

# Data collection Report



# Introduction

---



## What is Data collection?

Before we define what is data collection, it's essential to ask the question, "What is data?" The abridged answer is, data is various kinds of information formatted in a particular way.

Therefore, data collection is the process of gathering, measuring, and analyzing accurate data from a variety of relevant sources to find answers to research problems, answer questions, evaluate outcomes, and forecast trends and probabilities.

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

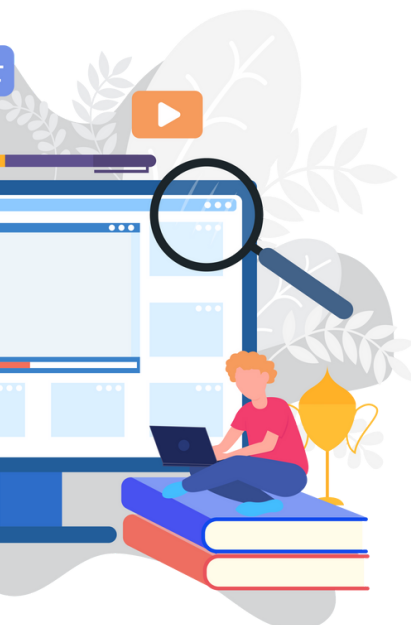


## Why Do We Need Data Collection?

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.

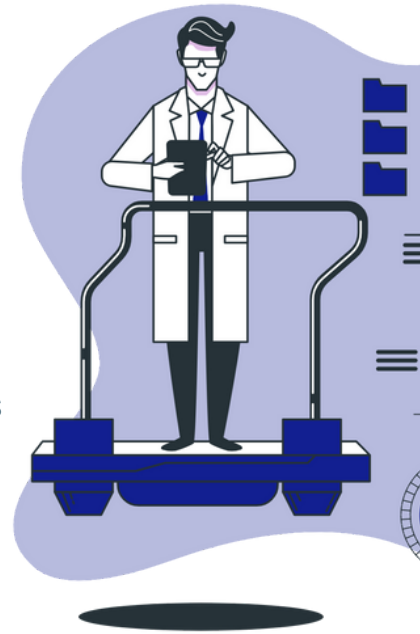
Whether you're in the world of academia, trying to conduct research, or part of the commercial sector, thinking of how to promote a new product, you need data collection to help you make better choices.



Similarities in sequences are scored in a matrix, and the algorithm allows for the detection of gaps in sequence alignment.

Another goal of bioinformatics is the extension of experimental data by predictions. A fundamental goal of computational biology is the prediction of protein structure from an amino acid sequence.

The spontaneous folding of proteins shows that this should be possible. Progress in the development of methods to predict protein folding is measured by biennial Critical Assessment of Structure Prediction (CASP) programs, which involve blind tests of structure prediction methods.



## What Are the Different Methods of Data Collection?

Carefully consider what method you will use to gather data that helps you directly answer your research questions.

METHOD	WHEN TO USE	HOW TO COLLECT DATA
Experiment	To test a causal relationship.	Manipulate variables and measure their effects on others.
Survey	To understand the general characteristics or opinions of a group of people.	Distribute a list of questions to a sample online, in person or over-the-phone.
focus group	To gain an in-depth understanding of perceptions or opinions on a topic.	Verbally ask participants open-ended questions in individual interviews or focus group discussions.
Observation	To understand something in its natural setting.	Measure or survey a sample without trying to affect them.
Ethnography	To study the culture of a community or organization first-hand.	Join and participate in a community and record your observations and reflections.
Archival research	To understand current or historical events, conditions or practices.	Access manuscripts, documents or records from libraries, depositories or the internet.

Data collection breaks down into two methods,  
The two methods are:

## Different Methods of Data Collection

### PRIMARY

As the name implies, this is original, first-hand data collected by the data researchers. This process is the initial information gathering step, performed before anyone carries out any further or related research. Primary data results are highly accurate provided the researcher collects the information. However, there's a downside, as first-hand research is potentially time-consuming and expensive.

### SECONDARY

Secondary data is second-hand data collected by other parties and already having undergone statistical analysis. This data is either information that the researcher has tasked other people to collect or information the researcher has looked up. Simply put, it's second-hand information. Although it's easier and cheaper to obtain than primary information, secondary information raises concerns regarding accuracy and authenticity.

## Specific Data Collection Techniques

### 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 Data Gathering

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.

### Delphi Technique

In the realm of data collection, researchers use the Delphi technique by gathering information from a panel of experts. Each expert answers questions in their field of specialty, and the replies are consolidated into a single opinion.

Unlike primary data collection, there are no specific collection methods. Instead, since the information has already been collected, the researcher consults various data sources, such as:

Financial Statements

Sales Reports

Deal Feedback

Business Magazines

Customer Personal Information

The internet

# Data Collection Tools

## 1 Word Association

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

## 2 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.

## 3 Online Surveys

These surveys are easy to accomplish, but some users may be unwilling to answer truthfully, if at all.



## The Importance of Ensuring Accurate and Appropriate Data Collection :-

Among the effects of data collection done incorrectly, include the following:

- Erroneous conclusions that squander resources.
- Decisions that compromise public policy.
- Incapacity to correctly respond to research inquiries.
- Deceiving other researchers into pursuing futile research avenues.
- The study's inability to be replicated and validated.

## What are Common Challenges in Data Collection?

There are some prevalent challenges faced while collecting data, let us explore a few of them to understand them better and avoid them.



## 1 Data Quality Issues

The main threat to the broad and successful application of machine learning is poor data quality. Data quality must be your top priority if you want to make technologies like machine learning work for you.

## 2 Inconsistent Data

When working with various data sources, it's conceivable that the same information will have discrepancies between sources. The differences could be in formats, units, or occasionally spellings.

The introduction of inconsistent data might also occur during firm mergers or relocations.

Inconsistencies in data have a tendency to accumulate and reduce the value of data if they are not continually resolved.

## 3 Data Downtime

Customer complaints and subpar analytical outcomes are only two ways that this data unavailability can have a significant impact on businesses.