Working with big data clusters?

Get understandable insights into large data sets with big data analytics software.

Browse software

❮ **Analytics Platforms Category**

What Is the Data Analysis Process? 5 Key Steps to Follow

October 6, 2021

by **Devin Pickell**

IN THIS POST

What are the 5 steps of the data analysis process

SHARE

Businesses generate and store tons of data every day, but what happens to this data after it's stored?

The short answer is that most of it sits in repositories and is almost never looked at again, which is quite counterintuitive.

Data can hold valuable insights into users, customer bases, and markets. When paired with **analytics software**, data can help businesses discover new product opportunities, marketing segments, industry verticals, and much more.

The problem isn't the lack of data available but the ambiguity in determining how exactly the data should be analyzed and used.

To clear up any uncertainties, businesses should understand the entire data analysis process in detail to make data-driven and

What is data analysis?

...ng, collecting, cleaning, examining, and modeling data to derive useful ...erstand the derived information for data-driven decision-making.

Now that you have a general overview of the data analysis process, it's time to dig deeper into each step.

What are the 5 steps of the data analysis process

The data analysis process is a collection of steps required to make sense of the available data. Identifying the critical stages in a data analysis process is a no-brainer. However, each step is equally important to ensure that the data is analyzed correctly and provides valuable and actionable information. Let's take a look at the five essential steps that make up a data analysis process fl...

Step 1: Define why you need data analysis

Before getting into the nitty-gritty of data analysis, a business must first define why it requires a well-founded process in the first place. The first step in a data analysis process is determining why you need data analysis. This need typically stems from a business problem or question, such as:

How can we reduce production costs without sacrificing quality?

In addition to finding a purpose, consider which metrics to track along the way. Also, be sure to identify sources of data when it's time to collect.

This process can be long and arduous, so building a roadmap will greatly prepare your data team for all the following steps.

... me to begin collecting the data needed for analysis. This step is important because the ... ermines how in-depth the analysis is.

... ces, also known as internal sources. This is typically **structured data** gathered from CRM software, ERP systems, marketing automation tools, and others. These sources contain information about customers, finances, gaps in sales, and more.

Then comes secondary sources, also known as external sources. This is both structured and **unstructured data** that can be gathered from many places.

For example, if you're looking to perform a sentiment analysis toward your brand, you could gather data from review sites or social media APIs.

Where is data collected?

Internal sources

- CRM software
- ERP software
- Marketing automation
- Databases and systems

External sources

- World health data
- Census bureaus
- Global finance data
- Google public data
- Google trends
- Amazon API
- Social media APIs
- Associated Press API
- Review websites

While it's not required to gather data from secondary sources, it could add another element to your data analysis. This is becoming more common in the age of big data.

Step 3: Clean unnecessary data

Once data is collected from all the necessary sources, your data team will be tasked with cleaning and sorting through it. Data cleaning is extremely important during the data analysis process, simply because not all data is *good* data.

Data scientists must identify and purge duplicate data, anomalous data, and other inconsistencies that could skew the analysis to generate accurate results.

With advances in **data science and machine learning platforms**, more intelligent automation can save a data analyst's valuable

[Data a]nalysis

[...] process is analyzing and manipulating the data. This can be done in a variety of ways.

One way is through **data mining**, which is defined as "knowledge discovery within databases". **Data mining techniques** like clustering analysis, anomaly detection, association rule mining, and others could unveil hidden patterns in data that weren't previously visible.

There's also business intelligence and **data visualization software**, both of which are optimized for decision-makers and business users. These options generate easy-to-understand reports, dashboards, scorecards, and charts.

Data scientists may also apply predictive analytics, which makes up one of the four data analytics used today (descriptive, diagnostic, predictive, prescriptive). **Predictive analysis** looks ahead to the future, attempting to forecast what will likely happen next based on [...] on.

What are the types of data analysis methods?

Data analysis methods can be broadly classified into the following categories:

- Quantitative data analysis
- Qualitative data analysis
- Statistical analysis
- Textual analysis
- Descriptive analysis
- Predictive analysis
- Prescriptive analysis
- Diagnostic analysis

Examples of data analysis techniques

Data analysts can use many data analysis techniques to extract meaningful information from aw data for real-world applications and computational purposes. Some of the notable data analysis techniques that aid a data analysis process are:

Exploratory data analysis

Exploratory data analysis is used to understand the messages within a dataset. This technique involves many iterative processes to ensure that the cleaned data is further sorted to better understand the useful meaning. Data visualization techniques such as analyzing data in an Excel sheet or other graphical format and descriptive analysis techniques such as calculating the mean or median are examples of exploratory data analysis.
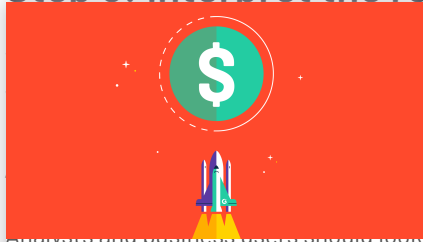
Using algorithms and models

variable). Such techniques are part of inferential statistics, the process of analyzing statistical data to draw conclusions about the relationship between different sets of data.

Step 5: Interpret the results

[...] from the data analysis. This part is essential because it's how a business will gain actual [...]

[...] validate why you conducted it, even if it's not 100 percent conclusive. For example, [...] to reduce production costs without sacrificing quality."

Analysts and business users should look to collaborate during this process. Also, when interpreting results, consider any challenges or limitations that may not have been present in the data. This will only bolster your confidence in the next steps.

Why is data analysis so important?

From small businesses to global enterprises, the amount of data businesses generate today is simply staggering, and this is why the term "big data" has become so buzzwordy.

However, this mountain of data hardly does much other than clog up cloud storage and databases without proper data analysis.

To [...] within your systems, learn more about **data analytics** and implement it to extract valuable insights.

Want to learn more about Analytics Platforms? **Explore Analytics Platforms products.**

Working with big data clusters?

Get understandable insights into large data sets with big data analytics software.

Browse software

Working with big data clusters?

Get understandable insights into large data sets with big data analytics software.

Browse software

Working with big data clusters?

Get understandable insights into large data sets with big data analytics software.

Browse software

Working with big data clusters?

Get understandable insights into large data sets with big data analytics software.

Browse software

Working with big data clusters?

Get understandable insights into large data sets with big data analytics software.

[Browse software]

Devin Pickell

Devin is a former senior content specialist at G2. Prior to G2, he helped scale early-stage startups out of Chicago's booming tech scene. Outside of work, he enjoys watching his beloved Cubs, playing baseball, and gaming. (he/him/his)

Explore More G2 Articles

ERP Statistics

Certified Ethical Hacking

Video Game Development

Discrete vs Continuous Data