



Amsource Technology [Follow](#)

IT / Technology Recruitment Consultancy. Passionate about connecting people, talent and opportunity in technology.

Dec 6 · 3 min read

## Infographic: Becoming a Data Scientist

In this infographic we break down some of the key skills candidates should be adopting if they are looking to enter the world of Data Science...

### What does it take to be a **Data Scientist?**

The explosion of Data Science cannot be underestimated. It is THE technology area organisations are developing and embracing, and the Harvard Business Review called it “the sexiest job of the 21st century”. So what does it take to be a Data Scientist?

#### Education

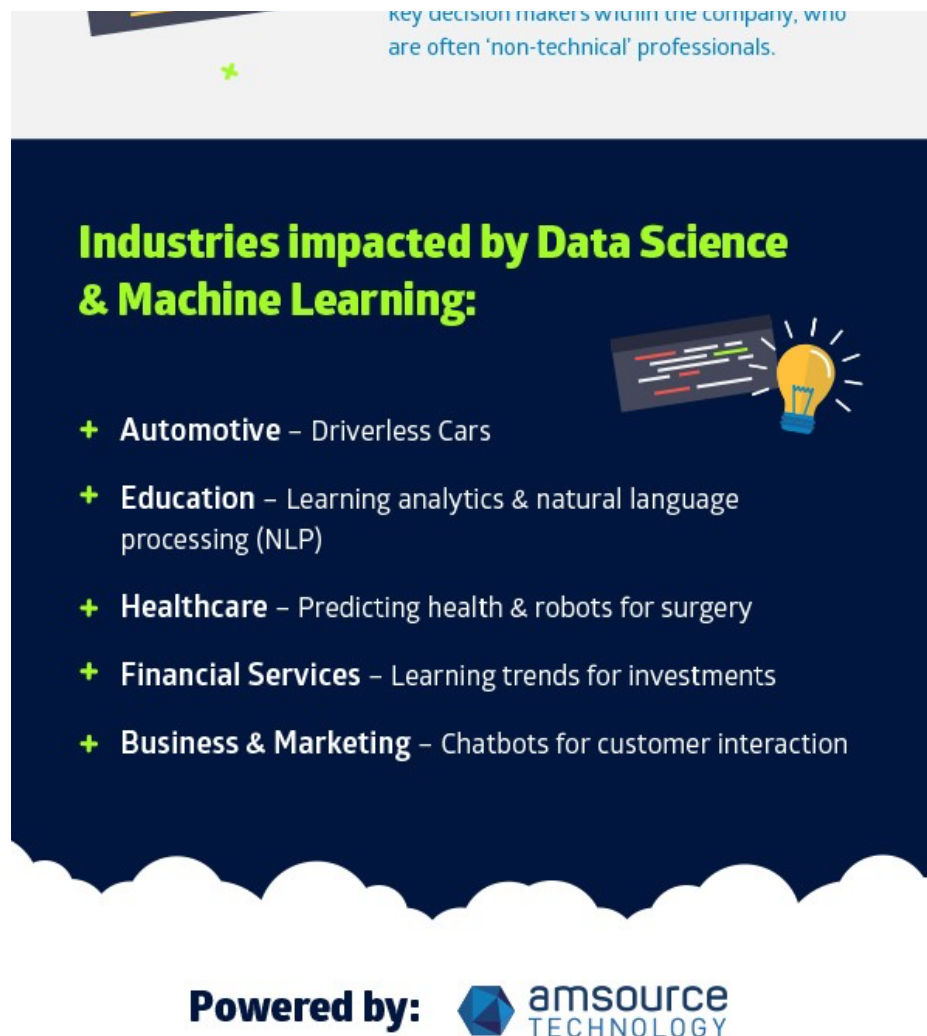
Ideally Masters or PhD educated in an Applied Mathematics based subject. Some universities have started offering Bachelor's degrees in Data Science.

#### Toolkits & Technologies

- + Python / R / SQL
- + Numpy / SciKit Learn / Pandas / Tensor Flow PySpark / Keras
- + Hadoop, Spark, Scala







### What is Data Science and Machine Learning?

By comparison **Data Science and Machine Learning** is a relatively new phenomenon that has swept the professional world over the last 10 years. Despite this, the statistical theory that underpins it has been around for far longer. Put simply, key advancements in technology such as programs like Hadoop, have allowed us to process vast amounts of data and analyse it through various machine learning methods. Prior to these advancements, the processes involved with Data Science were possible, but ultimately restricted by human shortcomings.

In a world where machines can now learn for themselves based on patterns in past data sets we find ourselves with the emergence of a new profession that is taking the commercial world by storm. Companies are now able to utilise this highly detailed insight to draw competitive advantage that can add real value like nothing else.

## **What skills does a Data Scientist need?**

There is some contention over what makes a Data Scientist and exactly what skills and experience they need to have to be qualified in this area. I also frequently get asked the question, ‘How does a graduate from a mathematical background make the step into the world of Data Science?’

Firstly, the difference between Data Analytics and Data Science is that analytics is about being descriptive. It’s about describing what the data is saying. The crucial element of Data Science however, is about building models to make predictions at the individual level. It’s these predictions that allow companies to better focus their approaches in marketing, finance, fraud detection, AI chatbots and so on...

With regard to getting into Data Science, there are a number of different pathways which someone could take. Usually, Data Scientists have a PHD/ Masters or higher in Computer Science, Mathematics, Neuroscience or related subjects. I have found that graduates who have also obtained a MSc in Data Science, especially those that include large amounts of Machine Learning elements, tend to have the experience required to move into the area within industry.

In addition to this, taking the initiative to complete relevant internships throughout your education will only help you to gain the skills required. Another important trait of a Data Scientist is that they have the personal traits to communicate across a number of levels. Often, they are required to translate their findings to non-technical people that use their insights to make business decisions. So, having said that, it’s important to have a business savviness that allows you to think in terms of a wider business objective.

**Experience and confidence with some of the following techniques are commonly required on many Data Science job specs...**

- **Toolkits and Technologies**

- Python / R / SQL
- Numpy / SciKit Learn / Pandas / Tensor Flow / PySpark / Keras
- Hadoop, Spark

- **Machine Learning Models and Algorithms**

- Linear regression / Logistic regression
- Support Vector Machines

- Decision Trees / Random Forest
- Neural Networks
- Naïve Bayes
- K-NN / K means clustering
- Natural Language Processing / Deep Learning / Computer Vision

