

What Skills are Needed to be a Data Scientist?

TechRounder PDF Edition

Live article: <https://www.techrounder.com/technology/what-skills-are-needed-to-be-a-data-scientist/>

By Vipin PG | Published June 26, 2020 | Updated June 26, 2020 | Format: Article | 5 min read

In brief

The recording of data has seen enormous growth for several years, due to the evolution of high technology devices. The data is often large and is termed as "Big Data".

The recording of data has seen enormous growth for several years, due to the evolution of high technology devices. The data is often large and is termed as "Big Data". So, this Big Data is put into use for further making decisions for the businesses. Analysis of Data is performed and insights are taken out from them, with the help of Data Science. That is why there is a high demand for data scientists all over the world. The smart devices make use of machine learning algorithms to predict human preferences and give them personalized recommendations based on their past actions. We can say that it makes the devices work smarter! Due to the high demand, data science offers high salary packages because there is a low supply of skilled data scientists. According to a report, in 2020, the job requirements for data science and analytics are projected to boom to by 364,000 openings to 2,720,000. And according to the U.S. Bureau of Labor Statistics, 11.5 million new jobs will be created by the year 2026. Data science is both preferred due to its high demand and because it comprises various skills such as machine learning, statistics, programming, computer science, etc. So it gives an opportunity for people to grow in multiple areas. This explains the growing demand for data scientist certification among professionals. This article explains about data science and what are the skills required to become a data scientist.

Who is a Data Scientist?

A data scientist is a professional who extracts, visualizes, analyzes, manages, and stores data with the help of advanced techniques, to take out some meaningful information from Big Data. Extracted information has proved to be helpful in making important business decisions for the future. Data science comprises mathematics, statistics, programming languages. Also, it consists of both structured and unstructured data. Data science has vast applications in the industries of healthcare, travel, marketing, sales, credit & Insurance, automation, social media. It is known to boost business performance.

Skills Required to Become a Data Scientist

Mathematics and Statistics

The concepts of mathematics and statistics are used for predicting the outcomes by using algorithms. Data is mainly analyzed, compared and insights are taken from them. Therefore, some of the concepts of statistics that are used majorly are descriptive statistics (mean, mode, median, etc), Bayes theorem, skewness, cumulative distribution function), calculus, matrices, linear algebra, and optimization methods

Programming skills

The programming languages are the roots of data science. The languages most popularly used for data science are Python and R. algorithms for data science are generally written with the help of these programming languages. So, gain expertise in any one of those.

Data Wrangling and processing

Data wrangling is a process of preparing the data for further use. It often happens that the data is unclear, unprocessed, and unstructured. This process tries to map one data from another to take out conclusions from it.

Data Visualization

Data visualization is a method of presenting data graphically to make it more easy and convenient to remember. Often when data is presented as actual numbers or statistics, it becomes confusing to understand. Data visualization is majorly used to compare data sets. Graphic representations such as histograms, pie charts, bar graphs, line charts, scatter plots, 3-D plots, time series, etc. The most important thing is to remember the right time at which these tools can be used otherwise it will fail to convey the right information to people. Another important thing is to understand the audience in which it is going to be presented. Some of the tools used for visualization are Tableau, google charts, kibana, Data wrapper.

Machine Learning

Machine learning is an important aspect of data science. It is used mainly for finding patterns and insights from large data. This is used when decision making is mostly dependent on the data. So some of the main concepts of machine learning that should be well aware are Random Forests, Naïve Bayes, Regression Models, K- nearest neighbours, and Clustering. The platforms used for this purpose are Python, Tensorflow, Keras.

Database Management

Since the work of a data scientist is done mostly on the data, it is important for the data scientist to be well aware of the database management systems(DBMS). The professional must know how to extract the right data and arrange them or edit them. SQL and MySQL are the languages used mainly for databases.

Cloud Computing

Cloud computing has been in trend since the past few years. Cloud computing offers users to store, analyze, and visualize the data. So, the data scientists must be well aware of cloud computing. Data Acquisition, Data Wrangling, Data Visualization, transferring, etc. are the applications that cloud computing offers. The platforms used mainly for cloud computing are Windows Azure, Google Cloud, Amazon Web Services, and IBM Cloud.

MS Excel

Excel is one of the basic platforms one must know well. Excel is used to store and manipulate large amounts of data. Though excel is not advanced as compared to other tools, it is still the basic platform that offers advantages such as manipulating and representing 2D data, excel sheets linked to python algorithms, data analytics, etc.

Soft Skills

Where hard skills demonstrate your knowledge about the subject, it is soft skills that ensure that you perform your work effectively. It is important that the professional must be passionate about his work. He must be a leader and a team player who is ambitious for the work. Also, good communication skills are a must. The professionals must have ethical and moral values which will be beneficial in the long run also.

Conclusion

Data Science is seeing growth since the past decade. The organizations have understood the importance of data and therefore they are now relying on data scientists for data-driven decision making. Applications of data science apart from technology are also in the industries of healthcare, agriculture, Aviation, cybersecurity, and manufacturing. According to Payscale, The average salary of a data scientist in India is rupees 6,98,413. If you have a keen interest in learning different aspects of the work, then data science is the one. Graduates or professionals from technical or management backgrounds must pursue it. So what are you waiting for? Start your career in Data Science by taking up a course that will make you industry ready for dynamic business environments. This will let you dig deeper into the concepts and apply them in real-life scenarios. Enter the path of data science which offers exciting career opportunities. It offers roles for entry as well as senior levels. Also read

- Are communication skills really important to become a Data Scientist?
- Now Is The Perfect Time To Up Your Content Marketing Game
- Best Strategies for Outsourcing Software Development

References

1. forbes.com - sites / louisacolumbus - <https://www.forbes.com/sites/louisacolumbus/2017/05/13/ibm-predicts-demand-for-data-scientists-will-soar-28-by-2020/>
2. simplilearn.com - big-data-and-analytics / data-scientist-certification-sas-r-excel-training - <https://www.simplilearn.com/big-data-and-analytics/data-scientist-certification-sas-r-excel-training>
3. builtin.com - data-science / data-science-applications-examples - <https://builtin.com/data-science/data-science-applications-examples>
4. payscale.com - research / IN - https://www.payscale.com/research/IN/Job=Data_Scientist%2C_IT/Salary