

He reasoned that a new name would help statistics shed inaccurate stereotypes, such as being synonymous with accounting or limited to describing data. In 2012, technologists Thomas H. Andrew Gelman of Columbia University has described statistics as a non-essential part of data science. 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. 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. Data science also integrates domain knowledge from the underlying application domain (e.g., natural sciences, information technology, and medicine). Data analysis typically involves working with smaller, structured datasets to answer specific questions or solve specific problems. 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. 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. A data scientist is a professional who creates programming code and combines it with statistical knowledge to create insights from data. The term "data science" has been traced back to 1974, when Peter Naur proposed it as an alternative name to computer science. 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. In summary, data analysis and data science are distinct yet interconnected disciplines within the broader field of data management and analysis. However, the definition was still in flux. 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. However, the definition was still in flux. 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. 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. The term "data science" has been traced back to 1974, when Peter Naur proposed it as an alternative name to computer science. Statistician Nathan Yau, drawing on Ben Fry, also links data science to human-computer interaction: users should be able to intuitively control and explore data. Jeff Wu used the term "data science" for the first time as an alternative name for statistics. Data analysis typically involves working with smaller, structured datasets to answer specific questions or solve specific problems. 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.