Meet your programme
Data Science MSc

Dr Paolo Guagliardo
(programme director)

Academic Year 2021/2022
Data

The **most important asset** of any enterprise

⇒ Enable and **support decision making**
Data

The **most important asset** of any enterprise

Must be **effectively, efficiently and reliably**

- collected and stored
- maintained and updated
- processed and analysed

to be **turned into meaningful information**

⇒ Enable and **support decision making**
Different kinds of data

Relational databases
Data organised in tables (relations) with typed attributes

Document stores
Text documents structured using tags (or other markers)

Graph databases
Data organised in graph structures with nodes and edges

Key-value stores
Data organised in associative arrays (a.k.a. dictionaries or maps)
Access, integration and manipulation of large volumes of data are at the core of data analysis.
Skills needed to become a Data Scientist

Skills needed to become a Data Scientist

Skills needed to become a Data Scientist

https://365datascience.com/skills-data-scientist/
Technology stack

**Python + cloud platform** (such as AWS or Azure)
+ **ML framework** (such as PyTorch or TensorFlow)

Data experience

Analysis of datasets with
- numerical frameworks such as MATLAB or R
- **Python libraries** such as NumPy, SciKit, Pandas, etc.

Environment

Unix/Linux, familiarity with **shell scripting** (e.g., Bash)
Skills overview

**Hard skills**
- Training and evaluating machine learning models
- Proficiency querying data with SQL
- A strong understanding of probability
  - statistical and systematic uncertainty

**Soft skills**
- Confidence in communication
  with both technical and non-technical audiences
- Self-motivation in the face of challenges
Programme structure [http://www.drps.ed.ac.uk/21-22/dpt/ptmscdatsc1f.htm]

**Mandatory courses** (80 credits)

- Informatics Research Review (IRR) 10 credits
- Informatics Project Proposal (IPP) 10 credits
- MSc Dissertation 60 credits

**Elective courses** (80 credits)

Selection from 3 areas:

- Machine Learning, Statistics and Optimization 10-60 credits
- Databases and Data Management 10-60 credits
- Applications 10-60 credits

**Other courses** (20 credits)

- Level 10/11 courses (can be from schools outwith Informatics)

Check out the programme builder in Path: https://path.is.ed.ac.uk/builder/PTMSCDATSC1F