

Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. Data science is an interdisciplinary academic field that uses statistics, scientific computing, scientific methods, processes, algorithms and systems to extract or extrapolate knowledge and insights from noisy, structured, and unstructured data. 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. 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. 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. 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. 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. A decade later, they reaffirmed it, stating that "the job is more in demand than ever with employers". Though it was used by the National Science Board in their 2005 report "Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century", it referred broadly to any key role in managing a digital data collection. Data scientists are often responsible for collecting and cleaning data, selecting appropriate analytical techniques, and deploying models in real-world scenarios. 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. They work at the intersection of mathematics, computer science, and domain expertise to solve complex problems and uncover hidden patterns in large datasets. Despite these differences, data science and data analysis are closely related fields and often require similar skill sets. A decade later, they reaffirmed it, stating that "the job is more in demand than ever with employers". Data science and data analysis are both important disciplines in the field of data management and analysis, but they differ in several key ways. However, data science is different from computer science and information science. Data scientists are responsible for breaking down big data into usable information and creating software and algorithms that help companies and organizations determine optimal operations. Data science is multifaceted and can be described as a science, a research paradigm, a research method, a discipline, a workflow, and a profession. 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. 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 again suggested that statistics should be renamed data science. In summary, data analysis and data science are distinct yet interconnected disciplines within the broader field of data management and analysis. 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 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.