Data Science

Reviewed by [Jake Frankenfield and Caroline Banton](#)   Updated Aug 27, 2019

What Is Data Science?

Data science provides meaningful information based on large amounts of complex data or [big data](#). Data science, or data-driven science, combines different fields of work in statistics and computation to interpret data for decision-making purposes.

Understanding Data Science

Data is drawn from different sectors, channels, and platforms including cell phones, social media, e-commerce sites, healthcare surveys, and Internet searches. The increase in the amount of data available opened the door to a new field of study based on big data—the massive data sets that contribute to the creation of better operational tools in all [sectors](#).

The continually increasing access to data is possible due to advancements in technology and collection techniques. Individuals buying patterns and behavior can be monitored and predictions made based on the information gathered.

However, the ever-increasing data is unstructured and requires parsing for effective decision making. This process is complex and time-consuming for companies—hence, the emergence of data science.

Data science, or data-driven science, uses big data and machine learning to interpret data for decision-making purposes.

A Brief History of Data Science

The term data science has existed for the better part of the last 30 years and was originally used as a substitute for "computer science" in 1960. Approximately 15 years later, the term was used to define the survey of data processing methods used in different applications. In 2001, data science was introduced as an independent discipline. The Harvard Business Review published an [article](#) in 2012 describing the role of the data scientist as the "sexiest job of the 21st century."

Key Takeaways

- Advances in technology, the Internet, social media, and the use of technology have all increased access to big data.
- Data science uses techniques such as machine learning and artificial intelligence to extract meaningful information and to predict future patterns and behaviors.
- The field of data science is growing as technology advances and big data collection and analysis techniques become more sophisticated.

How Data Science Is Applied

Data science incorporates tools from multiple disciplines to gather a data set, process, and derive insights from the data set, extract meaningful data from the set, and interpret it for decision-making purposes. The disciplinary areas that make up the data science field include mining, statistics, machine learning, analytics, and programming.

[Data mining](#) applies algorithms to the complex data set to reveal patterns that are then used to extract useful and relevant data from the set. Statistical measures or [predictive analytics](#) use this extracted data to gauge events that are likely to happen in the future based on what the data shows happened in the past.

Machine learning is an artificial intelligence tool that processes mass quantities of data that a human would be unable to process in a lifetime. Machine learning perfects the decision model presented under predictive analytics by matching the likelihood of an event happening to what actually happened at a predicted time.

Using analytics, the data analyst collects and processes the structured data from the machine learning stage using [algorithms](). The analyst interprets, converts, and summarizes the data into a cohesive language that the decision-making team can understand. Data science is applied to practically all contexts and, as the data scientist's role evolves, the field will expand to encompass data architecture, data engineering, and data administration.

Fast Fact

According to IBM, the demand for data scientists is expected to increase by 28% by 2020.

The Data Scientist Defined

A data scientist collects, analyzes, and interprets large volumes of data, in many cases, to improve a company's operations. Data scientist professionals develop statistical models that analyze data and detect patterns, trends, and relationships in data sets. This information can be used to predict consumer behavior or to identify business and operational risks. The data scientist is often a storyteller presenting data insights to decision makers in a way that is understandable and applicable to problem-solving.

Data Science Today

Companies are applying big data and data science to everyday activities to bring value to consumers. Banking institutions are capitalizing on big data to

enhance their fraud detection successes. Asset management firms are using big data to predict the likelihood of a security's price moving up or down at a stated time.

Companies such as Netflix mine big data to determine what products to deliver to its users. Netflix also uses algorithms to create personalized recommendations for users based on their viewing history. Data science is evolving at a rapid rate, and its applications will continue to change lives into the future.