first hand information and secondary sources provide secondary hand data.

The most common setback with primary data is that it is very expensive, which is not the case for secondary data. Secondary data, on the other hand, has authenticity issues.

In the following section each method of data collection will be discussed from the point of view of its applicability and suitability to a situation, and the problems and limitations associated with it.

#####