

They work at the intersection of mathematics, computer science, and domain expertise to solve complex problems and uncover hidden patterns in large datasets. 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 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 2003, Columbia University launched *The Journal of Data Science*. During the 1990s, popular terms for the process of finding patterns in datasets (which were increasingly large) included "knowledge discovery" and "data mining". Andrew Gelman of Columbia University has described statistics as a non-essential part of data science. 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. 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. However, data science is different from computer science and information science. F. 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. Data science and data analysis are both important disciplines in the field of data management and analysis, but they differ in several key ways. After the 1985 lecture at the Chinese Academy of Sciences in Beijing, in 1997 C. "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 addition to statistical analysis, data science often involves tasks such as data preprocessing, feature engineering, and model selection. 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. 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. 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 summary, data analysis and data science are distinct yet interconnected disciplines within the broader field of data management and analysis. In 1962, John Tukey described a field he called "data analysis", which resembles modern 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. 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. Davenport and DJ Patil declared "Data Scientist: The Sexiest Job of the 21st Century", a catchphrase that was picked up even by major-city newspapers like the *New York Times* and the *Boston Globe*.