

## What is Data and Data Science?

Course  
Level:



Number  
of days:



Delivery method:  
On-site



### Course aims

The aim of this course is to explain what is meant by data, big data and data science.

The course explains how data has evolved from simple marks on bones (43000 BCE) into today's big data, discussing the challenges, successes and failures of big data. To highlight the challenges we discuss a particularly (in)famous big data failure by Google.

The course also provides a contemporary introduction to data science, explaining its origins from datalogy, computer science and statistics, to its establishment as a stand-alone field and the emergence of the data science profession in the early 2000s. We then explore the work of data scientists and consider the skills a good data scientist should have.

The course concludes by considering two data science applications: graph networks and how they can detect financial fraud, and Natural Language Processing and the challenges involved understanding human language.

### Learning Outcomes

#### 1. Understand the core issues of data and big data.

- 1.1 Explain what is meant by the terms **data** and **big data**.
- 1.2 Explain the Vs of **big data**.
- 1.3 Understand the challenges and criticisms of **big data**.

#### 2. Understand the core issues of data science.

- 2.1 Understand what is meant by the terms **data science** and **data scientist**.
- 2.2 Understand how **data science** is related to other academic fields.
- 2.3 Understand the skills of a good **data scientist**.
- 2.4 Understand the challenges and applications of **data science**.

