

# **MODULE 2: DATA COLLECTION & SAMPLING METHODS**

Dr. A. B. Patankar (Athawale)

# Contents

- Primary & Secondary data
- Sources of data
- Methods of collecting data.
- Sampling – Census & Sample methods
- Methods of sampling
- Probability Sampling and Non-Probability Sampling

# Data Collection Methods: Definition, Examples and Sources



## DATA COLLECTION

Statistical Methods

Surveys

Polls

Interview

Delphi Technique

Focus Groups

Financial Reports

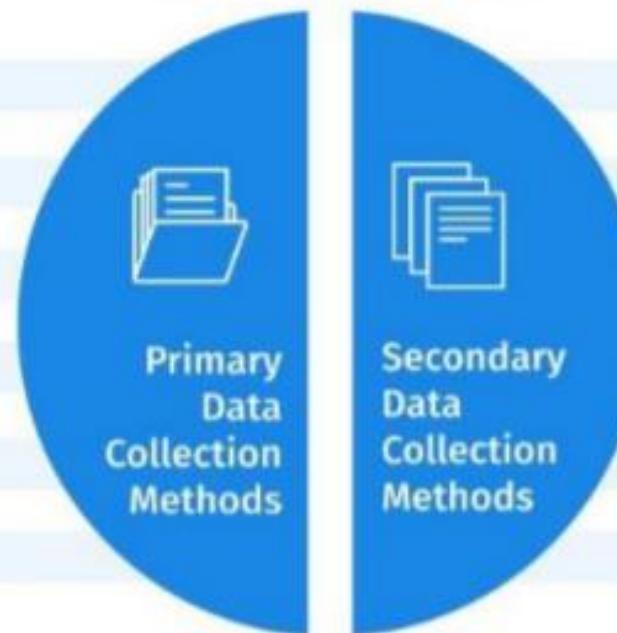
Sales Reports

Government Reports

Mission

Vision Statement

Internet



**Data** is a collection of facts, figures, objects, symbols, and events gathered from different sources. Organizations collect data to make better decisions. Without data, it would be difficult for organizations to make appropriate decisions, and so data is collected at various points in time from different audiences. For instance, before launching a new product, an organization needs to collect data on product demand, customer preferences, competitors, etc. In case data is not collected beforehand, the organization's newly launched product may lead to failure for many reasons, such as less demand and inability to meet customer needs.

Although data is a valuable asset for every organization, it does not serve any purpose until analyzed or processed to get the desired results.

**Data collection** is a process of collecting information from all the relevant sources to find answers to the research problem, test the hypothesis and evaluate the outcomes.

## **SOURCES OF DATA:**

Sources of Data can be classified into 2 types. Statistical sources refer to data that are gathered for some official purposes and incorporate censuses and officially administered surveys. Non-statistical sources refer to the collection of data for other administrative purposes or for the private sector.

What are the different sources of data?

Following are the two sources of data:

### **1. Internal Source**

When data are collected from reports and records of the organisation itself, it is known as  

---

the internal source.  

---

For example, a company publishes its ‘Annual Report’ on Profit and Loss, Total Sales, Loans, Wages etc.

## **2. External Source**

When data are collected from outside the organisation, it is known as the external source.

---

For example, if a Tour and Travels Company obtains information on ‘Karnataka Tourism’ from Karnataka Transport Corporation, it would be known as external sources of data.

## **Types of Data**

### **A) Primary Data**

---

Primary data means ‘First-hand information’ collected by an investigator.

It is collected for the first time.

It is original and more reliable.

For example Population census conducted by the government of India after every 10 years.

## **B) Secondary Data**

Secondary data refers to ‘Second-hand information’.

---

These are not originally collected rather obtained from already published or unpublished sources.

For example the Address of a person taken from the Telephone Directory or Phone number of a company taken from ‘Just Dial’.

Students can also refer to Meaning and Sources of Secondary Data

# **Methods of Collecting Primary Data**

**Direct Personal Investigation**

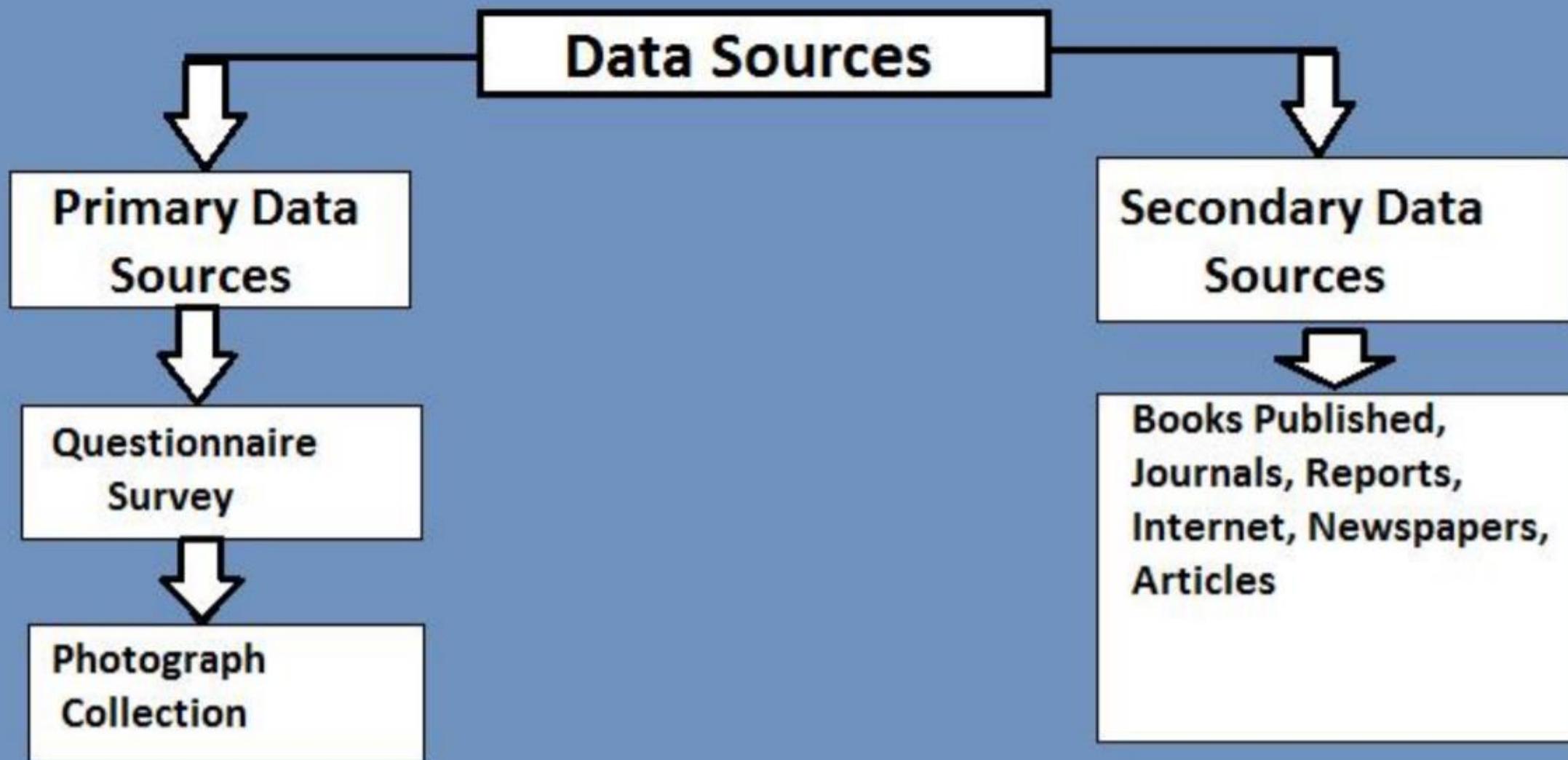
**Indirect Oral Investigation**

**Information Through Correspondents**

**Telephonic Interview**

**Mailed Questionnaire**

**The questionnaire filled by enumerators**



## **Procedure for data collection**

. There are various steps commonly involved in the Data Collection process.

### **1. Identifying opportunities/issues for data collection**

It involves identifying the particular opportunities/issues for collecting data. An in-depth analysis of external and internal processes is vital for getting a clear picture of all that occurs within an organization. By identifying and exploring specific issues in the data collection process, you can get the flexibility to make decisions about when and how you should collect information.

## **2. Set objectives**

Once you consider the important questions, you have to set clear goals individualized for every issue based on the collection analysis and techniques. Every opportunity or issue is reviewed in this step from the external and internal evaluation that is carried out earlier. A specific issue is chosen to start with.

### **3. Collecting Data**

Once every question is clearly defined and the goals are properly set, information is collected. A committee is set up and specific people are chosen. These people are likely to be held accountable for every important decision that is taken regarding the collection process including finances, coordination, logistics, and designs. The logistics, resources, and technology needed for the implementation of data collection, as well as the people in charge of the data collection, are identified and determined.

## **4. Data Analysis and interpretation**

Based on the qualitative or quantitative data collection techniques, data analysis could be intimidating at times. Businesses need to determine whether there are sufficient expertise and internal capacity for in-house data interpretation or whether a professional consultant should be hired for the task.

Often times, traditional data collection does not progress as fast as one would generally prefer, which is why it is important that CrowdForce provides access to fast and reliable data. Our network of field agents makes it possible to collect real-time data in different locations simultaneously. The hassle associated with data collection is drastically reduced as we combine steps 3 & 4 to reduce the amount of time used to gather the required data.

## **TOOLS OF DATA COLLECTION / INSTRUMENTS FOR DATA COLLECTION**

Data collection is an important step in the research process. The instrument you choose to collect the data will depend on the type of data you plan on collecting (qualitative or quantitative) and how you plan to collect it.

A number of common data-collecting instruments are used in construction research:

1. Questionnaires
2. Interviews
3. Observations
4. Archival documents and government sources
5. Laboratory experiments
6. Quasi experiment

## **Primary Data Collection Methods**

Primary data is collected from the first-hand experience and is not used in the past. The data gathered by primary data collection methods are specific to the research's motive and highly accurate.

Primary data collection methods can be divided into two categories: quantitative methods and qualitative methods.

## **Quantitative Methods:**

Quantitative techniques for market research and demand forecasting usually make use of statistical tools. In these techniques, demand is forecast based on historical data. These methods of primary data collection are generally used to make long-term forecasts. Statistical methods are highly reliable as the element of subjectivity is minimum in these methods.

## **Time Series Analysis**

The term time series refers to a sequential order of values of a variable, known as a trend, at equal time intervals. Using patterns, an organization can predict the demand for its products and services for the projected time.

## **Smoothing Techniques**

In cases where the time series lacks significant trends, smoothing techniques can be used. They eliminate a random variation from the historical demand. It helps in identifying patterns and demand levels to estimate future demand. The most common methods used in smoothing demand forecasting techniques are the simple moving average method and the weighted moving average method.

## **Barometric Method**

Also known as the leading indicators approach, researchers use this method to speculate future trends based on current developments. When the past events are considered to predict future events, they act as leading indicators.

## **Qualitative Methods:**

Qualitative methods are especially useful in situations when historical data is not available. Or there is no need of numbers or mathematical calculations. Qualitative research is closely associated with words, sounds, feeling, emotions, colors, and other elements that are non-quantifiable. These techniques are based on experience, judgment, intuition, conjecture, emotion, etc.

Quantitative methods do not provide the motive behind participants' responses, often don't reach underrepresented populations, and span long periods to collect the data. Hence, it is best to combine quantitative methods with qualitative methods.

## Surveys

Surveys are used to collect data from the target audience and gather insights into their preferences, opinions, choices, and feedback related to their products and services. Most survey maker software often a wide range of question types to select.

You can also use a ready-made survey template to save on time and effort. Online surveys can be customized as per the business's brand by changing the theme, logo, etc. They can be distributed through several distribution channels such as email, website, offline app, QR code, social media, etc. Depending on the type and source of your audience, you can select the channel.

Once the data is collected, survey software can generate various reports and run analytics algorithms to discover hidden insights. A survey dashboard can give you the statistics related to response rate, completion rate, filters based on demographics, export and sharing options, etc. You can maximize the effort spent on online data collection by integrating survey builder with third-party apps.

## Polls

Polls comprise of one single or multiple choice question. When it is required to have a quick pulse of the audience's sentiments, you can go for polls. Because they are short in length, it is easier to get responses from the people.

Similar to surveys, online polls, too, can be embedded into various platforms. Once the respondents answer the question, they can also be shown how they stand compared to others' responses.

## **Interviews**

In this method, the interviewer asks questions either face-to-face or through telephone to the respondents. In face-to-face interviews, the interviewer asks a series of questions to the interviewee in person and notes down responses. In case it is not feasible to meet the person, the interviewer can go for a telephonic interview. This form of data collection is suitable when there are only a few respondents. It is too time-consuming and tedious to repeat the same process if there are many participants.

## **Delphi Technique**

In this method, market experts are provided with the estimates and assumptions of forecasts made by other experts in the industry. Experts may reconsider and revise their estimates and assumptions based on the information provided by other experts. The consensus of all experts on demand forecasts constitutes the final demand forecast.

## **Focus Groups**

A small group of people, around 8-10 members, discuss the common areas of the problem. Each individual provides his insights on the issue concerned. A moderator regulates the discussion among the group members. At the end of the discussion, the group reaches a consensus.

## **Questionnaire**

A questionnaire is a printed set of questions, either open-ended or closed-ended. The respondents are required to answer based on their knowledge and experience with the issue concerned. The questionnaire is a part of the survey, whereas the questionnaire's end-goal may or may not be a survey.

## **Secondary Data Collection Methods**

Secondary data is the data that has been used in the past. The researcher can obtain data from the sources, both internal and external, to the organization.

**Internal sources of secondary data:**

1. Organization's health and safety records
2. Mission and vision statements
3. Financial Statements
4. Magazines
5. Sales Report
6. CRM Software
7. Executive summaries

**External sources of secondary data:**

1. Government reports
2. Press releases
3. Business journals
4. Libraries
5. Internet

The secondary data collection methods, too, can involve both quantitative and qualitative techniques. Secondary data is easily available and hence, less time-consuming and expensive as compared to the primary data. However, with the secondary data collection methods, the authenticity of the data gathered cannot be verified.

## **Secondary Data Collection Methods**

Secondary data is data collected by someone other than the actual user. It means that the information is already available, and someone analyses it. The secondary data includes magazines, newspapers, books, journals etc. It may be either published data or unpublished data.

1. Published data are available in various resources including
2. Government publications
3. Public records
4. Historical and statistical documents
5. Business documents
6. Technical and trade journals
7. Unpublished data includes
8. Diaries
9. Letters
10. Unpublished biographies etc.

## **Questionnaire method of data collection**

A questionnaire is a research instrument consisting of a series of questions for the purpose of gathering information from respondents. Questionnaires can be thought of as a kind of written interview. They can be carried out face to face, by telephone, computer or post. Questionnaires provide a relatively cheap, quick and efficient way of obtaining large amounts of information from a large sample of people. Questionnaire is as an instrument for research, which consists of a list of questions, along with the choice of answers, printed or typed in a sequence on a form used for acquiring specific information from the respondents. The questionnaire is prepared in such a way that it translates the required information into a series of questions, that informants can and will answer.

## **Characteristics of a Good Questionnaire**

The following are characteristics of good questionnaires:

1. It should consist of a well-written list of questions.
2. The questionnaire should deal with an important or significant topic to create interest among respondents.
3. It should seek only that data which cannot be obtained from other sources.
4. It should be as short as possible but should be comprehensive.
5. It should be attractive.
6. Directions should be clear and complete.
7. It should be represented in good psychological order proceeding from general to more specific responses.
8. Double negatives in questions should be avoided.
9. Putting two questions in one question also should be avoided. Every question should seek to obtain only one specific information.
10. It should be designed to collect information which can be used subsequently as data for analysis.

## **Format of Questions in Questionnaires**

The questions asked can take two forms:

**Restricted questions**, also called closed-ended, ask the respondent to make choices — yes or no, check items on a list, or select from multiple choice answers. Restricted questions are easy to tabulate and compile.

**Unrestricted questions** are open-ended and allow respondents to share feelings and opinions that are important to them about the matter at hand.

Unrestricted questions are not easy to tabulate and compile, but they allow respondents to reveal the depth of their emotions.

If the objective is to compile data from all respondents, then sticking with restricted questions that are easily quantified is better.

If degrees of emotions or depth of sentiment is to be studied, then develop a scale to quantify those feelings.

## **Advantages of Questionnaire**

1. One of the greatest benefits of questionnaires lies in their uniformity — all respondents see exactly the same questions.
2. It is an inexpensive method, regardless of the size of the universe.
3. Free from the bias of the interviewer, as the respondents answer the questions in his own words.
4. Respondents have enough time to think and answer.
5. Due to its large coverage, respondents living in distant areas can also be reached conveniently.
6. Comparability

## **Limitations of Questionnaire**

The risk of collection of inaccurate and incomplete information is high in the questionnaire, as it might happen that people may not be able to understand the question correctly.

The main demerits of this system can also be listed here:

1. Low rate of return of the duly filled in questionnaires; bias due to no-response is often Indeterminate.
2. It can be used only when respondents are educated and cooperating.
3. The control over questionnaire may be lost once it is sent.
4. There is inbuilt inflexibility because of the difficulty of amending the approach once
5. questionnaires have been dispatched.
6. There is also the possibility of ambiguous replies or omission of replies altogether to certain
7. Questions; interpretation of omissions is difficult.
8. It is difficult to know whether willing respondents are truly representative.
9. This method is likely to be the slowest of all.

## Interview Schedules

Data Collection through Schedules – Very similar to the Questionnaire method. The main difference is that a schedule is filled by the trained enumerator who is specially appointed for the purpose. Enumerator goes to the respondents, asks them the questions from the Performa in the order listed, and records the responses in the space provided.

**Open-ended** – Questions in which the respondent answers in his own words.

**Closed-ended** (or Fixed Alternative) – Question in which respondent selects one or more options from pre-determined set of responses.

**Simple dichotomy** → Closed ended question with only two response alternatives

**Multiple Choice** → Closed ended question with more than two response alternatives.

**Determinant choice** – Multiple choice question in which respondent must select only one of the response alternatives.

**Checklist question** - Multiple choice question in which respondent can select more than one of the response alternatives.

**Investigator** - One who conducts the investigation i.e. statistical enquiry and seeks information is known as Investigator. It can be an individual person or an organization.

**Enumerators**- Enumerators are the persons who help the Investigators in the collection of data.

**Informant** - Informants are the respondents who supply the information to the investigator or enumerators.

## **Interview Method of Data Collection**

### **Types of interviews**

An interview is generally a qualitative research technique which involves asking open-ended questions to converse with respondents and collect elicit data about a subject. The interviewer in most cases is the subject matter expert who intends to understand respondent opinions in a well-planned and executed series of questions and answers. Interviews are similar to focus groups and surveys when it comes to garnering information from the target market but are entirely different in their operation – focus groups are restricted to a small group of 6-10 individuals whereas surveys are quantitative in nature. Interviews are conducted with a sample from a population and the key characteristic they exhibit is their conversational tone.

## **Fundamental Types of Interviews in Research**

A researcher has to conduct interviews with a group of participants at a juncture in the research where information can only be obtained by meeting and personally connecting with a section of their target audience. Interviews offer the researchers with a platform to prompt their participants and obtain inputs in the desired detail. There are three fundamental types of interviews in research:

## **Structured Interviews:**

Structured interviews are defined as research tools that are extremely rigid in their operations and allows very little or no scope of prompting the participants to obtain and analyze results. It is thus also known as a standardized interview and is significantly quantitative in its approach. Questions in this interview are pre-decided according to the required detail of information.

Structured interviews are excessively used in survey research with the intention of maintaining uniformity throughout all the interview sessions.

They can be closed-ended as well as open-ended – according to the type of target population. Closed-ended questions can be included to understand user preferences from a collection of answer options whereas open-ended can be included to gain details about a particular section in the interview.

## **Advantages of structured interviews:**

Structured interviews focus on the accuracy of different responses due to which extremely organized data can be collected. Different respondents have different type of answers to the same structure of questions – answers obtained can be collectively analyzed.

1. They can be used to get in touch with a large sample of the target population.
2. The interview procedure is made easy due to the standardization offered by structured interviews.
3. Replication across multiple samples becomes easy due to the same structure of interview.
4. As the scope of detail is already considered while designing the interview, better information can be obtained and the researcher can analyze the research problem in a comprehensive manner by asking accurate research questions.
5. Since the structure of the interview is fixed, it often generates reliable results and is quick to execute.
6. The relationship between the researcher and the respondent is not formal due to which the researcher can clearly understand the margin of error in case the respondent either decides to be a part of the survey or is just not interested in providing the right information.

### **Disadvantages of structured interviews:**

1. Limited scope of assessment of obtained results.
2. The accuracy of information overpowers the detail of information.
3. Respondents are forced to select from the provided answer options.
4. The researcher is expected to always adhere to the list of decided questions irrespective of how interesting the conversation is turning out to be with the participants.
5. A significant amount of time is required for a structured interview.
6. Learn more: Market Research

## **Semi-Structured Interviews:**

Semi-structured interviews offer a considerable amount of leeway to the researcher to probe the respondents along with maintaining basic interview structure. Even if it is a guided conversation between researchers and interviewees – an appreciable flexibility is offered to the researchers. A researcher can be assured that multiple interview rounds will not be required in the presence of structure in this type of research interview.

Keeping the structure in mind, the researcher can follow any idea or take creative advantage of the entire interview. Additional respondent probing is always necessary to garner information for a research study. The best application of semi-structured interview is when the researcher doesn't have time to conduct research and requires detailed information about the topic.

## **Advantages of semi-structured interviews:**

Questions of semi-structured interviews are prepared before the scheduled interview which provides the researcher with time to prepare and analyze the questions.

1. It is flexible to an extent while maintaining the research guidelines.
2. Researchers can express the interview questions in the format they prefer, unlike the structured interview.
3. Reliable qualitative data can be collected via these interviews.
4. Flexible structure of the interview.

## **Disadvantages of semi-structured interviews:**

Participants may question the reliability factor of these interviews due to the flexibility offered.

Comparing two different answers becomes difficult as the guideline for conducting interviews is not entirely followed. No two questions will have the exact same structure and the result will be an inability to compare and infer results.

## **Unstructured Interviews:**

Also called as in-depth interviews, unstructured interviews are usually described as conversations held with a purpose in mind – to gather data about the research study. These interviews have the least number of questions as they lean more towards a normal conversation but with an underlying subject.

The main objective of most researchers using unstructured interviews is to build a bond with the respondents due to which there are high chances that the respondents will be 100% truthful with their answers. There are no guidelines for the researchers to follow and so, they can approach the participants in any ethical manner to gain as much information as they possibly can for their research topic.

Since there are no guidelines for these interviews, a researcher is expected to keep their approach in check so that the respondents do not sway away from the main research motive. For a researcher to obtain the desired outcome, he/she must keep the following factors in mind:

### **Intent of the interview.**

The interview should primarily take into consideration the participant's interest and skills.

All the conversations should be conducted within permissible limits of research and the researcher should try and stick by these limits.

The skills and knowledge of the researcher should match the purpose of the interview.

Researchers should understand the do's and don'ts of unstructured interviews.

## **Intent of the interview.**

The interview should primarily take into consideration the participant's interest and skills.

All the conversations should be conducted within permissible limits of research and the researcher should try and stick by these limits.

The skills and knowledge of the researcher should match the purpose of the interview.

Researchers should understand the do's and don'ts of unstructured interviews.

## **Advantages of Unstructured Interviews:**

Due to the informal nature of unstructured interviews – it becomes extremely easy for researchers to try and develop a friendly rapport with the participants. This leads to gaining insights in extreme detail without much conscious effort.

The participants can clarify all their doubts about the questions and the researcher can take each opportunity to explain his/her intention for better answers.

There are no questions which the researcher has to abide by and this usually increases the flexibility of the entire research process.

## **Disadvantages of Unstructured Interviews:**

As there is no structure to the interview process, researchers take time to execute these interviews.

The absence of a standardized set of questions and guidelines indicates that the reliability of unstructured interviews is questionable.

In many cases, the ethics involved in these interviews are considered borderline upsetting.

## **Methods of Research Interviews:**

There are three methods to conduct research interviews, each of which is peculiar in its application and can be used according to the research study requirement.

### **Personal Interviews:**

Personal interviews are one of the most used types of interviews, where the questions are asked personally directly to the respondent. For this, a researcher can have a guide online surveys to take note of the answers. A researcher can design his/her survey in such a way that they take notes of the comments or points of view that stands out from the interviewee.

## **Advantage:**

1. More information and that too in greater depth can be obtained.
2. Interviewer by his own skill can overcome the resistance, if any, of the respondents; the interview method can be made to yield an almost perfect sample of the general population.
3. There is greater flexibility under this method as the opportunity to restructure questions is always there, specially in case of unstructured interviews.
4. Observation method can as well be applied to recording verbal answers to various questions.
5. Personal information can as well be obtained easily under this method.

6. Samples can be controlled more effectively as there arises no difficulty of the missing returns; non-response generally remains very low.
7. The interviewer can usually control which person(s) will answer the questions. This is not possible in mailed questionnaire approach. If so desired, group discussions may also be held.
8. The interviewer may catch the informant off-guard and thus may secure the most spontaneous reactions than would be the case if mailed questionnaire is used.
9. The language of the interview can be adopted to the ability or educational level of the person interviewed and as such misinterpretations concerning questions can be avoided.
10. The interviewer can collect supplementary information about the respondent's personal characteristics and environment which is often of great value in interpreting results.

## **Disadvantages:**

1. It is a very expensive method, especially when large and widely spread geographical sample is taken.
2. There remains the possibility of the bias of interviewer as well as that of the respondent; there also remains the headache of supervision and control of interviewers.
3. Certain types of respondents such as important officials or executives or people in high income groups may not be easily approachable under this method and to that extent the data may prove inadequate.
4. This method is relatively more-time-consuming, specially when the sample is large and recalls upon the respondents are necessary.

5. The presence of the interviewer on the spot may over-stimulate the respondent, sometimes even to the extent that he may give imaginary information just to make the interview interesting.
6. Under the interview method the organisation required for selecting, training and supervising the field-staff is more complex with formidable problems.
7. Interviewing at times may also introduce systematic errors.
8. Effective interview presupposes proper rapport with respondents that would facilitate free and frank responses. This is often a very difficult requirement.

## **Telephonic Interviews:**

Telephone interviews: This method of collecting information consists in contacting respondents on telephone itself. It is not a very widely used method, but plays important part in industrial surveys, particularly in developed regions. The chief merits of such a system are:

## **Advantages:**

1. To find the interviewees it is enough to have their telephone numbers on hand.
2. They are usually lower cost.
3. The information is collected quickly.
4. Having a personal contact can also clarify doubts, or give more details of the questions.

## **Disadvantages:**

1. It is more flexible in comparison to mailing method.
2. It is faster than other methods i.e., a quick way of obtaining information.
3. It is cheaper than personal interviewing method; here the cost per response is relatively low.
4. Recall is easy; callbacks are simple and economical.
5. There is a higher rate of response than what we have in mailing method; the non-response is generally very low.
6. Replies can be recorded without causing embarrassment to respondents.
7. Interviewer can explain requirements more easily.
8. At times, access can be gained to respondents who otherwise cannot be contacted for one reason or the other.
9. No field staff is required.
10. Representative and wider distribution of sample is possible.

## **Email or Web Page Interviews:**

Online research is growing more and more because consumers are migrating to a more virtual world and it is best for each researcher to adapt to this change.

The increase in people with Internet access has made it popular that interviews via email or web page stand out among the types of interviews most used today. For this nothing better than an online survey.

More and more consumers are turning to online shopping, which is why they are a great niche to be able to carry out an interview that will generate information for the correct decision making.

### **Advantages of email surveys:**

Speed in obtaining data

The respondents respond according to their time, at the time they want and in the place they decide.

Online surveys can be mixed with other research methods or using some of the previous interview models. They are tools that can perfectly complement and pay for the project.

A researcher can use a variety of questions, logics, create graphs and reports immediately.

Undoubtedly, the objective of the research will set the pattern of what types of interviews are best for data collection. Based on the research design, a research can plan and test the questions, for instance, if the questions are the correct and if the survey flows in the best way.

In addition there are other types of research that can be used under specific circumstances, for example in the case of no connection or adverse situations to carry out surveys, in these types of occasions it is necessary to conduct a field research, which can not be considered an interview if not rather a completely different methodology.

To summarize the discussion, an effective interview will be one that provides researchers with the necessary data to know the object of study and that this information is applicable to the decisions researchers make.  
Dr. A. B. Patankar (Athawale)