

The professional title of "data scientist" has been attributed to DJ Patil and Jeff Hammerbacher in 2008. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. Data analysis typically involves working with smaller, structured datasets to answer specific questions or solve specific problems. For instance, a data scientist might develop a recommendation system for an e-commerce platform by analyzing user behavior patterns and using machine learning algorithms to predict user preferences. However, data science is different from computer science and information science. The professional title of "data scientist" has been attributed to DJ Patil and Jeff Hammerbacher in 2008. While both fields involve working with data, data science is more of an interdisciplinary field that involves the application of statistical, computational, and machine learning methods to extract insights from data and make predictions, while data analysis is more focused on the examination and interpretation of data to identify patterns and trends. For instance, a data scientist might develop a recommendation system for an e-commerce platform by analyzing user behavior patterns and using machine learning algorithms to predict user preferences. This can involve tasks such as data cleaning, data visualization, and exploratory data analysis to gain insights into the data and develop hypotheses about relationships between variables. In 2015, the American Statistical Association identified database management, statistics and machine learning, and distributed and parallel systems as the three emerging foundational professional communities. During the 1990s, popular terms for the process of finding patterns in datasets (which were increasingly large) included "knowledge discovery" and "data mining". As such, it incorporates skills from computer science, statistics, information science, mathematics, data visualization, information visualization, data sonification, data integration, graphic design, complex systems, communication and business. Andrew Gelman of Columbia University has described statistics as a non-essential part of data science. In 1985, in a lecture given to the Chinese Academy of Sciences in Beijing, C. It uses techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, information science, and domain knowledge. Data science is a "concept to unify statistics, data analysis, informatics, and their related methods" to "understand and analyze actual phenomena" with data. A decade later, they reaffirmed it, stating that "the job is more in demand than ever with employers". F. The professional title of "data scientist" has been attributed to DJ Patil and Jeff Hammerbacher in 2008. Both fields require a solid foundation in statistics, programming, and data visualization, as well as the ability to communicate findings effectively to both technical and non-technical audiences. "Data science" became more widely used in the next few years: in 2002, the Committee on Data for Science and Technology launched the Data Science Journal. The modern conception of data science as an independent discipline is sometimes attributed to William S. This can involve tasks such as data cleaning, data visualization, and exploratory data analysis to gain insights into the data and develop hypotheses about relationships between variables. After the 1985 lecture at the Chinese Academy of Sciences in Beijing, in 1997 C. In 2015, the American Statistical Association identified database management, statistics and machine learning, and distributed and parallel systems as the three emerging foundational professional communities.