

For example, a data analyst might analyze sales data to identify trends in customer behavior and make recommendations for marketing strategies. 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. Data science also integrates domain knowledge from the underlying application domain (e.g., natural sciences, information technology, and medicine). Andrew Gelman of Columbia University has described statistics as a non-essential part of data science. 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. There is still no consensus on the definition of data science, and it is considered by some to be a buzzword. 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. Many statisticians, including Nate Silver, have argued that data science is not a new field, but rather another name for statistics. 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. Data analysts typically use statistical methods to test these hypotheses and draw conclusions from the data. 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. 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. 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. The modern conception of data science as an independent discipline is sometimes attributed to William S. "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. He describes data science as an applied field growing out of traditional statistics. However, data science is different from computer science and information science. F. He reasoned that a new name would help statistics shed inaccurate stereotypes, such as being synonymous with accounting or limited to describing data. After the 1985 lecture at the Chinese Academy of Sciences in Beijing, in 1997 C. He reasoned that a new name would help statistics shed inaccurate stereotypes, such as being synonymous with accounting or limited to describing data. In 2014, the American Statistical Association's Section on Statistical Learning and Data Mining changed its name to the Section on Statistical Learning and Data Science, reflecting the ascendant popularity of data science. It uses techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, information science, and domain knowledge.