



LLOYD'S

# Fundamentals of Data Science (Technical)

---

## Course Syllabus and Learning Outcomes

 [Lloyd's Courses](#)

2020

## CONTENTS

### COURSE OVERVIEW

03

### COURSE STRUCTURE

- Weeks 1 - 8

04

### AIMS AND LEARNING OUTCOMES

- Hands-on Experience

- Technology Stack used for Exercises and Assignments

05

### MODULES AND TOPICS

- Module 1: Welcome and Course Information

- Module 2: Introduction to Core Concepts and Technologies

- Module 3: Data Collection and Management

- Module 4: Data Analysis

- Module 5: Data Visualisation

- Module 6: Module Reflections on the Future of Data Science

- Week 7 & 8: Hands-on Practice - Interactive Visualisation

06 - 11

# Course Overview

## Fundamentals of Data Science - Technical

**Duration:** 8 weeks

**Number of hours:** 40 (includes tutorial support + self-study, assignments and peer discussions)

**Assignments:** Three assignments (each worth 33% of the final mark)

This course equips you with the theoretical knowledge and both practical and technical skills to participate in the flourishing data revolution, helping you to contribute to and benefit from the new data-driven economy. The course emphasises a hands-on approach to learning data skills, offering a number of interactive, online exercises that will let you try out many of the techniques and concepts covered in the taught material.

The course is broken into eight weeks.

# Course Structure

## Week 1

In Week 1, you will meet your tutor and the other participants on this course and find out more about what you will be doing over the next eight weeks and how we will be supporting you. You will get "hands-on" experience of **Jupyter**, the web-based learning environment which you will use for the course exercises and assignments. This week also contains a **Python Primer** activity for those of you who are unfamiliar with the programming language or would like a refresher.

## Week 2

In Week 2, you will learn about the fundamental terminology and processes in Data Science, discovering the technology landscape that has helped fuel the data explosion, and the tools that data scientists use to unlock the hidden value in these vast amounts of data. This week also contains an introduction to using Python for Data Science.

## Week 3

You will begin gaining hands-on experience of Data Science in Week 3, focusing on collecting, storing and managing data. You will learn about the different sources of data and how they can be combined in order to increase the potential insights available.

## Week 4 & 5

Week 4 and 5 will help you understand how this data is analysed, covering a range of techniques that any Data Science team will encounter, from statistics to machine learning. You will use Python to analyse some given data.

## Week 6 & 7

In Week 6 and 7, you will learn about how the findings from Data Science work can be reported using different data visualisation techniques. You will discover the various ways in which particular types of data can be displayed in order to highlight a key finding and improve the impact of your reports.

## Week 8

In Week 8, you will look at the future of Data Science and we will support you to finish off your assignments.

# Aims and Learning Outcomes

This module aims to provide you with the knowledge and expertise to become a proficient data scientist.

Having successfully completed this module, you will be able to:

- Understand the key concepts in Data Science, including their real-world applications and the toolkit used by data scientists
- Explain how data is collected, managed and stored for Data Science
- Implement data collection and management scripts using NodeJS and MongoDB
- Demonstrate an understanding of statistics and machine learning concepts that are vital for Data Science
- Produce Python code to statistically analyse a dataset
- Critically evaluate data visualisations based on their design and use for communicating stories from data
- Plan and generate visualisations from data using Python and Bokeh

## Hands-on Experience

**Week 1** includes an (optional) introduction/'refresher' on **Python**. This includes online exercises for you to work through at your own pace. This not graded.

**Week 2** contains further **Python** practice exercises. You are encouraged to do these as these will help you with your assignments. Again, these practice exercises are not graded.

**Weeks 3 - 8** each include online exercises (ungraded) and a related graded coursework assignment.

## Technology Stack used for Exercises and Assignments

<b>Visualising</b>	Bokeh (Python)
<b>Stats / Analysis</b>	NumPy / SciPy
<b>Management / Querying</b>	MongoDB (using Python)
<b>Base</b>	Python

# Modules and Topics

## Module 1: Welcome and Course Information

### TOPICS

- Welcome and introduction
- Learning outcomes of the week
- What Data Science is and why it is important
- Course syllabus and learning outcomes
- Using discussion forums
- Introduce yourself
- Help and tutoring support
- Course assignment details
- A hands-on Jupyter familiarisation activity
- Python Primer
- Glossary of terminology

### LEARNING OUTCOMES

- Describe the concepts of 'big data' and 'open data'
- Identify and explain the core concepts of the Data Science Pipeline
- Utilise Python programming language to perform basic data management and interrogation operations

## Module 2: Introduction to Core Concepts and Technologies

### TOPICS

- Introduction
- Learning outcomes of the week
- Data Science in a nutshell
- Terminology
- The Data Science process
- A Data Science toolkit
- Types of data
- Example applications
- Further reading
- Summary

### LEARNING OUTCOMES

- Describe what Data Science is
- Explain why Data Science is different to other disciplines, including Statistics and Computer Science
- Summarise the Data Science process and the steps involved
- Reflect on what counts as Data Science
- Classify types of data from a range of sources
- Summarise the broad skillset required for doing Data Science

## Module 3: Data Collection and Management

### TOPICS

- Introduction
- Learning outcomes of the week
- Sources of data
- Data collection and APIs
- Exploring and fixing data
- Data storage and management
- Using multiple data sources
- Further reading
- Summary

### LEARNING OUTCOMES

- Explain the different sources of data for use in Data Science applications
- Compare the different licensing options for data
- Describe the role of APIs for data collection
- Explain the role of data cleansing
- Apply data collection techniques to store and manage data in MongoDB
- Assess the benefits and issues with collecting data from multiple sources
- Write queries to extract data from MongoDB



## Module 4: Data Analysis

### TOPICS

- Introduction
- Learning outcomes of the week
- Terminology and concepts
- Introduction to statistics
  - Nature of statistics and introduction
  - Central tendencies and distributions
  - Variance
  - Distribution properties and arithmetic
  - Samples/CLT
- Basic machine learning algorithms
  - Linear regression
  - SVM
  - Naive Bayes
- Further reading
- Summary

### LEARNING OUTCOMES

- Describe the role of statistics in Data Science
- Solve statistical problems using essential techniques for Data Science
- Produce Python code to programmatically apply knowledge of statistics
- Describe the purpose of machine learning and the different types of algorithm
- Compare three basic machine learning algorithms

## Module 5: Data Visualisation

### TOPICS

- Introduction
- Learning outcomes of the week
- Types of data visualisation
  - Exploratory
  - Explanatory
- Data for visualisation
  - Data types
  - Data encodings
  - Retinal variables
  - Mapping variables to encodings
  - Visual encodings
- Technologies for visualisation
  - Bokeh (Python)
- Further reading
- Summary

### LEARNING OUTCOMES

- Explain the purpose of data visualisation and its role within Data Science
- Explain the differences between exploratory and explanatory visualisations
- Compare different data types and the sorts of visualisation they can be used in
- Describe the process of encoding data in a visual, and the different encodings that can be used
- Evaluate the effectiveness of particular data encodings
- Produce visualisations using Python (Bokeh)

## Module 6: Module Reflections on the Future of Data Science

### TOPICS

- Introduction
- Learning outcomes for the week
- The future of Data Science

### LEARNING OUTCOMES

- Explain current technical and ethical challenges for Data Science
- Describe possible future directions for the field of Data Science

## Hands-on Practice - Interactive Visualisation (Week 7 & 8)

### TOPICS

- Analysis and visualisation with Python

### LEARNING OUTCOMES

- Implement the Data Science Pipeline process efficiently utilising the Python programming language

# Ready to start?

Visit [our website](#) to secure your place now.

If you have any questions about this course or the other courses available for Lloyd's marketplace professionals, call **+44 (0)12 2344 7775** to speak to a member of our team.

For data science news and insights follow Southampton Data Science Academy on:

