

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. 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. The modern conception of data science as an independent discipline is sometimes attributed to William S. The term "data science" has been traced back to 1974, when Peter Naur proposed it as an alternative name to computer science. 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. In 1985, in a lecture given to the Chinese Academy of Sciences in Beijing, C. Jeff Wu again suggested that statistics should be renamed data science. In 2012, technologists Thomas H. In 2003, Columbia University launched The Journal of Data Science. Many statisticians, including Nate Silver, have argued that data science is not a new field, but rather another name for statistics. In 1998, Hayashi Chikio argued for data science as a new, interdisciplinary concept, with three aspects: data design, collection, and analysis. The term "data science" has been traced back to 1974, when Peter Naur proposed it as an alternative name to computer science. In 1985, in a lecture given to the Chinese Academy of Sciences in Beijing, C. 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. 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 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. 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. 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. While both fields involve working with data, data science is more of an interdisciplinary field that involves the application of statistical, computational, and machine learning methods to extract insights from data and make predictions, while data analysis is more focused on the examination and interpretation of data to identify patterns and trends. 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 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 scientists are responsible for breaking down big data into usable information and creating software and algorithms that help companies and organizations determine optimal operations. In 2003, Columbia University launched The Journal of Data Science. 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.