

A data scientist is a professional who creates programming code and combines it with statistical knowledge to create insights from data. A data scientist is a professional who creates programming code and combines it with statistical knowledge to create insights from data. 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. During the 1990s, popular terms for the process of finding patterns in datasets (which were increasingly large) included "knowledge discovery" and "data mining". However, data science is different from computer science and information science. 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. 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 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. 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. Jeff Wu used the term "data science" for the first time as an alternative name for statistics. 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. F. "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. 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. 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" became more widely used in the next few years: in 2002, the Committee on Data for Science and Technology launched the Data Science Journal. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. 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. For example, a data analyst might analyze sales data to identify trends in customer behavior and make recommendations for marketing strategies. 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. In 1998, Hayashi Chikio argued for data science as a new, interdisciplinary concept, with three aspects: data design, collection, and analysis. Many statisticians, including Nate Silver, have argued that data science is not a new field, but rather another name for statistics. Data analysis typically involves working with smaller, structured datasets to answer specific questions or solve specific problems. F.