

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. Many statisticians, including Nate Silver, have argued that data science is not a new field, but rather another name for statistics. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. In 1962, John Tukey described a field he called "data analysis", which resembles modern data science. 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 also integrates domain knowledge from the underlying application domain (e.g., natural sciences, information technology, and medicine). 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. Data science and data analysis are both important disciplines in the field of data management and analysis, but they differ in several key ways. Moreover, both fields benefit from critical thinking and domain knowledge, as understanding the context and nuances of the data is essential for accurate analysis and modeling. Data science and data analysis are both important disciplines in the field of data management and analysis, but they differ in several key ways. 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. A decade later, they reaffirmed it, stating that "the job is more in demand than ever with employers". He describes data science as an applied field growing out of traditional statistics. Despite these differences, data science and data analysis are closely related fields and often require similar skill sets. 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. A data scientist is a professional who creates programming code and combines it with statistical knowledge to create insights from data. In 1962, John Tukey described a field he called "data analysis", which resembles modern 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. 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. 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. 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. 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. He describes data science as an applied field growing out of traditional statistics.