Foundations of Data Science

When you'll study it
Semester 1

CATS points
15

ECTS points
7.5

Level
Level 7

Module lead
Adriane Chapman

Academic year
2023-24

On this page

Syllabus

The course will introduce students to the data scientist toolkit and the underlying core concepts. It will cover the full technical pipeline from data collection (sampling methods, crawling) to processing and basic notions of statistical analysis and visualization. The module will also include advanced topics in high-performance computing, including non-relational databases and MapReduce. By taking this course the students will be provided with the basic toolkit to work with data (CSV, R, MongoDB). To support these learning objectives, the coursework will include exercises and a group project in which students will use existing open data sets and build their own application.

The course will cover the following concepts:

- Fundamentals and core terminology
- Technology pipeline and methods
- Application scenarios and state of the art
- Data collection (sampling, crawling)
- Data analytics (statistical modeling, basic concepts, experiment design, pitfalls, R)

Data interpretation and use (visualization techniques, pitfalls, D3)

High-performance computing (parallel databases, MapReduce, Hadoop, NoSQL)

Cloud computing (principles, architectures, existing technologies)

Back
Aims and Objectives

Next
Learning and Teaching

Part of the

Information

For visitors

For staff & students

For schools & colleges

For researchers

For employers & recruiters

For parents & guardians

For international students

Contact us

Tel: +44(0)23 8059 5000
Fax: +44(0)23 8059 3131

University of Southampton
University Road
Southampton
SO17 1BJ
United Kingdom

Get directions →

Connect with us

Download a copy of our prospectus or order a printed copy to be delivered to your door.

Get a prospectus →

Site map

Accessibility