Imperial College London

University of the Year 2022 by The Times and Sunday Times' Good University Guide presents

Summer Online Programme

An Introduction to Data Science: Concepts and Applications

Learn data science without coding!

PROGRAMME OVERVIEW

Data science courses seem to be everywhere but what about an avenue to learning data science for the non-specialist? Here we deliver the solution with a Non-Technical Introduction to Data Science which covers concepts and applications in an approachable way, without coding nor previous experience required.

Data science is not only for coders and we feel strongly that knowledge of data science concepts is for all of us. Understanding these opens up the world of data-driven decision making which benefits both our lives and our careers by using data analysis to support decisions even when we are not the ones coding it. We want to make sure everyone can join the discussion on data science and that this engagement is not done only by the specialists. We bring the topics to you in an accessible, approachable way, without the need to understand coding or its syntax.

This master class will provide participants with an understanding of these topics, addressing concepts of data science and the applications where it is making an impact.

Topics covered include:

- **What is data science?**
  To introduce data science and explain how to break down a problem such that a computer can understand and make sense of it. Humans make decisions using judgment and intuition, but computers need problems presented in binary format.
- **What is Machine Learning?**
  To introduce the concepts of supervised & unsupervised machine learning, as well as reinforcement learning. To understand the different types of input data for those.
- **Real world applications of machine learning**
  To learn about applications of machine learning across different industries such as climate, drug discovery and robotics.
- **What are machine ethics and why does it matter?**
  To introduce the concepts of machine ethics and to understand how misrepresentative data can alter your outcomes with serious consequences
- **Simple steps for preparing a dataset**
  To understand the steps that need to be performed on a raw dataset to make it usable for analysis

Team based learning via group project:

As part of this programme, students will have the opportunity to work in small project teams to explore a novel idea for a machine learning application that builds on concepts learned in this course. No coding is required.

On completion of this programme, students will:

- Understand what data science and machine learning are.
- Understand the concept of Computational Thinking and why it is relevant.
- Describe three main types of machine learning.
- Understand applications of machine learning and the impact it has.
- Understand the importance of machine ethics, data bias and its consequences.
- Understand the steps necessary to clean a raw dataset for analysis.
- Design and develop an idea for a machine learning application.

PROGRAMME STRUCTURE AND TEACHING METHODS

- 5 live lectures delivering via Microsoft Teams with group discussions and questions.
- Web based exercises and quizzes will be provided for formative feedback.
- Group projects for assessing the learning outcomes, supported by 3 hour tutorials.
- 1 session on final day for project presentation.

Project work will be done through team-based learning and tutorials. Final projects will be presented in groups on the last day of the programme. A prize will be awarded to the team with the best project.

The programme will be delivered over Microsoft Teams. Online project channels will be allocated to each team for project work. Students will be able to use the channel at any time to work on their project.

The entire programme will be taught in English.

CERTIFICATION

Students will receive a verified Imperial College London digital certificate on successful completion of this programme and a prize will be awarded to the best project team. Each student will also receive a transcript for their project marks.

ENTRY REQUIREMENTS

All students are expected to be studying an undergraduate or a postgraduate degree in a **non-technical subject** at a well-recognised university in China. This is for those with **NO** coding, programming nor technical experience, yet for those wishing to learn data science and its concepts and applications.

English requirements:

All students are required to have a good command of English, and if it is not their first language, they will need to satisfy the College requirement as follows:

- a minimum score of IELTS (Academic Test) 6.5 overall (with no less than 6.0 in any element) or equivalent.
- TOEFL (iBT) 92 overall (minimum 20 in all elements)
- CET- 4 (China) minimum score of 550
- CET- 6 (China) minimum score of 520

Students will need to have access to a computer with a webcam, microphone and good internet connection to attend the live classes.

PROVISIONAL SCHEDULE

Session 1: What is data science?
Content • What is data science? • Computational thinking • How humans process information • How computers need problems broken down • Interactive exercise on computational thinking • Q&A
Session 2: What is machine learning?
Content • What is machine learning? • Supervised machine learning • Unsupervised machine learning • Reinforcement learning • Interactive demonstrations of supervised machine learning • Q&A
Session 3: Real world applications
Content Applications of Machine Learning in: • Climate • Drug discovery • Robotics • Q&A
Session 4: What are machine ethics & why does it matter?
Content • Machine ethics • Data bias • Interactive demonstration of supervised machine learning with misrepresentative data • Consequences • Q&A
Session 5: Simple steps for preparing a dataset
Content • Steps needed to make a dataset usable for analysis • Data Checklist- what is needed? • Q&A
Session 6: Project Presentations & Career Pathways Discussion
Content • Group project presentations

- Q&A and feedback
- Announcement of winning project group
- Discussion on career pathways

THE PRESENTER

Dr Susan Mulcahy
Director of the Data Spark Programme
Imperial College London

Dr Susan Mulcahy is the Director of the Data Sparks Programme, the innovative student placement programme matching a real world industry project on data science with a team of our postgraduate students. This programme sits within Imperial Business Analytics, the research centre focused on bringing data science research closer to the world of business. Susan was previously the Senior Education Fellow of the Data Science Institute (DSI) at Imperial where she developed the educational offering of the DSI for internal students and external industry engagements. She is also a Lecturer in Data Analytics at Ada National College for Digital Skills. Having also facilitated technical courses for corporate clients since 2013, Susan enjoys teaching/facilitating/presenting technical topics to a general audience.

Susan received her data-driven PhD from Imperial's Bioengineering Department in 2016 where she researched indicators of traumatic brain injury using MATLAB on datasets collecting over 500 million data points per patient per day. In addition to this, she has an MBA from INSEAD in France and a BSc in Mechanical Engineering from Purdue University in the USA. Susan has been a Fellow of the Royal Geographical Society since 2002.

For outside interests, Susan seeks out adventure. In 1999, she spent three months riding her bicycle across the USA. These days, she can be found rowing weekly on the Thames (anything from singles to 8s, outside of lockdown), hiking up a rugged mountain in the Scottish Highlands, or sleeping in a tent in her back garden in London (which she did for 82 consecutive nights in lockdown v1.0 in search of a local adventure.)

Organsied by Imperial's Centre for Continuing Professional Development
www.imperial.ac.uk/cpd