



Data Science Fundamentals Course Descriptor

Course Title	Data Science Fundamentals	Faculty	EDGE Innovation Unit (London)
Course code	NCHMV465	Course Leader	Marta Rodriguez-Martinez
Credit points	30	Teaching Period	12 weeks
FHEQ level	4	Date approved	March 2021
Compulsory/Optional	Compulsory		
Prerequisites	None		
Co-requisites	None		

COURSE SUMMARY

This course will introduce the Python programming language and its use in data science. Learners will start by understanding the basics of Python programming and the use of a version control software. Learners will be taught how to use Python to clean and visualise data as well as develop basic predictive models. Learners will identify potential predictive relationships between variables, and apply linear and logistic regression models to validate their hypotheses. They will learn how to evaluate, optimise and interpret these models. Learners will also understand the fundamentals of natural language processing and how to manipulate time series data using Python libraries.

COURSE AIMS

- Train learners to use version control software.
- Train learners to write basic Python scripts and employ libraries to manipulate data.
- Train learners to build basic predictive models such as linear and logistic regressions.
- Introduce learners to the fundamentals of natural language processing and time series analysis.

LEARNING OUTCOMES

On successful completion of the course, learners will be able to:

KNOWLEDGE AND UNDERSTANDING

- K2a Understand the foundations of linear and logistic regressions.
- K2a Understand the foundations of natural language processing and time series analysis.
- K2a Describe the use of descriptive and inferential statistics.

SUBJECT SPECIFIC SKILLS

- S1a Collate data for analysis using Python.
- S2a Develop predictive models using linear and logistic regression.
- S3a Apply the scientific method to test hypotheses.

TRANSFERABLE AND PROFESSIONAL SKILLS

- T2a Identify correct tools/methods to solve a problem at hand.
- T2a Apply analytical thinking to test a hypothesis.
- T3a Select visualisations to present findings accurately.

TEACHING AND LEARNING

The course learning and teaching hours will be structured as follows:

- Seminar = 30 hours
- Applied Seminar = 15 hours
- Coaching session = 2 hours
- Academic Drop-ins = 8 hours
- Reading and Assessment Prep = 65 hours
- Application of learning = 180 hours

Total = 300 hours

ASSESSMENT

FORMATIVE

Learners will be formatively assessed during the course through a range of techniques. During seminars, applied learning seminars and one to one coaching, questioning and quizzing are used as formative assessment. The purpose of the formative assessment is to provide learners with developmental feedback to support their learning.

SUMMATIVE

AE	Assessment Type	Weighting	Online submission	Duration	Length
1	Set exercise	30%	Yes	Requiring on average 5-15 hours to complete	1500 words +/- 10%
2	Set exercise	50%	Yes	Requiring on average 15-25 hours to complete	2000 words +/- 10%
3	Oral	20%	No	30 mins +/- 10%	

The summative assessment will be assessed in accordance with the assessment aims set out in the Programme Specification.

FEEDBACK

Learners will receive formal feedback in a variety of ways. For summative assessments, feedback is written and provided within 20 days of submission, and for formative assessment it could include verbal or written feedback. Learners will also have a progress review each quarter, see above, with their Multiverse Coach and their Line Manager. These reviews are an opportunity to discuss progress to date both in their role and on the degree programme.

INDICATIVE READING

Note: Comprehensive and current reading lists for courses are produced annually in the Course Syllabus or other documentation provided to learners; the indicative reading list provided below is used as part of the approval/modification process only.

BOOKS

Schneider, D.I., 2015. *An Introduction to Programming Using Python*. Pearson.

Boschetti, A, 2018, *Python Data Science Essentials*, Packt

ELECTRONIC RESOURCES

Li, S. 2017. Building A Logistic Regression in Python, Step by Step, *Towards Data Science*, <https://towardsdatascience.com/building-a-logistic-regression-in-python-step-by-step-becd4d56c9c8>

Lopez Yse, D. 2019. Your Guide to Natural Language Processing (NLP), *Towards Data Science*, <https://towardsdatascience.com/your-guide-to-natural-language-processing-nlp-48ea2511f6e1>

INDICATIVE TOPICS

Learners will study the following topics:

- Version control software
- Descriptive and inferential statistics
- Introduction to Linear regression
- Introduction to Logistic regression
- Introduction to Natural Language Processing
- Basic manipulation of time series data

Title: NCHMV465 Data Science Fundamentals Course Descriptor					
Approved by: Academic Board					
Version number	Date approved	Date published	Owner	Location	Proposed next review date
1.0	March 2021	March 2021	Marta Rodriguez-Martinez	Multiverse website	March 2026
Modifications (As per AQF4)					
Version number	Date approved	Date published	Modification (including category number)		