What is Data Science?

Before answering this question, the question that we should be asking is, why do you need data science and if you are considering a career in the field, what should you expect before diving into it?

A data scientist as the name suggests is a scientist. And a scientist by definition discovers as opposed to an engineer who builds. So if you are someone who likes to make discoveries and gather insights through data, you might find data science to your liking.

This article will now explain the differences between common buzzwords heard in data based jobs:

1. Data Engineer

Data Engineers are responsible for obtaining data from IOT devices, 3rd party APIs etc. and keeping the data in a warehouse. The skills needed for the job are advanced scripting and understanding on how devices work.

The role of a data engineer is as versatile as the project requires them to be. It will correlate with the overall complexity of a data platform. If you look at the *Data Science Hierarchy of Needs*, you can grasp a simple idea: The more advanced technologies like machine learning or artificial intelligence are involved, the more complex and resource-heavy data platforms become.

1. Data Scientist

Data Scientist is the one who analyses and interpret complex digital data. While there are several ways to get into a data scientist’s role, the most seamless one is by acquiring enough experience. These skills include advanced statistical analyses, a complete understanding of machine learning, data conditioning etc.

1. Data Analyst

Most entry-level professionals interested in getting into a data-related job start off as **Data analysts**. Qualifying for this role is as simple as it gets. All you need is a bachelor’s degree and good statistical knowledge. Strong technical skills would be a plus and can give you an edge over most other applicants. Other than this, companies expect you to understand data handling, modeling and reporting techniques along with a strong understanding of the business.

The skills required to become data scientist are:

1. Mathematical skills

* Statistics
* Probability
* Linear Algaebra
* Multi-variable Calculus
* Graph Theory

1. Programming skills

* Python (Numpy, Pandas, Matplotlib, Tensorflow, Scipy)
* R

1. Computer Science skills

* Data Structures and Algorithms(basics)
* Databases (SQL)
* Web scrapping (HTML and Javascript)

1. Technologies

* Machine Learning
* Deep Learning
* Computer Vision
* Natural Language Processing
* Time Series Analysis

1. Visualization

* Microsoft Excel
* Power BI
* Tableau
* Python (Matplotlib)
* R (ggplot2)

References:

1. [Edureka](https://www.edureka.co/blog/data-analyst-vs-data-engineer-vs-data-scientist/)
2. [Krish Nayak (YouTube)](https://www.youtube.com/watch?v=017B07EHe2M)
3. [Tina Huang (YouTube)](https://www.youtube.com/watch?v=JUkSeqe1eYQ)