

Both fields play vital roles in leveraging the power of data to understand patterns, make informed decisions, and solve complex problems across various domains. In 2012, technologists Thomas H. Davenport and D.J. Patil defined a data scientist as a professional who creates programming code and combines it with statistical knowledge to create insights from data. In contrast, data science deals with quantitative and qualitative data (e.g., from images, text, sensors, transactions, customer information, etc.) and emphasizes prediction and action. In addition to statistical analysis, data science often involves tasks such as data preprocessing, feature engineering, and model selection. In 1998, Hayashi Chikio argued for data science as a new, interdisciplinary concept, with three aspects: data design, collection, and analysis. As such, it incorporates skills from computer science, statistics, information science, mathematics, data visualization, information visualization, data sonification, data integration, graphic design, complex systems, communication and business. 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. "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. 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. Many statisticians, including Nate Silver, have argued that data science is not a new field, but rather another name for statistics. Data science and data analysis are both important disciplines in the field of data management and analysis, but they differ in several key ways. In contrast, data science deals with quantitative and qualitative data (e.g., from images, text, sensors, transactions, customer information, etc.) and emphasizes prediction and action. The modern conception of data science as an independent discipline is sometimes attributed to William S. Cleveland. Both fields require a solid foundation in statistics, programming, and data visualization, as well as the ability to communicate findings effectively to both technical and non-technical audiences. In addition to statistical analysis, data science often involves tasks such as data preprocessing, feature engineering, and model selection. In contrast, data science deals with quantitative and qualitative data (e.g., from images, text, sensors, transactions, customer information, etc.) and emphasizes prediction and action. Data scientists are often responsible for collecting and cleaning data, selecting appropriate analytical techniques, and deploying models in real-world scenarios. In 2003, Columbia University launched The Journal of Data Science. Cleveland. 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. Both fields require a solid foundation in statistics, programming, and data visualization, as well as the ability to communicate findings effectively to both technical and non-technical audiences. A decade later, they reaffirmed it, stating that "the job is more in demand than ever with employers". 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. Many statisticians, including Nate Silver, have argued that data science is not a new field, but rather another name for statistics.