

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. It uses techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, information science, and domain knowledge. 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. There is still no consensus on the definition of data science, and it is considered by some to be a buzzword. Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. For example, a data analyst might analyze sales data to identify trends in customer behavior and make recommendations for marketing strategies. Jeff Wu used the term "data science" for the first time as an alternative name for statistics. 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. Both fields play vital roles in leveraging the power of data to understand patterns, make informed decisions, and solve complex problems across various domains. However, the definition was still in flux. 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. In a 2001 paper, he advocated an expansion of statistics beyond theory into technical areas; because this would significantly change the field, it warranted a new name. F. A decade later, they reaffirmed it, stating that "the job is more in demand than ever with employers". Others argue that data science is distinct from statistics because it focuses on problems and techniques unique to digital data. 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. 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. Data analysis focuses on extracting insights and drawing conclusions from structured data, while data science involves a more comprehensive approach that combines statistical analysis, computational methods, and machine learning to extract insights, build predictive models, and drive data-driven decision-making. Data science is multifaceted and can be described as a science, a research paradigm, a research method, a discipline, a workflow, and a profession. 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. After the 1985 lecture at the Chinese Academy of Sciences in Beijing, in 1997 C. 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. In 1998, Hayashi Chikio argued for data science as a new, interdisciplinary concept, with three aspects: data design, collection, and analysis. This can involve tasks such as data cleaning, data visualization, and exploratory data analysis to gain insights into the data and develop hypotheses about relationships between variables. He reasoned that a new name would help statistics shed inaccurate stereotypes, such as being synonymous with accounting or limited to describing data.