

## Data Collection: What you need to know and why

**Author: Tatenda Emma Matika . December 2020**

Data is an important tool for any organization that aims to provide better services to customers and to make sure employees are happy at work. If used well, it gives a competitive advantage to an organization. But how can organizations collect such data?

When I hear the term data collection, the first thing that I think of is questionnaires and online surveys. I have carried out an online survey using google forms before. It sounds easy when thinking about it but the actual process is not. Before collecting data, one should know the following:

- Why they are collecting data
- What type of data they want
- The type of respondents
- What tool should be used to collect the data
- How long the data collection should last
- How the data is going to be used
- What tools will be used to process the data
- What tools will be used to report the results

To make things clear, we should start by defining what data collection is. According to [Wikipedia](#), Data Collection is a process of gathering and measuring information about certain variables in established systems which allows one to answer questions and derive meaning from the data.

### **Why data is collected**

We cannot just collect data for the sake of collecting it. Identifying the reason behind data collection makes it easy to make decisions on other parts of the collection process. The main reasons why organizations collect data can fall under the following:

- To understand people of interest to the organization, e.g. employees or customers
- To identify areas to improve e.g. remuneration or product size or colour
- To predict the future e.g. predicting employee turnover
- To make decisions e.g. which customers to target, which people to recruit

Identifying why we are collecting data indirectly identifies who is going to provide the data. If the data is going to be used for determining employee engagement, we know that it is going to be provided by employees. So, we have an idea of the tool we can use for collection. We cannot use an online survey if we know that our employees do not have enough resources to access the internet.

### **The type of data to be collected**

Data to be collected can be categorized into **Primary data** and **Secondary data**. Primary data is data collected by the user for the user. For example, when a business collects data from their employees in order to understand their needs, this is usually primary data. Secondary data is data collected by a one party to be used by other parties, i.e. the one who collects is not the user of the data. This can also be data from other organizations. For example, when we conduct salary surveys, we use data recorded by other companies.

With primary data, we have control over what kind of information we want from respondents. This helps when structuring the questions to ask. The data can be either **Qualitative** or **Quantitative**. Qualitative data is descriptive and mostly contains people's opinions for example product reviews. It is difficult to measure and open-ended.

Quantitative data is numeric and easy to measure for example the number of customers who walked into a shop on a certain day. It is easy to analyse statistically and therefore regarded as more reliable compared to qualitative. It is always good to accompany quantitative data with qualitative data so that the qualitative sheds more light on the quantitative. A practical example is when carrying out employee engagement surveys. Allowing employees to enter other comments can give an insight on why they disagree with certain issues mentioned in the quantitative section.

### **The method of collection**

The method depends on most of the factors listed above. The type of data to be collected, type of respondents, timeframe and so on all determine the method of collection. The main methods used for data collection are Interviews, Questionnaires, Observation and Focus Groups.

### **Interviews**

Interviews can be in-person (physically or virtually) or over the phone. They can also be structured, unstructured or semi-structured. For structured interviews, the questions are written down before the interview. The interviewer then asks these questions during the interview and collects responses. For unstructured interviews, there are no pre-set questions, the interviewer and interviewee discuss and responses are collected during the interview. A semi-structured interview is a combination of the two methods.

Interviews are useful when the number of respondents is very small and the respondents have no problem with their identity being disclosed. They are also useful when the respondents do not have time to complete questionnaires or when it is likely that more information can be obtained through discussion rather than using pre-written questions.

### **Tools used for interviews**

Interviews are carried out in a short space of time so it might be difficult to collect all the data. Some of the tools that can be used are Video cameras, Audio recorders or Recordings of online meetings. After the interview, the interviewer can then replay the recordings and capture the data.

## **Questionnaires**

These are the most common data collection method. The person collecting data sets questions and then sends them to the respondents. They can be online surveys or physical papers. They are useful when respondents do not want their identities to be disclosed and when there is a large number of respondents. However, it is difficult to make sure that all of the targeted respondents give feedback. Also, some do not respond honestly as they fear that their identities may be discovered.

### **Tools used for questionnaires**

Online tools such as Google Forms, Survey Monkey and others can be used. Physical papers can be used for respondents who do not have access to the internet or for those who are not knowledgeable about such tools.

## **Observation**

This can be done by either going through a process then capturing data about it or by observing other people going through the process. This is first-hand information and is likely to be reliable. However, a problem can arise when the subjects being observed do not act like they would normally do if they know they are being observed. An example of observation is when determining employee engagement - an external person can go to a workplace and observe employees as they do their work. Another problem that could also arise is that the data could be biased according to the observer's way of thinking.

### **Tools used for observation**

Cameras, for example CCTV cameras can be used to observe environments. The data collected can then be captured manually by the data collector.

## **Focus Groups**

These are interviews but for people in the same category or demographic.

## **Problems Faced During Data Collection**

When collecting data, it is inevitable to encounter some challenges. Some organizations' policies do not allow sharing of data and this might affect the activity for which the data has to be collected. Sometimes respondents are worried about their privacy and will not provide honest responses. Also, the data collection might go well but the responses will need to be cleaned, especially for open ended questions. This process usually takes time and needs the person performing it to be very careful.

*Tatenda Emma Matika is a Business Analytics Trainee at Industrial Psychology Consultants (Pvt) Ltd a management and human resources consulting firm.*

## **References**

[https://en.wikipedia.org/wiki/Data\\_collection](https://en.wikipedia.org/wiki/Data_collection)

<https://www.lotame.com/what-are-the-methods-of-data-collection/>

Kongmany Chaleunvong, (2009), Data Collection Techniques.

<https://www.scribbr.com/methodology/data-collection/>

[https://ori.hhs.gov/education/products/n\\_illinois\\_u/datamanagement/dctopic.html](https://ori.hhs.gov/education/products/n_illinois_u/datamanagement/dctopic.html)

<https://thehumancapitalhub.com/articles/Data-Collection-What-You-Need-To-Know-And-Why>