

Overview of data analysis

6 minutes

Before data can be used to tell a story, it must be run through a process that makes it usable in the story. Data analysis is the process of identifying, cleaning, transforming, and modeling data to discover meaningful and useful information. The data is then crafted into a story through reports for analysis to support the critical decision-making process.



As the world becomes more data-driven, storytelling through data analysis is becoming a vital component and aspect of large and small businesses. It is the reason that organizations continue to hire data analysts.

Data-driven businesses make decisions based on the story that their data tells, and in today's data-driven world, data is not being used to its full potential, a challenge that most businesses face. Data analysis is, and should be, a critical aspect of all organizations to help determine the impact to their business, including evaluating customer sentiment, performing market and product research, and identifying trends or other data insights.

While the process of data analysis focuses on the tasks of cleaning, modeling, and visualizing data, the concept of data analysis and its importance to business should not be understated. To analyze data, core components of analytics are divided into the following categories:

- Descriptive
- Diagnostic

- Predictive
- Prescriptive
- Cognitive

Descriptive analytics

Descriptive analytics help answer questions about what has happened based on historical data. Descriptive analytics techniques summarize large datasets to describe outcomes to stakeholders.

By developing key performance indicators (KPIs), these strategies can help track the success or failure of key objectives. Metrics such as return on investment (ROI) are used in many industries, and specialized metrics are developed to track performance in specific industries.

An example of descriptive analytics is generating reports to provide a view of an organization's sales and financial data.

Diagnostic analytics

Diagnostic analytics help answer questions about why events happened. Diagnostic analytics techniques supplement basic descriptive analytics, and they use the findings from descriptive analytics to discover the cause of these events. Then, performance indicators are further investigated to discover why these events improved or became worse. Generally, this process occurs in three steps:

1. Identify anomalies in the data. These anomalies might be unexpected changes in a metric or a particular market.
2. Collect data that's related to these anomalies.
3. Use statistical techniques to discover relationships and trends that explain these anomalies.

Predictive analytics

Predictive analytics help answer questions about what will happen in the future. Predictive analytics techniques use historical data to identify trends and determine if they're likely to recur. Predictive analytical tools provide valuable insight into what might happen in the future. Techniques include a variety of statistical and machine learning techniques such as neural networks, decision trees, and regression.

Prescriptive analytics

Prescriptive analytics help answer questions about which actions should be taken to achieve a goal or target. By using insights from predictive analytics, organizations can make data-driven decisions. This technique allows businesses to make informed decisions in the face of uncertainty. Prescriptive analytics techniques rely on machine learning strategies to find patterns in large datasets. By analyzing past decisions and events, organizations can estimate the likelihood of different outcomes.

Cognitive analytics

Cognitive analytics attempt to draw inferences from existing data and patterns, derive conclusions based on existing knowledge bases, and then add these findings back into the knowledge base for future inferences, a self-learning feedback loop. Cognitive analytics help you learn what might happen if circumstances change and determine how you might handle these situations.

Inferences aren't structured queries based on a rules database; rather, they're unstructured hypotheses that are gathered from several sources and expressed with varying degrees of confidence. Effective cognitive analytics depend on machine learning algorithms, and will use several natural language processing concepts to make sense of previously untapped data sources, such as call center conversation logs and product reviews.

Example

By enabling reporting and data visualizations, a retail business uses descriptive analytics to look at patterns of purchases from previous years to determine what products might be popular next year. The company might also look at supporting data to understand why a particular product was popular and if that trend is continuing, which will help them determine whether to continue stocking that product.

A business might determine that a certain product was popular over a specific timeframe. Then, they can use this analysis to determine whether certain marketing efforts or online social activities contributed to the sales increase.

An underlying facet of data analysis is that a business needs to trust its data. As a practice, the data analysis process will capture data from trusted sources and shape it into something that is consumable, meaningful, and easily understood to help with the decision-making process. Data analysis enables businesses to fully understand their data through data-driven processes and decisions, allowing them to be confident in their decisions.

As the amount of data grows, so does the need for data analysts. A data analyst knows how to organize information and distill it into something relevant and comprehensible. A data analyst knows how to gather the right data and what to do with it, in other words, making sense of the data in your data overload.

Next unit: Roles in data

[Continue >](#)