

Data Analysis

What is data analysis?

Sefam Agbenyega | IIPGH Coding Class – Data Analytics | 8th June, 2023

# 

# What is Data Analysis?

Data analysis is the process of inspecting, cleansing, transforming, and modelling data to discover useful information, inform conclusions, and support decision-making.

Data Analysis is also the process of systematically applying statistical and/or logical techniques to describe and illustrate, condense and recap, and evaluate data.

While data analysis in qualitative research can include statistical procedures, many times analysis becomes an ongoing iterative process where data is continuously collected and analyzed almost simultaneously. Indeed, researchers generally analyze for patterns in observations throughout the entire data collection phase (Savenye, Robinson, 2004). The form of the analysis is determined by the specific qualitative approach taken (field study, ethnography content analysis, oral history, biography, unobtrusive research) and the form of the data (field notes, documents, audiotape, videotape**).**

An essential component of ensuring data integrity is the accurate and appropriate analysis of research findings. Improper statistical analyses distort scientific findings, mislead casual readers (Shepard, 2002), and may negatively influence the public perception of research. Integrity issues are just as relevant to the analysis of non-statistical data as well. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, and is used in different business, science, and social science domains. In today's business world, data analysis plays a role in making decisions more scientific and helping businesses operate more effectively.

# Types Of Data Analysis

Several types of data analysis techniques exist based on business and technology. However, the major data analysis methods are:

* **Descriptive Analysis**

Descriptive analysis involves summarizing and describing the main features of a dataset. It focuses on organizing and presenting the data in a meaningful way, often using measures such as mean, median, mode, and standard deviation. It provides an overview of the data and helps identify patterns or trends.

* **Inferential Analysis**

Inferential analysis aims to make inferences or predictions about a larger population based on sample data. It involves applying statistical techniques such as hypothesis testing, confidence intervals, and regression analysis. It helps generalize findings from a sample to a larger population.

* **Exploratory Data Analysis (EDA)**

EDA focuses on exploring and understanding the data without preconceived hypotheses. It involves visualizations, summary statistics, and data profiling techniques to uncover patterns, relationships, and interesting features. It helps generate hypotheses for further analysis.

* **Diagnostic Analysis**

Diagnostic analysis is a type of data analysis that focuses on identifying and understanding the causes of problems or issues. It involves examining data to diagnose the root causes of a particular situation or to determine why something is not working as expected. In diagnostic analysis, the emphasis is on investigating and understanding the underlying factors or variables that contribute to a problem or outcome. It aims to answer questions like "What went wrong?" or "Why did this happen?"

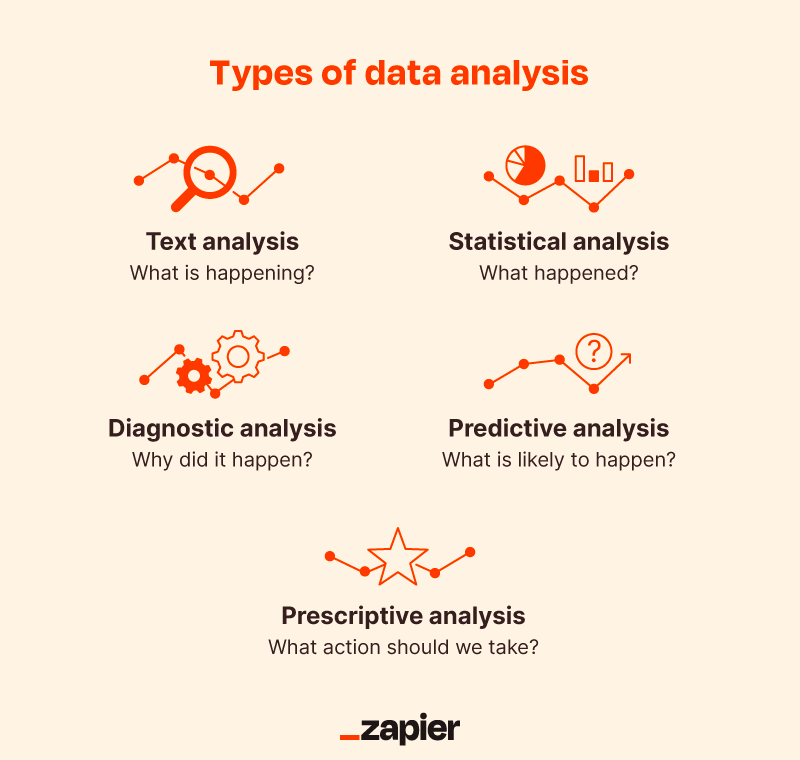
* **Predictive Analysis**

Predictive analysis involves using historical data to make predictions or forecasts about future outcomes. It utilizes statistical modelling techniques, machine learning algorithms, and time series analysis to identify patterns and build predictive models. It is often used for forecasting sales, predicting customer behaviour, or estimating risk.

* **Prescriptive Analysis**

Prescriptive analysis goes beyond predictive analysis by recommending actions or decisions based on the predictions. It combines historical data, optimization algorithms, and business rules to provide actionable insights and optimize outcomes. It helps in decision-making and resource allocation.

* **Statistical Analysis**

Statistical Analysis shows “What happen?” by using past data in the form of dashboards. Statistical Analysis includes the collection, analysis, interpretation, presentation, and modelling of data. It analyses a set of data or a sample of data. There are two categories of this type of Analysis – Descriptive Analysis and Inferential Analysis.

# Processes Of Data Analysis

Answering the question “What is data analysis” is only the first step. Now we will look at how it’s performed. The process of data analysis, or alternately, data analysis steps, involves gathering all the information, processing it, exploring the data, and using it to find patterns and other insights. The process of data analysis consists of:

* **Data Requirement Gathering**

Ask yourself why you’re doing this analysis, what type of data you want to use, and what data you plan to analyze.

* **Data Collection**

Guided by your identified requirements, it’s time to collect the data from your sources. Sources include case studies, surveys, interviews, questionnaires, direct observation, and focus groups. Make sure to organize the collected data for analysis.

* **Data Cleaning**

Not all of the data you collect will be useful, so it’s time to clean it up. This process is where you remove white spaces, duplicate records, and basic errors. Data cleaning is mandatory before sending the information for analysis.

* **Data Analysis**

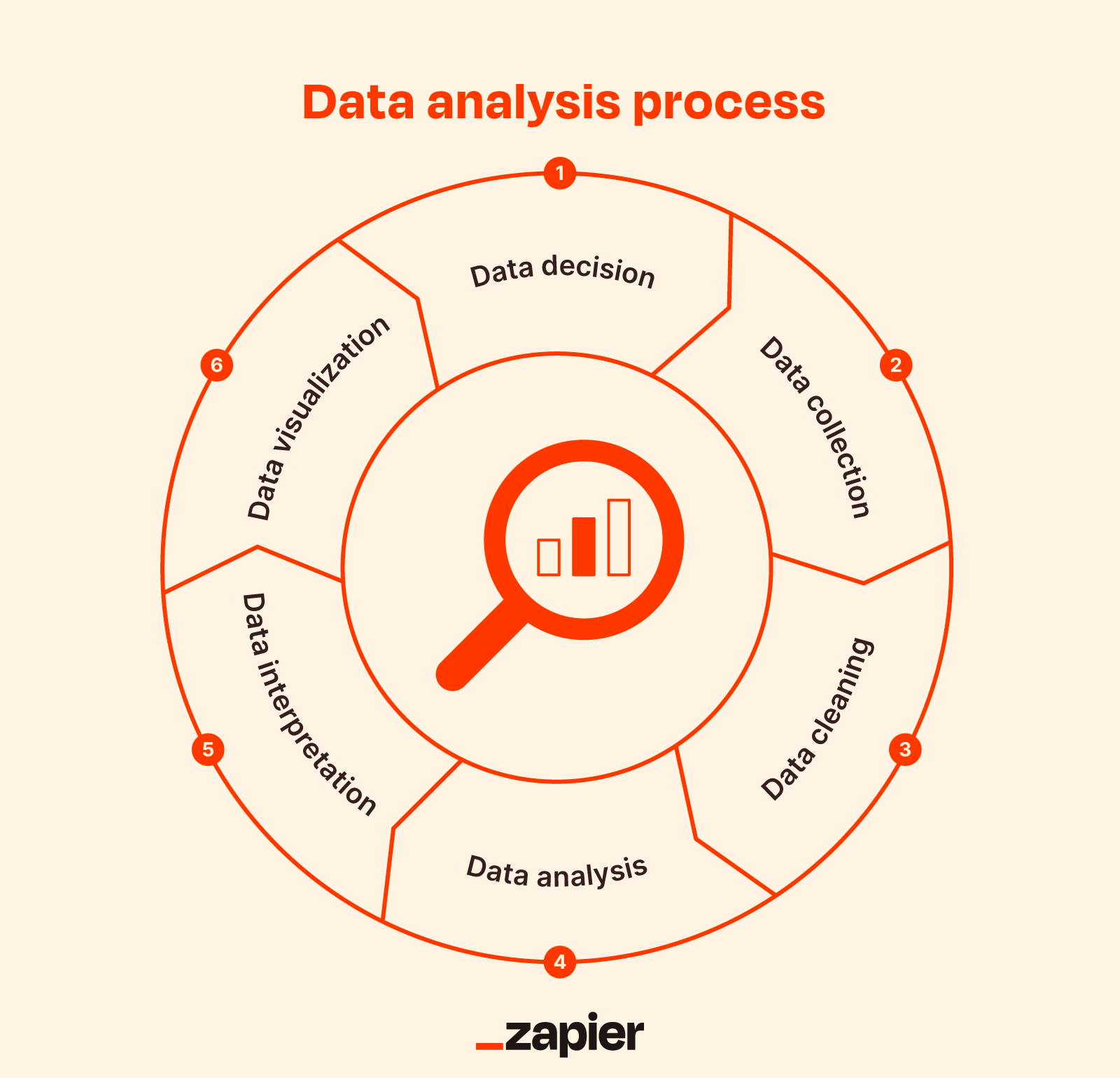
Here is where you use data analysis software and other tools to help you interpret and understand the data and arrive at conclusions. Data analysis tools include Excel, Python, R, Looker, Rapid Miner, Chartio, Metabase, Redash, and Microsoft Power BI.

* **Data Interpretation**

Now that you have your results, you need to interpret them and come up with the best courses of action based on your findings.

* **Data Visualization**

Data visualization is a fancy way of saying, “graphically show your information in a way that people can read and understand it.” You can use charts, graphs, maps, bullet points, or a host of other methods. Visualization helps you derive valuable insights by helping you compare datasets and observe relationships.



# Data Analysis Methods

Some professionals use the terms “data analysis methods” and “data analysis techniques” interchangeably. To further complicate matters, sometimes people throw in the previously discussed “data analysis types” into the fray as well! Our hope here is to establish a distinction between what kinds of data analysis exist, and the various ways it’s used.

Although there are many data analysis methods available, they all fall into one of two primary types: qualitative analysis and quantitative analysis.

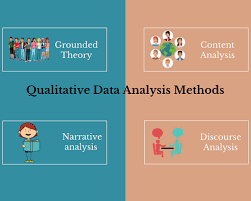
* **Qualitative Data Analysis**

The qualitative data analysis method derives data via words, symbols, pictures, and observations. This method doesn’t use statistics. The most common qualitative methods include:

* Content Analysis, for analyzing behavioral and verbal data.
* Narrative Analysis, for working with data culled from interviews, diaries, surveys.
* Grounded Theory, is for developing causal explanations of a given event by studying and extrapolating from one or more past cases.
* **Quantitative Data Analysis**

Also known as statistical data analysis methods collect raw data and process it into numerical data. Quantitative analysis methods include:

* Hypothesis Testing, is for assessing the truth of a given hypothesis or theory for a data set or demographic.
* Mean, or average determines a subject’s overall trend by dividing the sum of a list of numbers by the number of items on the list.
* Sample Size Determination uses a small sample taken from a larger group of people and analyzed. The results gained are considered representative of the entire body.

We can further expand our discussion of data analysis by showing various techniques, broken down by different concepts and tools.

# Top Data Analysis Tools

Here is a list of the top seven data analysis tools in terms of popularity, learning, and performance.

* **Tableau Public**
* **R Programming**
* **Python**
* **Apache Spark**
* **SAS**
* **Excel**
* **RapidMiner**

****

# Uses Of Data Analysis

Data analysis is indeed a versatile tool used in various sectors for extracting insights, making informed decisions, and solving complex problems. Here are a few sectors where data analysis finds widespread application:

* **Science and Research**

Data analysis is crucial in scientific research. Scientists collect data during experiments or observations and then analyze it to draw meaningful conclusions. It helps them identify patterns, test hypotheses, and make discoveries in fields like biology, physics, chemistry, and environmental science.

* **Business and Marketing**

Data analysis plays a significant role in business decision-making. Companies collect data on customer behaviour, sales, and market trends. By analyzing this data, businesses can understand customer preferences, target their marketing efforts, optimize pricing strategies, and make informed decisions to improve their products or services.

* **Healthcare and Medicine**

Data analysis is vital in healthcare to improve patient outcomes and develop new treatments. Medical researchers analyze large datasets to identify risk factors, understand disease patterns, and evaluate the effectiveness of medications or therapies. Data analysis also helps healthcare providers optimize patient care, predict disease outbreaks, and identify trends for public health interventions.

* **Finance and Economics**

Data analysis is extensively used in finance and economics to make informed financial decisions and predict market trends. Analysts examine financial data, stock market trends, and economic indicators to identify investment opportunities, assess risks, and forecast market behaviour.

* **Government and Public Policy**

Data analysis supports evidence-based decision-making in government and public policy by analyzing demographic trends, economic indicators, social welfare programs, crime patterns, and environmental data.

* **Transportation and Logistics**

Data analysis assists in optimizing transportation routes, predicting demand, improving logistics operations, and enhancing overall efficiency in areas such as shipping, logistics, and public transportation.

* **Sports Analytics**

Data analysis is increasingly used in sports to evaluate player performance, team strategies, match outcomes, and fan engagement, aiding in decision-making for coaches, managers, and sports organizations.

These are just a few examples, but data analysis is applicable in almost every field where data is collected. It helps us make sense of information, uncover insights, and make informed decisions based on evidence rather than guesswork.

# 

# Importance Of Data Analysis

Here is a list of reasons why data analysis is crucial to doing business today.

* **Better Customer Targeting**

You don’t want to waste your business’s precious time, resources, and money putting together advertising campaigns targeted at demographic groups that have little to no interest in the goods and services you offer. Data analysis helps you see where you should be focusing your advertising and marketing efforts.

* **You Will Know Your Target Customers Better**

Data analysis tracks how well your products and campaigns are performing within your target demographic. Through data analysis, your business can get a better idea of your target audience’s spending habits, disposable income, and most likely areas of interest. This data helps businesses set prices, determine the length of ad campaigns, and even help project the number of goods needed.

* **Reduce Operational Costs**

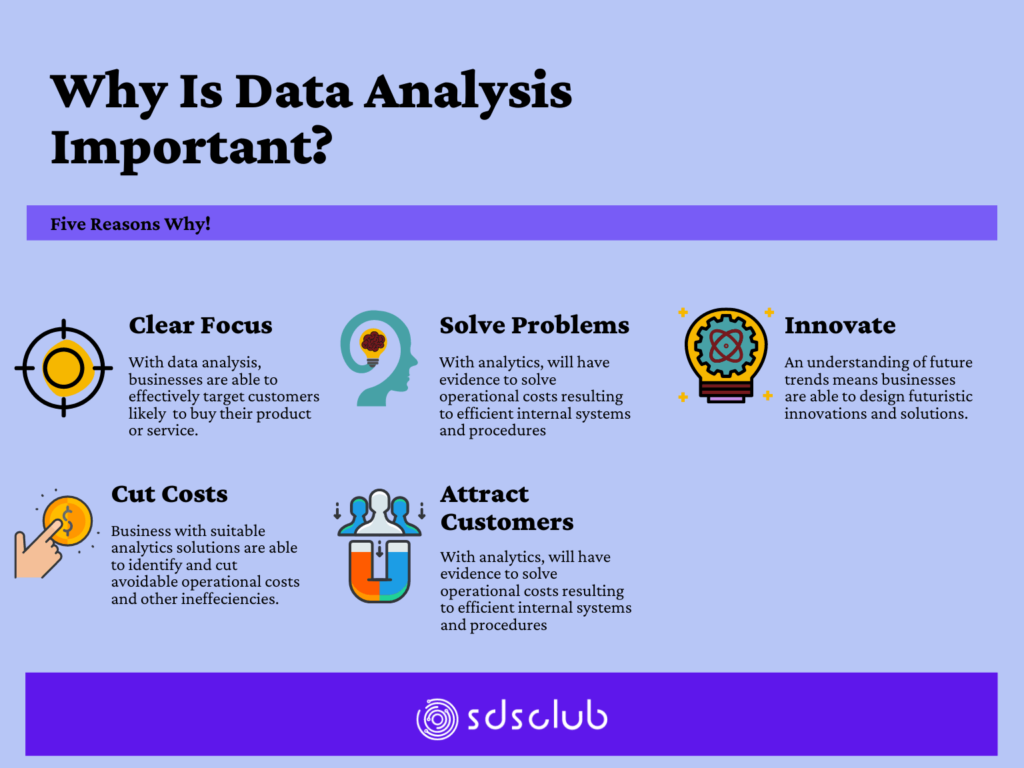
Data analysis shows you which areas in your business need more resources and money, and which areas are not producing and thus should be scaled back or eliminated outright.

* **Better Problem-Solving Methods**

Informed decisions are more likely to be successful decisions. Data provides businesses with information. You can see where this progression is leading. Data analysis helps businesses make the right choices and avoid costly pitfalls.

* **You Get More Accurate Data**

If you want to make informed decisions, you need data, but there’s more to it. The data in question must be accurate. Data analysis helps businesses acquire relevant, accurate information, suitable for developing future marketing strategies, business plans, and realigning the company’s vision or mission.



# What Is the Importance of Data Analysis in Research?

A huge part of a researcher’s job is to sift through data. That is literally the definition of “research.” However, today’s Information Age routinely produces a tidal wave of data, enough to overwhelm even the most dedicated researcher. From a bird’s eye view, data analysis:

1. Plays a key role in distilling this information into a more accurate and relevant form, making it easier for researchers to do their job.

2. Provides researchers with a vast selection of different tools, such as descriptive statistics, inferential analysis, and quantitative analysis.

3. Offers researchers better data and better ways to analyze and study said data.



# Examples

### Scenarios in which Data Analysis is Used

Here are some scenarios in which data analysis is commonly used:

* **Supply Chain Optimization**

Analyzing supply chain data, including inventory levels, transportation costs, and supplier performance, to optimize logistics and streamline operations, leading to cost savings and improved efficiency.

* **Social Media Analytics**

Analyzing social media data to understand customer behaviour, engagement levels, and sentiment towards a brand, helping companies refine their social media strategies and improve customer satisfaction.

* **Energy Consumption Analysis**

Analyzing energy consumption data to identify patterns, peak demand periods, and opportunities for energy efficiency improvements, enabling businesses and households to reduce energy costs and environmental impact.

* **Predictive Maintenance**

Analyzing equipment sensor data to detect patterns and anomalies that can predict potential failures, allowing businesses to schedule maintenance proactively, minimize downtime, and reduce maintenance costs.

* **Market Research**
* Conducting surveys, analyzing market data, and utilizing statistical techniques to gain insights into consumer preferences, market trends, and competitive landscapes, aiding in product development and marketing strategies.
* **Human Resources Analytics**

Analyzing HR data, including employee performance, engagement, and turnover rates, to identify patterns and factors influencing workforce productivity, enabling businesses to optimize talent management strategies.

* **Traffic Analysis**

Analyzing traffic data from sensors, GPS systems, and transportation networks to identify congestion patterns, optimize traffic flow, and improve transportation infrastructure planning.

* **Environmental Data Analysis**

Analyzing environmental data, such as weather patterns, air quality measurements, and ecosystem monitoring, to understand climate change, assess environmental impact, and inform policy decisions.

* **Sports Performance Analysis**

Analyzing sports performance data, including player statistics, match data, and biometric measurements, to evaluate player performance, optimize strategies, and enhance training programs.

These scenarios further exemplify the wide-ranging applications of data analysis across industries, demonstrating its potential to uncover valuable insights, improve decision-making, and drive business outcomes.

### Manufacturing Efficiency in a Company.

# Summary

Data analysis is a powerful tool that helps us understand the world better, make informed decisions, and uncover interesting insights. It's used in various fields, such as science, business, and even sports. By analyzing data, we can discover patterns, make predictions, and learn new things about the world around us.

#### References

* **What is Data Analysis?**

*Retrieved from:* [*https://en.wikipedia.org/wiki/Data\_analysis*](https://en.wikipedia.org/wiki/Data_analysis)

*Retrieved from:* [*https://ori.hhs.gov/education/products/n\_illinois\_u/datamanagement/datopic*](https://ori.hhs.gov/education/products/n_illinois_u/datamanagement/datopic)

* **Types Of Data Analysis**

*Retrieved from:* [*https://www.simplilearn.com/data-analysis-methods-process-types-article*](https://www.simplilearn.com/data-analysis-methods-process-types-article)

* **Processes Of Data Analysis**

*Retrieved from:* [*https://www.simplilearn.com/data-analysis-methods-process-types-article*](https://www.simplilearn.com/data-analysis-methods-process-types-article)

* **Uses of Data Analysis**

*Open AI. (2023, June 2). Chat with ChatGPT (Version 3.5) [Understanding Data Analysis]. Retrieved from [https://chat.openai.com/]*

* **Importance Of Data Analysis**

*Retrieved from:* [*https://www.simplilearn.com/data-analysis-methods-process-types-article*](https://www.simplilearn.com/data-analysis-methods-process-types-article)

* **Data Analysis Methods**

*Retrieved from:* [*https://www.simplilearn.com/data-analysis-methods-process-types-article*](https://www.simplilearn.com/data-analysis-methods-process-types-article)

* **Top Data Analysis Tools**

*Retrieved from:* [*https://www.simplilearn.com/data-analysis-methods-process-types-article*](https://www.simplilearn.com/data-analysis-methods-process-types-article)

* **What Is the Importance of Data Analysis in Research?**

*Retrieved from: https://www.simplilearn.com/data-analysis-methods-process-types-article*

* **Examples**

**Scenarios in which Data Analysis is Used**

*Open AI. (2023, June 9). Chat with ChatGPT (Version 3.5) [Scenarios in which Data Analysis is used]. Retrieved from [https://chat.openai.com/]*

* **Images**

Retrieved from: Google