Data Science

Data science is the study of where information comes from, what it represents and how it can be turned into a valuable resource in the creation of business and [IT strategies](https://searchcio.techtarget.com/definition/IT-strategy-information-technology-strategy). Mining large amounts of structured and unstructured data to identify patterns can help an organization rein in costs, increase efficiencies, recognize new market opportunities and increase the organization's competitive advantage.

The data science field employs mathematics, statistics and computer science disciplines, and incorporates techniques like [machine learning](https://searchenterpriseai.techtarget.com/definition/machine-learning-ML), cluster analysis, [data mining](https://searchsqlserver.techtarget.com/definition/data-mining) and visualization.

### Data scientists

As the amount of data generated by the typical modern business increases, so does the prominence of data scientists hired by organizations to help them turn [raw data](https://searchdatamanagement.techtarget.com/definition/raw-data) into valuable business [information](https://searchsqlserver.techtarget.com/definition/information). Data extraction is the act of retrieving specific data from unstructured or poorly structured data sources for further processing and investigation. Data scientists must possess a combination of analytic, machine learning, data mining and statistical skills, as well as experience with algorithms and coding. Along with managing and interpreting large amounts of data, many data scientists are also tasked with creating data visualization models that help illustrate the business value of digital information.

To be effective, however, data scientists must possess [emotional intelligence](https://searchcio.techtarget.com/definition/emotional-intelligence) in addition to education and experience in [data analytics](https://searchdatamanagement.techtarget.com/definition/data-analytics). Perhaps the most important skill a data scientist must possess is the ability to present the data insights to others, including C-suite executives, and explain the significance of the data in a way that can be easily understood.

Data scientists draw the digital information they are studying from a growing list of channels and sources, including smartphones, internet of things [(IoT) devices](https://internetofthingsagenda.techtarget.com/definition/IoT-device), social media, surveys, purchases, and internet searches and behavior. By sorting through these large data sets, data scientists can identify patterns to solve problems through data analysis -- a process known as data mining.

### Benefits of data science

The main advantage of enlisting data science in an organization is the empowerment and facilitation of decision-making. Organizations with data scientists can factor in quantifiable, data-based evidence into their business decisions. These [data-driven decisions](https://whatis.techtarget.com/definition/data-driven-decision-management-DDDM) can ultimately lead to increased profitability and improved operational efficiency, business performance and workflows. In customer-facing organizations, data science helps identify and refine target audiences. Data science can also assist recruitment: Internal processing of applications and data-driven aptitude tests and games can help an organization's human resources team make quicker and more accurate selections during the hiring process.

The specific benefits of data science vary depending on the company's goal and the industry. Sales and marketing departments, for example, can mine customer data to improve conversion rates or create one-to-one marketing campaigns. Banking institutions are mining data to enhance fraud detection. Streaming services like Netflix mine data to determine what its users are interested in, and use that data to determine what TV shows or films to produce. Data-based algorithms are also used at Netflix to create personalized recommendations based on a user's viewing history. Shipment companies like DHL, FedEx and UPS use data science to find the best delivery routes and times, as well as the best modes of transport for their shipments.

Data science is still an emerging field within the enterprise because the identification and analysis of vast amounts of [unstructured data](https://searchbusinessanalytics.techtarget.com/definition/unstructured-data) can prove too complex, expensive and time-consuming for companies.

### Data science and machine learning

Machine learning is often incorporated in data science. Machine learning is an [artificial intelligence](https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence) (AI) tool that essentially automates the data-processing portion of data science. Machine learning integrates advanced algorithms that learn on their own and can process massive amounts of data in a fraction of the time it would take a human.