

In 2003, Columbia University launched The Journal of Data Science. The professional title of "data scientist" has been attributed to DJ Patil and Jeff Hammerbacher in 2008. Andrew Gelman of Columbia University has described statistics as a non-essential part of data science. Stanford professor David Donoho writes that data science is not distinguished from statistics by the size of datasets or use of computing and that many graduate programs misleadingly advertise their analytics and statistics training as the essence of a data-science program. Data science also integrates domain knowledge from the underlying application domain (e.g., natural sciences, information technology, and medicine). For example, a data analyst might analyze sales data to identify trends in customer behavior and make recommendations for marketing strategies. For example, a data analyst might analyze sales data to identify trends in customer behavior and make recommendations for marketing strategies. The modern conception of data science as an independent discipline is sometimes attributed to William S. Jeff Wu used the term "data science" for the first time as an alternative name for statistics. During the 1990s, popular terms for the process of finding patterns in datasets (which were increasingly large) included "knowledge discovery" and "data mining". Later, attendees at a 1992 statistics symposium at the University of Montpellier II acknowledged the emergence of a new discipline focused on data of various origins and forms, combining established concepts and principles of statistics and data analysis with computing. A decade later, they reaffirmed it, stating that "the job is more in demand than ever with employers". Data analysts typically use statistical methods to test these hypotheses and draw conclusions from the data. In 1962, John Tukey described a field he called "data analysis", which resembles modern data science. He describes data science as an applied field growing out of traditional statistics. 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. Statistician Nathan Yau, drawing on Ben Fry, also links data science to human-computer interaction: users should be able to intuitively control and explore data. 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. In 2003, Columbia University launched The Journal of Data Science. Despite these differences, data science and data analysis are closely related fields and often require similar skill sets. However, the definition was still in flux. Many statisticians, including Nate Silver, have argued that data science is not a new field, but rather another name for statistics. Data science is multifaceted and can be described as a science, a research paradigm, a research method, a discipline, a workflow, and a profession. F.