ASSIGNMENT

16/mhs02/005

DATA COLLECTION

Whether it is business, marketing, humanities, physical sciences, social sciences, or other fields of study or discipline, data plays a very important role, serving as their respective starting points. That is why, in all of these processes that involve the usage of information and knowledge, one of the very first steps is data collection.

Data collection is described as the “process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer queries, stated research questions, test hypotheses, and evaluate outcomes.”

Depending on the discipline or field, the nature of the information being sought, and the objective or goal of users, the methods of data collection will vary. The approach to applying the methods may also vary, customized to suit the purpose and prevailing circumstances, without compromising the integrity, accuracy and reliability of the data.

There are two main types of data that users find themselves working with – and having to collect.

* **Quantitative Data.** These are data that deal with quantities, values or numbers, making them measurable. Thus, they are usually expressed in numerical form, such as length, size, amount, price, and even duration. The use of statistics to generate and subsequently analyze this type of data add credence or credibility to it, so that quantitative data is overall seen as more reliable and objective.
* **Qualitative Data.** These data, on the other hand, deals with quality, so that they are descriptive rather than numerical in nature. Unlike quantitative data, they are generally not measurable, and are only gained mostly through observation. Narratives often make use of adjectives and other descriptive words to refer to data on appearance, color, texture, and other qualities.

In most cases, these two data types are used as preferences in choosing the method or tool to be used in data collection. As a matter of fact, data collection methods are classified into two, and they are based on these types of data. Thus, we can safely say that there are two major classifications or categories of data collection methods: the quantitative data collection methods and the qualitative data collection methods.

**Data collection improves quality of expected results or output.**

Just as having data will improve decision-making and the quality of the decisions, it will also improve the quality of the results or output expected from any endeavor or activity. For example, a manufacturer will be able to produce high quality products after designing them using reliable data gathered. Consumers will also find the claims of the company about the product to be more reliable because they know it has been developed after conducting significant amount of research.

**QUANTITATIVE DATA COLLECTION METHOD**

Quantitative research methods describe and measure the level of occurrences on the basis of numbers and calculations. Moreover, the questions of “how many?” and “how often?” are often asked in quantitative studies. Accordingly, quantitative data collection methods are based on numbers and mathematical calculations.

Quantitative research can be described as ‘entailing the collection of numerical data and exhibiting the view of relationship between theory and research as deductive, a predilection for natural science approach, and as having an objectivist conception of social reality’. In other words, quantitative studies mainly examine relationships between numerically measured variables with the application of statistical techniques.

Quantitative data collection methods are based on random sampling and structured data collection instruments. Findings of quantitative studies are usually easy to present, summarize, compare and generalize. Quantitative studies on the contrary, are usually based on non-random sampling methods and use non-quantifiable data such as words, feelings, emotions e.t.c.

Data can be readily quantified and generated into numerical form, which will then be converted and processed into useful information mathematically. The result is often in the form of statistics that is meaningful and, therefore, useful. Unlike qualitative methods, these quantitative techniques usually make use of larger sample sizes because its measurable nature makes that possible and easier.

**Quantitative Surveys**

Unlike the open-ended questions asked in qualitative questionnaires, quantitative paper surveys pose closed questions, with the answer options provided. The respondents will only have to choose their answer among the choices provided on the questionnaire.

* 1. Similarly, these are ideal for use when surveying large numbers of respondents.
* 2.The standardized nature of questionnaires enable researchers to make generalizations out of the results.

3.This can be very limiting to the respondents, since it is possible that his actual answer to the question may not be in the list of options provided on the questionnaire.

* 4.While data analysis is still possible, it will be restricted by the lack of details.

**Interviews**

Personal one-on-one interviews may also be used for gathering quantitative data. In collecting quantitative data, the interview is more structured than when gathering qualitative data, comprised of a prepared set of standard questions.

These interviews can take the following forms:

* **Face-to-face interviews:** Much like when conducting interviews to gather qualitative data, this can also yield quantitative data when standard questions are asked.
  + 1.The face-to-face setup allows the researcher to make clarifications on any answer given by the interviewee.
  + 2.This can be quite a challenge when dealing with a large sample size or group of interviewees. If the plan is to interview everyone, it is bound to take a lot of time, not to mention a significant amount of money.
* **Telephone and/or online, web-based interviews.** Conducting interviews over the telephone is no longer a new concept. Rapidly rising to take the place of telephone interviews is the video interview via internet connection and web-based applications, such as Skype.
  + 1.The net for data collection may be cast wider, since there is no need to travel through distances to get the data. All it takes is to pick up the phone and dial a number, or connect to the internet and log on to Skype for a video call or video conference.
  + 2.Quality of the data may be questionable, especially in terms of impartiality. The net may be cast wide, but it will only be targeting a specific group of subjects: those with telephones and internet connections and are knowledgeable about using such technologies.
* **Computer-assisted interviews.** This is called CAPI, or Computer-Assisted Personal Interviewing where, in a face-to-face interview, the data obtained from the interviewee will be entered directly into a database through the use of a computer.
  + 1. The direct input of data saves a lot of time and other resources in converting them into information later on, because the processing will take place immediately after the data has been obtained from the source and entered into the database.

2. The use of computers, databases and related devices and technologies does not come cheap. It also requires a certain degree of being tech-savvy on the part of the data gatherer.

**Quantitative Observation**

This is straightforward enough. Data may be collected through systematic observation by, say, counting the number of users present and currently accessing services in a specific area, or the number of services being used within a designated vicinity.

When quantitative data is being sought, the approach is naturalistic observation, which mostly involves using the senses and keen observation skills to get data about the “what”, and not really about the “why” and “how”.

* 1. It is a quite simple way of collecting data, and not as expensive as the other methods.
* 2.The problem is that senses are not infallible. Unwittingly, the observer may have an unconscious grasp on his senses, and how they perceive situations and people around. Bias on the part of the observer is very possible.

**Experiments**

Have you ever wondered where clinical trials fall? They are considered to be a form of experiment, and are quantitative in nature. These methods involve manipulation of an independent variable, while maintaining varying degrees of control over other variables, most likely the dependent ones. Usually, this is employed to obtain data that will be used later on for analysis of relationships and correlations.

Quantitative researches often make use of experiments to gather data, and the types of experiments are:

* **Laboratory experiments.** This is your typical scientific experiment setup, taking place within a confined, closed and controlled environment (the laboratory), with the data collector being able to have strict control over all the variables. This level of control also implies that he can fully and deliberately manipulate the independent variable.
* **Field experiments.** This takes place in a natural environment, “on field” where, although the data collector may not be in full control of the variables, he is still able to do so up to a certain extent. Manipulation is still possible, although not as deliberate as in a laboratory setting.
* **Natural experiments.** This time, the data collector has no control over the independent variable whatsoever, which means it cannot be manipulated. Therefore, what can only be done is to gather data by letting the independent variable occur naturally, and observe its effects.

You can probably name several other data collection methods, but the ones discussed are the most commonly used approaches. At the end of the day, the choice of a collection method is only 50% of the whole process. The correct usage of these methods will also have a bearing on the quality and integrity of the data being sought.

**Qualitative Data Collection Methods**

Exploratory in nature, these methods are mainly concerned at gaining insights and understanding on underlying reasons and motivations, so they tend to dig deeper. Since they cannot be quantified, measurability becomes an issue. This lack of measurability leads to the preference for methods or tools that are largely unstructured or, in some cases, maybe structured but only to a very small, limited extent.

Generally, qualitative methods are time-consuming and expensive to conduct, and so researchers try to lower the costs incurred by decreasing the sample size or number of respondents.

Qualitative data collection methods are exploratory in nature and are mainly concerned with gaining insights and understanding on underlying reasons and motivations. Qualitative data collection methods emerged after it has become known that traditional quantitative data collection methods were unable to express human feelings and emotions.

Monette et al (2010) credit qualitative methods with the acknowledgement of abstraction and generalization. Polonsky and Waller (2011) categorize vision, images, forms and structures in various media, as well as spoken and printed word and recorded sound into qualitative data collection methods.

It is noted that “qualitative methods are often regarded as providing rich data about real life people and situations and being more able to make sense of behaviour and to understand behaviour within its wider context. However, qualitative research is often criticized for lacking generalizability, being too reliant on the subjective interpretations of researchers and being incapable of replication by subsequent researchers.

Popular qualitative data collection methods used in business studies include interviews, focus groups, observation and action research. Moreover, grounded theory and document analysis can be also used as data collection method in qualitative studies.

**Face-to-Face Personal Interviews**

This is considered to be the most common data collection instrument for qualitative research, primarily because of its personal approach. The interviewer will collect data directly from the subject (the interviewee), on a one-on-one and face-to-face interaction. This is ideal for when data to be obtained must be highly personalized.

The interview may be informal and unstructured – conversational, even – as if taking place between two casual to close friends. The questions asked are mostly unplanned and spontaneous, with the interviewer letting the flow of the interview dictate the next questions to be asked.

However, if the interviewer still wants the data to be standardized to a certain extent for easier analysis, he could conduct a semi-structured interview where he asks the same series of open-ended questions to all the respondents. But if they let the subject choose her answer from a set of options, what just took place is a closed, structured and fixed-response interview.

* 1.This allows the interviewer to probe further, by asking follow-up questions and getting more information in the process.
* 2. The data will be highly personalized (particularly when using the informal approach).
* 3. This method is subject to certain limitations, such as language barriers, cultural differences, and geographical distances.
* 4. The person conducting the interview must have very good interviewing skills in order to elicit responses.

**Qualitative Surveys**

* **Paper surveys or questionnaires.** Questionnaires often utilize a structure comprised of short questions and, in the case of qualitative questionnaires, they are usually open-ended, with the respondents asked to provide detailed answers, in their own words. It’s almost like answering essay questions.
  + 1. Since questionnaires are designed to collect standardized data, they are ideal for use in large populations or sample sizes of respondents.
  + 2. The high amount of detail provided will aid analysis of data.
  + 3.On the other hand, the large number of respondents (and data), combined with the high level and amount of detail provided in the answers, will make data analysis quite tedious and time-consuming.
* **Web-based questionnaires.** This is basically a web-based or internet-based survey, involving a questionnaire uploaded to a site, where the respondents will log into and accomplish electronically. Instead of a paper and a pen, they will be using a computer screen and the mouse.
  + 1.Data collection is definitely quicker. This is often due to the questions being shorter, requiring less detail than in, say, a personal interview or a paper questionnaire.
  + 2. It is also uncomplicated, since the respondents can be invited to answer the questionnaire by simply sending them an email containing the URL of the site where the online questionnaire is available for answering.
  + 3.There is a limitation on the respondents, since the only ones to be able to answer are those who own a computer, have internet connection, and know their way around answering online surveys.

4. The lesser amount of detail provided means the researcher may end up with mostly surface data, and no depth or meaning, especially when the data is processed.

**Focus Groups**

Focus groups method is basically an interview method, but done in a group discussion setting. When the object of the data is behaviors and attitudes, particularly in social situations, and resources for one-on-one interviews are limited, using the focus group approach is highly recommended. Ideally, the focus group should have at least 3 people and a moderator to around 10 to 13 people maximum, plus a moderator.

Depending on the data being sought, the members of the group should have something in common. For example, a researcher conducting a study on the recovery of married mothers from alcoholism will choose women who are (1) married, (2) have kids, and (3) recovering alcoholics. Other parameters such as the age, employment status, and income bracket do not have to be similar across the members of the focus group.

The topic that data will be collected about will be presented to the group, and the moderator will open the floor for a debate.

* 1.There may be a small group of respondents, but the setup or framework of data being delivered and shared makes it possible to come up with a wide variety of answers.
* 2. The data collector may also get highly detailed and descriptive data by using a focus group.
* 3. Much of the success of the discussion within the focus group lies in the hands of the moderator. He must be highly capable and experienced in controlling these types of interactions.

**Documental Revision**

This method involves the use of previously existing and reliable documents and other sources of information as a source of data to be used in a new research or investigation. This is likened to how the data collector will go to a library and go over the books and other references for information relevant to what he is currently researching on.

* 1. The researcher will gain better understanding of the field or subject being looked into, thanks to the reliable and high quality documents used as data sources.
* 2. Taking a look into other documents or researches as a source will provide a glimpse of the subject being looked into from different perspectives or points of view, allowing comparisons and contrasts to be made.
* 3. Unfortunately, this relies heavily on the quality of the document that will be used, and the ability of the data collector to choose the right and reliable documents. If he chooses wrong, then the quality of the data he will collect later on will be compromised.

**Observation**

In this method, the researcher takes a participatory stance, immersing himself in the setting where his respondents are, and generally taking a look at everything, while taking down notes.

Aside from note-taking, other documentation methods may be used, such as video and audio recording, photography, and the use of tangible items such as artifacts, mementoes, and other tools.

* 1. The participatory nature may lead to the researcher getting more reliable information.
* 2. Data is more reliable and representative of what is actually happening, since they took place and were observed under normal circumstances.
* 3. The participation may end up influencing the opinions and attitudes of the researcher, so he will end up having difficulty being objective and impartial as soon as the data he is looking for comes in.
* 4. Validity may arise due to the risk that the researcher’s participation may have an impact on the naturalness of the setting. The observed may become reactive to the idea of being watched and observed. If he planned to observe recovering alcoholic mothers in their natural environment (e.g. at their homes with their kids), their presence may cause the subjects to react differently, knowing that they are being observed. This may lead to the results becoming impaired.

**Longitudinal studies**

This is a research or data collection method that is performed repeatedly, on the same data sources, over an extended period of time. It is an observational research method that could even cover a span of years and, in some cases, even decades. The goal is to find correlations through an empirical or observational study of subjects with a common trait or characteristic.

An example of this is the Terming study of the gifted conducted by Lewis Terman at Stanford University. The study aimed to gather data on the characteristics of gifted children – and how they grow and develop – over their lifetime. Terman started in 1921, and it extended over the lifespan of the subjects, more than 1,500 boys and girls aged 3 to 19 years old, and with IQs higher than 135. To this day, this study is the world’s “oldest and longest-running” longitudinal study.

* 1. This is ideal when seeking data meant to establish a variable’s pattern over a period of time, particularly over an extended period of time.
* 2. As a method to find correlations, it is effective in finding connections and relationships of cause and effect.
* 3. The long period may become a setback, considering how the probability of the subjects at the beginning of the research will still be complete 10, 20, or 30 years down the road is very low.
* 4. Over the extended period, attitudes and opinions of the subjects are likely to change, which can lead to the dilution of data, reducing their reliability in the process.

**Case Studies**

In this qualitative method, data is gathered by taking a close look and an in-depth analysis of a “case study” or “case studies” – the unit or units of research that may be an individual, a group of individuals, or an entire organization. This methodology’s versatility is demonstrated in how it can be used to analyze both simple and complex subjects.

However, the strength of a case study as a data collection method is attributed to how it utilizes other data collection methods, and captures more variables than when a single methodology is used. In analyzing the case study, the researcher may employ other methods such as interviewing, floating questionnaires, or conducting group discussions in order to gather data.

* 1. It is flexible and versatile, analyzing both simple and complex units and occurrence, even over a long period of time.
* 2. Case studies provide in-depth and detailed information, thanks to how it captures as many variables as it can.

3. Reliability of the data may be put at risk when the case study or studies chosen are not representative of the sample or population.

The main sources and procedures associated with the most popular qualitative methods are presented on the table below as proposed by Yamagata-Lynch (2010).

|  |  |  |
| --- | --- | --- |
| **Methodology** | **Sources** | **Procedure** |
| Document analysis | Reports, newsletters, publications | Read all materials and documented and descriptive statistics related to the research issue |
| Interviews | Primary participants Secondary participants | Tape recorded semi-structured interviews, then transcribed the interviews for the participants to review |
| Observations | Observed participants’ interactions | Took notes and videotaped the observations |
| Exit interviews | Primary participants  Secondary participants | Presented findings to participants during individual or group interview sessions |

The table below illustrates the difference between qualitative data collection methods and qualitative research:

|  |  |  |  |
| --- | --- | --- | --- |
|  | | **Quantitative** | **Qualitative** |
| Requirement | Question | Hypothesis | Interest |
| Method | Control and randomization | Curiosity and reflexivity |
| Data collection | Response | Viewpoint |
| Outcome | Dependent variable | Account |
| Ideal | Data | Numerical | Textual |
| Sample size | Large (power) | Small (saturation) |
| Context | Eliminated | Highlighted |
| Analysis | Rejection on null | Synthesis |