

Vasant Dhar writes that statistics emphasizes quantitative data and description. The professional title of "data scientist" has been attributed to DJ Patil and Jeff Hammerbacher in 2008. Many statisticians, including Nate Silver, have argued that data science is not a new field, but rather another name for statistics. 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. The modern conception of data science as an independent discipline is sometimes attributed to William S. Turing Award winner Jim Gray imagined data science as a "fourth paradigm" of science (empirical, theoretical, computational, and now data-driven) and asserted that "everything about science is changing because of the impact of information technology" and the data deluge. However, data science is different from computer science and information science. In summary, data analysis and data science are distinct yet interconnected disciplines within the broader field of data management and analysis. However, the definition was still in flux. In addition to statistical analysis, data science often involves tasks such as data preprocessing, feature engineering, and model selection. Data scientists often work with unstructured data such as text or images and use machine learning algorithms to build predictive models and make data-driven decisions. Cleveland. A data scientist is a professional who creates programming code and combines it with statistical knowledge to create insights from data. In 1998, Hayashi Chikio argued for data science as a new, interdisciplinary concept, with three aspects: data design, collection, and analysis. 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, on the other hand, is a more complex and iterative process that involves working with larger, more complex datasets that often require advanced computational and statistical methods to analyze. Jeff Wu used the term "data science" for the first time as an alternative name for statistics. 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. Jeff Wu again suggested that statistics should be renamed data science. Both fields play vital roles in leveraging the power of data to understand patterns, make informed decisions, and solve complex problems across various domains. He describes data science as an applied field growing out of traditional statistics. 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. Jeff Wu again suggested that statistics should be renamed data science. While data analysis focuses on extracting insights from existing data, data science goes beyond that by incorporating the development and implementation of predictive models to make informed decisions.