Andrew Gelman of Columbia University has described statistics as a non-essential part of data science. 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 is a "concept to unify statistics, data analysis, informatics, and their related methods" to "understand and analyze actual phenomena" with data. 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. The modern conception of data science as an independent discipline is sometimes attributed to William S. Both fields play vital roles in leveraging the power of data to understand patterns, make informed decisions, and solve complex problems across various domains. Andrew Gelman of Columbia University has described statistics as a non-essential part of data science. He reasoned that a new name would help statistics shed inaccurate stereotypes, such as being synonymous with accounting or limited to describing data. For example, a data analyst might analyze sales data to identify trends in customer behavior and make recommendations for marketing strategies. 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. In 1962, John Tukey described a field he called "data analysis", which resembles modern data science. There is still no consensus on the definition of data science, and it is considered by some to be a buzzword. 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. Cleveland. In 2012, technologists Thomas H. "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 1962, John Tukey described a field he called "data analysis", which resembles modern data science. Vasant Dhar writes that statistics emphasizes quantitative data and description. 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. 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. A decade later, they reaffirmed it, stating that "the job is more in demand than ever with employers". In 1998, Hayashi Chikio argued for data science as a new, interdisciplinary concept, with three aspects: data design, collection, and analysis. In 1985, in a lecture given to the Chinese Academy of Sciences in Beijing, C. 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.