

Jobs for data science professionals

Author: Benjamin Sombi . September 2021

What is Data Science?

Data science is still popular among trained professionals and organizations focused on collecting data and extracting valuable insights to help businesses flourish. Any organization can benefit from a large amount of data, but only if it is processed effectively. When we [entered the age of big data](#), the demand for storage increased dramatically. Until 2010, the primary focus was on developing cutting-edge infrastructure to store this valuable data, which would subsequently be accessed and processed to generate business insights. The attention has switched to processing this data now that frameworks like Hadoop have taken care of the storage portion. [In this article, we will discuss what data science](#) is, how it fits into the current state of big data and how to become a data scientist.

Data science is a branch of science that use scientific methods, procedures, [algorithms, and systems to extract knowledge and insights from noisy, structured, and unstructured data](#).

Both businesses and consumers gain from data science. According to the McKinsey Global Institute, big data can boost a retailer's profit margin by 60%, and ["services enabled by personal-location data can allow consumers to capture \\$600 billion in economic surplus"](#). Data science can increase company revenues while also saving customers money, resulting in a win-win situation for the economy.

In 2008, businesses saw the need for data professionals to organize and analyze massive amounts of data. That's how the title "data scientist" was established. Hal Varian, Google's top economist and a UC Berkeley professor of information sciences, business, and economics, foresaw the significance of adjusting to technology's effect and reconfiguration of diverse industries in a 2009 McKinsey & Company essay.

Influential data scientists can formulate pertinent questions, collect data from a range of sources, organize the data, convert results into solutions, and present their findings in a way that positively influences business decisions. Skilled data scientists are becoming increasingly valuable for businesses as these skills are required in almost every industry.

Why do businesses need Data Science?

We've progressed from working with small collections of structured data to enormous mines of unstructured and semi-structured data from various sources. When it comes to processing this massive amount of unstructured data, typical Business Intelligence technologies fall short. As a result, Data Science includes more complex tools for working with enormous volumes of data from various sources, including financial records, multimedia files, marketing forms, sensors and instruments, and text files.

Relevant use-cases are listed below, as well as the reasons behind Data Science's growing popularity among businesses:

Predictive analytics uses data science in a variety of ways. In weather forecasting, data from satellites, radars, ships, and aircraft is used to create models that can accurately forecast weather and predict imminent natural disasters. This assists in taking suitable measures at the right moment and avoiding reducing risks associated with specific trends.

Traditional models based on browser history, purchase history, and basic demographic characteristics have never been as precise in their product recommendations. With data science, large amounts of data and various data can be used to better train models and provide more exact suggestions.

Data science can also help you make better decisions. A classic example is self-driving or intelligent cars. To construct a picture (map) of its surroundings, a smart vehicle collects data in real-time from its surroundings using various sensors such as radars, cameras, and lasers. Based on this data and a robust Machine Learning algorithm, it makes critical driving decisions like turning, stopping, speeding, and so on.

What Does a Data Scientist Do?

During the previous decade, data scientists have become crucial assets in almost every company. These [professionals](#) are well-rounded, data-driven individuals with advanced technical capabilities who can construct complicated quantitative algorithms to organize and synthesize vast amounts of data to answer questions and drive strategy in their company. This is combined with the communication and leadership skills required to provide tangible results to numerous stakeholders throughout a company or organization.

Data scientists must be curious and results-oriented, with extensive industry knowledge and communication skills that allow them to communicate highly technical results to non-technical colleagues. They have a solid quantitative foundation in statistics and linear algebra and programming skills focusing on data warehousing, mining, and modeling, which they use to build and analyze algorithms.

Why Become a Data Scientist?

Large tech companies are no longer the only ones needing data scientists as more data becomes more accessible. A shortage of skilled people available to fill open positions poses a challenge to the expanding need for data science specialists across large and small sectors.

The demand for data scientists is not expected to decrease in the coming years. In 2017 and 2018, LinkedIn named data scientist as one of the most promising careers and the most in-demand by employers.

Steps to Become a Data Scientist

A strong data scientist's skill set includes modules in data mining, data analysis, programming, mathematics and statistics, machine learning, business, data hacking, data visualization, database & (big) data.

1. Learn Python

Learning a computer language should be the first and most crucial step toward Data Science (i.e. Python). Python is a free and open-source programming language that comes with a number of libraries. Python is the most frequent scripting because of its simplicity, versatility, and pre-installation of strong libraries (such as NumPy, SciPy, and Pandas) essential in data analysis and other parts of Data Science language used by the majority of Data Scientists.

2. Learn Statistics

Statistics is the grammar if Data Science is a language. Statistics is the process of studying and

interpreting substantial data sets. Statistics are as essential to us as air when it comes to data processing and gathering insights. We can use statistics to decipher the hidden details in massive datasets.

3. Data Collection

In the discipline of Data Science, this is one of the most crucial tasks. This ability requires familiarity with various tools for importing data from local systems such as CSV files and scraping data from websites using the BeautifulSoup Python library. Scraping can also be done using an API. Knowledge of Query Language or Python ETL pipelines can help with data collection.

4. Data Cleaning

As a Data Scientist, you'll spend the majority of your time on this step. Data cleaning is the process of removing undesired variables, missing values, category values, outliers, and incorrectly reported records from raw data to be used for work and analysis. Data cleaning is critical since real-world data is dirty, and attaining it with the help of numerous Python modules (such as Pandas and NumPy) is crucial for aspiring Data Scientists.

5. Exploratory Data Analysis

In the enormous subject of data science, EDA (exploratory data analysis) is the most significant part. It entails examining a variety of data, variables, data patterns, and trends and extracting relevant insights from them using a variety of graphical and statistical tools. EDA detects a variety of patterns that a machine learning program could miss. All data manipulation, analysis, and visualization are included.

6. Machine Learning & Deep Learning

A Data Scientist's most important talent in machine learning. Machine learning is used to create numerous predictive models, categorization models, and other models, and it is utilized by large corporations to optimize their planning based on forecasts. They are predicting the price of a car, for example.

Deep Learning, on the other hand, is a more advanced version of Machine Learning that uses Neural Networks to train data. Neural Networks are a framework that incorporates several machine learning algorithms for addressing various problems. Recurrent neural networks (RNN) and convolutional neural networks (CNN) are examples of neural networks.

7. Learn Deploying of ML model

The process of making your [Machine Learning Model available](#) for use to end-users is known as deployment. This is accomplished by integrating the model with a variety of existing production environments, allowing for the practical application of the ML model for a variety of business applications.

[Flask](#), [Pythoneverywhere](#), [MLOps](#), [Microsoft Azure](#), [Google Cloud](#), [Heroku](#), and other platforms are available for deploying your machine learning model.

8. Real-World Testing

After deployment, the Machine Learning Model should be tested and validated to ensure its effectiveness and correctness. Testing is essential in data science since it ensures that the ML model's efficiency and effectiveness are maintained.

There are various types of testing, such as A/B, AAB, and so on.

9. Explore and Practice using online datasets

The world's largest data science communities, such as Kaggle and Analytics Vidhya, are beneficial for connecting with various datasets and may thus be used to practice different data analysis approaches and machine learning algorithms. Competitions hosted in these communities are also helpful for honing data science abilities, assisting us in becoming proficient in Data Science more quickly.

10. Analytical Curiosity

Because data science is a rapidly growing topic, it necessitates an insatiable desire to learn more about it and a constant need to update and learn new skills and approaches.

This is the primary talent that will keep us up to date on new skills and concepts, preventing us from falling behind on numerous Data Science technology breakthroughs.

11. Non-Technical Skills

Teamwork, communication skills, task management, business understanding, and other non-technical skills are examples of non-technical skills.

Communication abilities enable us to communicate our technical thoughts and concepts to the Firm's non-technical workers and authorities.

Task management entails the correct administration and planning of the solution's delivery.

Business knowledge/aptitude, or an awareness of the industry in which we operate, is critical for various analyses and successful solutions to problems in those businesses.

Jobs for data science professionals

Here are some of the leading data science careers you can break into with an advanced degree.

1. Data Scientist

Average Salary: \$139,840 per year

Typical Job Requirements: For businesses, find, clean, and organize data. Data scientists will be required to examine enormous amounts of complex raw and processed data to uncover patterns that will benefit an organization and aid in critical business choices. Data scientists are far more technical than data analysts.

2. Machine Learning Engineer

Average Salary: \$114,826 per year

Typical Job Requirements: Data channels and software solutions are created by machine learning developers. They are responsible for executing tests and experiments to monitor the performance and usefulness of machine learning systems in addition to designing and building them. Solid statistics and programming skills, as well as knowledge of software engineering, are often required.

3. Machine Learning Scientist

Average Salary: \$114,121 per year

Typical Job Requirements: New data methodologies and algorithms, such as supervised, unsupervised, and deep learning techniques, are being investigated for use in adaptive systems. Research Scientist or Research Engineer are common titles for machine learning scientists.

4. Applications Architect

Average Salary: \$113,757 per year

Typical Job Requirements: Track the behavior of business applications and how they interact with one another and with users. Applications architects are also responsible for developing the architecture of

applications, which includes components such as the user interface and infrastructure.

5. Enterprise Architect

Average Salary: \$110,663 per year

Typical Job Requirements: An enterprise architect is in charge of coordinating an organization's strategy with the technology required to achieve its goals. To develop the systems architecture needed to meet those demands, they must understand the business and its technology requirements.

6. Data Architect

Average Salary: \$108,278 per year

Typical Job Requirements: Ascertain that data solutions are developed for performance and create analytics apps that can be used across numerous platforms. Data architects often strive to improve the performance and usefulness of existing database systems, offer access to database administrators and analysts, and design new database systems.

7. Infrastructure Architect

Average Salary: \$107,309 per year

Typical Job Requirements: Ensure that all business systems are operating at peak performance and capable of supporting the development of new technologies and system requirements. Cloud Infrastructure Architect is a job title that supervises a company's cloud computing strategy.

8. Data Engineer

Average Salary: \$102,864 per year

Typical Job Requirements: On acquired and stored data, do batch or real-time processing. Data engineers are also in charge of creating and maintaining data pipelines that enable data scientists to access information by creating a robust and integrated data ecosystem within a business.

9. Business Intelligence (BI) Developer

Average Salary: \$81,514 per year

Typical Job Requirements: Business intelligence (BI) developers create techniques to help business users quickly discover the data they need to make better business decisions. They employ BI tools or design custom BI analytic solutions to help end-users comprehend their systems because they are highly data-savvy.

10. Statistician

Average Salary: \$76,884 per year

Typical Job Requirements: Statisticians collect, analyze, and evaluate data to uncover trends and patterns that can be utilized to guide organizational decision-making. In addition, statisticians' everyday tasks may include designing data gathering systems, disseminating findings to stakeholders, and advising on organizational strategy.

11. Data Analyst

Average Salary: \$62 453 per year

Typical Job Requirements: Large data sets can be transformed and manipulated to meet a company's study needs. This function might also include tracking site statistics and evaluating A/B testing in many companies. Data analysts also help with decision-making by generating reports for organizational leaders that effectively explain patterns and insights gathered from their research.

This article was written by Benjamin Sombi, a consultant at the Industrial Psychology Consultants, a management and human resources consultants company.

<https://thehumancapitalhub.com/articles/jobs-for-data-science-professionals>