A REVIEW OF RESEARCH PROCESS: DATA COLLECTION AND ANALYSIS

Dr. Anjana Kumari
Assistant Professor, Department of Education,
Sri Ramakrishna Sarada Ashrama Teacher's Training College, Jharkhand
E-mail: anjanasinha1111@gmail.com

ABSTRACT

Research is a highly specialized activity that entails much more than simply gathering data or writing a description. It entails gathering information in a specific manner, which is then thoroughly analyzed to provide answers to research questions and evaluate results. Regardless of the subject of investigation, data collecting is at the center of any research strategy. Any research project begins with a set of questions that must be answered. Data collection is the process of carefully acquiring the desired information with the least amount of distortion possible so that the analysis can produce responses that are believable and logical. Data collection tools refer to the devices/instruments used to collect data, such as a paper questionnaire or computer-assisted interviewing system. Case Studies, Checklists, Interviews, Observation sometimes, and Surveys or Questionnaires are all tools used to collect data.

Keywords: Data Collection, Experimental, Primary Data, Observational, Secondary Data, Simulation

INTRODUCTION

Data collection is the process of acquiring and evaluating the information on variables of interest in a systematic manner that allows researchers to answer research questions, test hypotheses, and assess outcomes. All fields of study, including physical and social sciences, humanities, business, and others, use data collection as part of their research. While the methodologies differ depending on the discipline, the emphasis on accurate and honest data collection stays the same. The process of collecting, measuring, and evaluating correct insights for research using established approved procedures is referred to as data collection. On the basis of the facts gathered, a researcher might evaluate their hypothesis. Regardless of the subject of study, data collecting is usually the first and most significant phase in the research process. Depending on the information needed, different approaches to data gathering are used in different disciplines of study.

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. It is important to decide the tools for data

collection because research is carried out in different ways and for different purposes. The objective behind data collection is to capture quality evidence that allows analysis to lead to the formulation of convincing and credible answers to the posed questions.
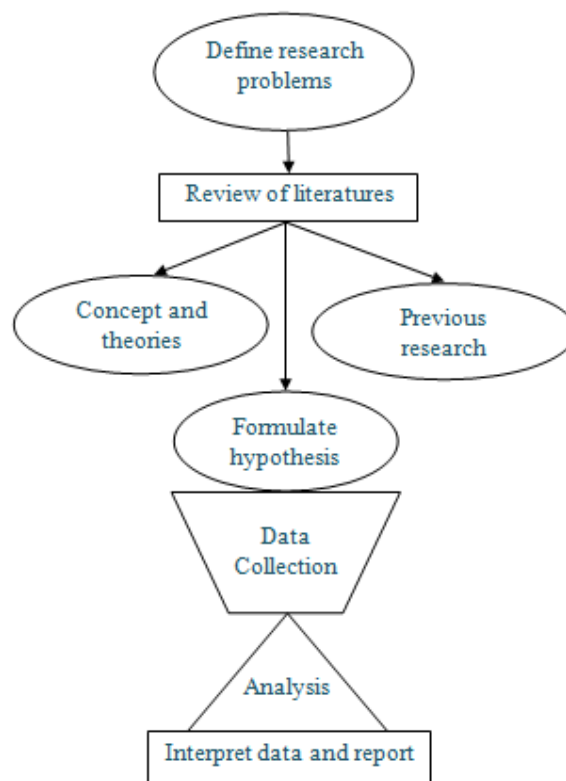
What is the origin of data?
There are two kinds of data: primary data and secondary data. Secondary data can be gathered from various government departments such as health, education, and population (e.g. hospitals, clinics, and school records) and used in our own research. Secondary sources include private foundation databases, local and county government data, surveillance data from government programmes, and statistics from federal agencies such as the Census and the National Institutes of Health (NIH). If the information was obtained from a government agency, this secondary data could save us money and time while also ensuring accuracy. It does, however, have some disadvantages. Secondary data may be found 'out of date' during analysis. It's possible that the data wasn't collected for a long enough period of time to detect trends, such as a two-month-old organism pattern in a hospital. We must formulate the research, which is a significant constraint. Additionally, we can break up data into qualitative and quantitative types. Qualitative data covers descriptions such as color, size, quality, and appearance. Quantitative data, unsurprisingly, deals with numbers, such as statistics, poll numbers, percentages, etc.

Primary Data Collection Methods
Primary data is collected from 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
- Qualitative methods

In primary data collection, you collect the data yourself using qualitative and quantitative methods. The key point here is that the data you collect is unique to you and your research and, until you publish, no one else has access to it. There are many methods of collecting primary data.

The main methods include:

- Questionnaires
- Interviews
- Focus Group Interviews
- Observation
- Survey
- Case-studies
- Diaries
- Activity Sampling Technique
- Memo Motion Study
- Process Analysis
- Link Analysis
- Time and Motion Study
- Experimental Method
- Statistical Method etc.

Questionnaires:
Questionnaires are a simple, straightforward data collection method. Respondents get a series of questions, either open or close-ended, related to the matter at hand.

Interviews:
The researcher asks questions of a large sampling of people, either by direct interviews or means of mass communication such as by phone or mail. This method is by far the most common means of data gathering.

Projective Technique:
Projective data gathering is an indirect interview, used when potential respondents know why they're being asked questions and hesitate to answer. For instance, someone may be reluctant to answer questions about their phone service if a cell phone carrier representative poses the questions. With projective data gathering, the interviewees get an incomplete question, and they must fill in the rest, using their opinions, feelings, and attitudes.

Focus Groups:
Focus groups, like interviews, are a commonly used technique. The group consists of anywhere from a half-dozen to a dozen people, led by a moderator, brought together to discuss the issue.

Data Collection Tools

Now that we've explained the various techniques, let's narrow our focus even further by looking at some specific tools. For example, we mentioned interviews as a technique, but we can further break that down into different interview types (or "tools").

Word Association - The researcher gives the respondent a set of words and asks them what comes to mind when they hear each word.

Sentence Completion - Researchers use sentence completion to understand what kind of ideas the respondent has. This tool involves giving an incomplete sentence and seeing how the interviewee finishes it.

Role-Playing - Respondents are presented with an imaginary situation and asked how they would act or react if it was real.

In-Person Surveys - The researcher asks questions in person.

Online/Web Surveys - These surveys are easy to accomplish, but some users may be unwilling to answer truthfully, if at all.

Mobile Surveys - These surveys take advantage of the increasing proliferation of mobile technology. Mobile collection surveys rely on mobile devices like tablets or smartphones to conduct surveys via SMS or mobile apps.

Phone Surveys - No researcher can call thousands of people at once, so they need a third party to handle the chore. However, many people have call screening and won't answer.

Advantages of Using Primary Data:

- The investigator collects data specific to the problem under study.
- There is no doubt about the quality of the data collected (for the investigator).
- If required, it may be possible to obtain additional data during the study period.

Disadvantages of Using Primary Data:

1. The investigator has to contend with all the hassles of data collection-
- deciding why, what, how, when to collect;
- getting the data collected (personally or through others);
- getting funding and dealing with funding agencies;
- ethical considerations (consent, permissions, etc.).

2. Ensuring the data collected is of a high standard-
- all desired data is obtained accurately, and in the format, it is required in;
- there is no fake/ cooked up data;
- unnecessary/ useless data has not been included.

3. Cost of obtaining the data is often the major expense in studies.

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. Sources of secondary data include books, personal sources, journals, newspapers, websites, government records, etc. Secondary data are known to be readily available compared to that of primary data. It requires very little research and needs for manpower to use these sources.

In other words, data collected from a source that has already been published in any form is called secondary data. The review of literature in any research is based on secondary data. It is collected by someone else for some other purpose (but being utilized by the investigator for another purpose). For example, Census data is used to analyze the impact of education on career choice and earning. Common sources of secondary data for social science include censuses, organizational records, anddata collected through qualitative methodologies or qualitative research. Secondary data is essential since it is impossible to conduct a new survey that can adequately capture past changes and/or developments.

Sources of Secondary Data:
The following are some ways of collecting secondary data –

- Books
- Records
- Biographies
- Newspapers
- Published censuses or other statistical data
- Data archives
- Internet articles
- Research articles by other researchers (journals)
- Databases, etc.

Books - Books are one of the most traditional ways of collecting data. Today, there are books available for all topics you can think of. When carrying out research, all you have to do is look for a book on the topic being researched, then select from the available repository of books in that area. Books, when carefully chosen are an authentic source of authentic data and can be useful in preparing a literature review.

Published Sources - There are a variety of published sources available for different research topics. The authenticity of the data generated from these sources depends majorly on the writer and publishing company.

Unpublished Personal Sources - This may not be readily available and easily accessible compared to the published sources. They only become accessible if the researcher shares with another researcher who is not allowed to share it with a third party. For example, the product management team of an organization may need data on customer feedback to assess what customers think about their product and improvement suggestions. They will need to collect the data from the customer service department, which primarily collected the data to improve customer service.

Journal - Journals are gradually becoming more important than books these days when data collection is concerned. This is because journals are updated regularly with new publications

on a periodic basis, therefore giving to date information. Also, journals are usually more specific when it comes to research.

Newspapers - In most cases, the information passed through a newspaper is usually very reliable. Hence, making it one of the most authentic sources of collecting secondary data. The kind of data commonly shared in newspapers is usually more political, economic, and educational than scientific. Therefore, newspapers may not be the best source for scientific data collection.

Websites - The information shared on websites is mostly not regulated and as such may not be trusted compared to other sources. However, there are some regulated websites that only share authentic data and can be trusted by researchers. Most of these websites are usually government websites or private organizations that are paid, data collectors.

Blogs - Blogs are one of the most common online sources for data and may even be less authentic than websites. These days, practically everyone owns a blog, and a lot of people use these blogs to drive traffic to their website or make money through paid ads. Therefore, they cannot always be trusted. For example, a blogger may write good things about a product because he or she was paid to do so by the manufacturer even though these things are not true.
.

Diaries - They are personal records and as such rarely used for data collection by researchers. Also, diaries are usually personal, except for these days when people now share public diaries containing specific events in their life.

Government Records - Government records are a very important and authentic source of secondary data. They contain information useful in marketing, management, humanities, and social science research. Some of these records include; census data, health records, education institute records, etc. They are usually collected to aid proper planning, allocation of funds, and prioritizing of projects.
What are the Secondary Data Collection Tools?

Popular tools used to collect secondary data include; bots, devices, libraries, etc. In order to ease the data collection process from the sources of secondary data highlighted above, researchers use these important tools which are explained below.

Bots:
There are a lot of data online and it may be difficult for researchers to browse through all these data and find what they are actually looking for. In order to ease this process of data collection, programmers have created bots to do an automatic web scraping for relevant data.

Internet-Enabled Devices:
This could be a mobile phone, PC, or tablet that has access to an internet connection. They are used to access journals, books, blogs, etc. to collect secondary data.

Library:

This is a traditional secondary data collection tool for researchers. The library contains relevant materials for virtually all the research areas you can think of, and it is accessible to everyone. A researcher might decide to sit in the library for some time to collect secondary data or borrow the materials for some time and return when done collecting the required data.

Radio:

Radio stations are one of the secondary sources of data collection, and one needs radio to access them. The advent of technology has even made it possible to listen to the radio on mobile phones, deeming it unnecessary to get a radio.

Secondary data has various uses in research, business, and statistics. Researchers choose secondary data for different reasons, with some of it being due to price, availability, or even needs of the research. Although old, secondary data may be the only source of data in some cases. This may be due to the huge cost of performing research or due to its delegation to a particular body (e.g. national census).

Importance of Secondary Data

Secondary data can be less valid but its importance is still there. Sometimes it is difficult to obtain primary data; in these cases getting information from secondary sources is easier and possible. Sometimes primary data does not exist in such a situation one has to confine the research to secondary data. Sometimes primary data is present but the respondents are not willing to reveal it in such a case too secondary data can suffice. For example, if the research is on the psychology of transsexuals first it is difficult to find out transsexuals and second they may not be willing to give information you want for your research, so you can collect data from books or other published sources. A clear benefit of using secondary data is that much of the background work needed has already been carried out. For example, literature reviews, case studies might have been carried out, published texts and statistics could have been already used elsewhere, media promotion and personal contacts have also been utilized. This wealth of background work means that secondary data generally have a pre-established degree of validity and reliability which need not be re-examined by the researcher who is re-using such data. Furthermore, secondary data can also be helpful in the research design of subsequent primary research and can provide a baseline with which the collected primary data results can be compared to. Therefore, it is always wise to begin any research activity with a review of the secondary data.

Advantages of Using Secondary Data:
- No hassles of data collection.
- It is less expensive.
- The investigator is not personally responsible for the quality of data ('I didn't do it').

Disadvantages of Using Secondary Data:
- The data collected by the third party may not be a reliable party so the reliability and accuracy of data go down.

- Data collected in one location may not be suitable for the other one due to variable environmental factors.
- With the passage of time, the data becomes obsolete and very old.
- Secondary data collected can distort the results of the research. For using secondary data special care is required to amend or modify for use.
- Secondary data can also raise issues of authenticity and copyright.

Keeping in view the advantages and disadvantages of sources of data requirement of the research study and time factor, both sources of data i.e. primary and secondary data have been selected. 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.

Importance of Data Collection

There are a bunch of underlying reasons for collecting data, especially for a researcher. Walking you through them, here are a few reasons;

The integrity of the Research -A key reason for collecting data, be it through quantitative or qualitative methods is to ensure that the integrity of the research question is indeed maintained.

Reduce the likelihood of errors - The correct use of appropriate data collection of methods reduces the likelihood of errors consistent with the results.

Decision Making - To minimize the risk of errors in decision-making, it is important that accurate data is collected so that the researcher doesn't make uninformed decisions.

Save Cost and Time - Data collection saves the researcher time and funds that would otherwise be misspent without a deeper understanding of the topic or subject matter.

To support a need for a new idea, change, and/or innovation - To prove the need for a change in the norm or the introduction of new information that will be widely accepted, it is important to collect data as evidence to support these claims.

Research Data Types

Based on the techniques of collection, data can be classified into four categories:
1. Observational
2. Experimental
3. Simulation
4. Derived.

The way you collect the research data may have an impact on how you manage it. Data that is difficult or impossible to replace, e.g. the recording of an event at a certain time and location, necessitates additional backup measures to minimize the risk of data loss. Alternatively, if you

need to merge data points from many sources, you must follow best practices to avoid data corruption.

Observational Data:

Observational data is gathered by watching a person or thing doing something. It's gathered by methods, including human observation, open-ended surveys, or the use of a sensor or equipment to monitor and record data. The observation method lets the researcher look for nonverbal cues about feelings, check the pattern of interaction (who talks to whom), observe how participants interact, and also record the time they spend in different activities.

Some characteristics of observation are:

- Observations are guided by the research questions. Therefore the observations are conscious and planned. They differ from casual everyday observations of behavior which are often casual, selective, and inaccurate.
- Observations are systematically recorded, often using an observation checklist.
- Data are analyzed using both quantitative and qualitative data analysis methods.

Experimental Data:

Experimental data is collected by the researcher's active intervention during change of variables to produce and measure change or create a difference. This type of data is frequently repeatable, although doing so can be costly. Experimental data typically allows the researcher to determine a causal relationship and is typically projectable to a larger population.

Simulation Data:

Simulation data is created by employing computer test models to simulate the operation of a real-world process or system over time. The basic definition of data simulation is taking a large amount of data and using it to simulate or mirror real-world conditions to either predict a future instance, determine the best course of action or validate a model.

Derived/ compiled Data:

Derived data is created by transforming existing data points, which are often from disparate data sources, into new data using arithmetic formulas or aggregation. Derived data involves using existing data points, often from different data sources, to create new data through some sort of transformation, such as an arithmetic formula or aggregation.

SUMMARY

Data collecting is an unavoidable aspect of any research project. If the data obtained is valid and relevant, good conclusions can be drawn from a study with adequate analysis. Data collection can be used for creating improvements or bridging gaps. Data can be gathered in a variety of methods, including the internet, government collections, and archives. Individuals should also acquire useful data, otherwise, they risk wasting a lot of time and effort. One can select the appropriate media and source from which to gather specific data. Our society is highly dependent on data, which underscores the importance of collecting it. Accurate data collection

is necessary to make informed business decisions, ensure quality assurance, and keep research integrity.

The concept of data collection isn't a new one, as we'll see later, but the world has changed. There is far more data available today, and it exists in forms that were unheard of a century ago. The data collection process has had to change and grow with the times, keeping pace with technology.

REFERENCES

1. Ahuja R. (2015). Research methods. New Delhi: Rawat Publications.
2. Bailey K. D. (2008). Methods of Social Research (4. Ed.). Free Press. Pg. 194.
3. Bloor M. Frankland, J. Thomas, M., &Robson, K. (2001).Focus groups in social research. London: Sage Publications.
4. Brewer J. D. (2000). Ethnography. Buckingham, MK: Open University Press.
5. Cargan L. (2007). Doing Social Research. Plymouth, UK: Rowman and Littlefield Publishers.
6. Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Methods of data collection in qualitative research: interviews and focus groups. British Dental Journal, 204, 291-295.
7. Niraula SR. A review of the research process, data collection, and analysis. Insights Biol Med. 2019; 3: 001-006. https://doi.org/10.29328/journal.ibm.1001014.