

After the 1985 lecture at the Chinese Academy of Sciences in Beijing, in 1997 C. 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. The field encompasses preparing data for analysis, formulating data science problems, analyzing data, developing data-driven solutions, and presenting findings to inform high-level decisions in a broad range of application domains. Both fields play vital roles in leveraging the power of data to understand patterns, make informed decisions, and solve complex problems across various domains. In contrast, data science deals with quantitative and qualitative data (e.g., from images, text, sensors, transactions, customer information, etc.) and emphasizes prediction and action. He describes data science as an applied field growing out of traditional statistics. 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. Data analysis typically involves working with smaller, structured datasets to answer specific questions or solve specific problems. Data science is an interdisciplinary field focused on extracting knowledge from typically large data sets and applying the knowledge and insights from that data to solve problems in a wide range of application domains. 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. In 1996, the International Federation of Classification Societies became the first conference to specifically feature data science as a topic. For example, a data analyst might analyze sales data to identify trends in customer behavior and make recommendations for marketing strategies. 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. 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, 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. In addition to statistical analysis, data science often involves tasks such as data preprocessing, feature engineering, and model selection. Vasant Dhar writes that statistics emphasizes quantitative data and description. Data science is an interdisciplinary field focused on extracting knowledge from typically large data sets and applying the knowledge and insights from that data to solve problems in a wide range of application domains. They work at the intersection of mathematics, computer science, and domain expertise to solve complex problems and uncover hidden patterns in large datasets. In 2014, the American Statistical Association's Section on Statistical Learning and Data Mining changed its name to the Section on Statistical Learning and Data Science, reflecting the ascendant popularity 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. 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. Jeff Wu used the term "data science" for the first time as an alternative name for statistics. Davenport and DJ Patil declared "Data Scientist: The Sexiest Job of the 21st Century", a catchphrase that was picked up even by major-city newspapers like the New York Times and the Boston Globe. In contrast, data science deals with quantitative and qualitative data (e.g., from images, text, sensors, transactions, customer information, etc.) and emphasizes prediction and action.