New College of the Humanities

Foundations of Data Science

Course Title	Foundations of Data Science	Faculty	Philosophy
Course Code	NCHAI759	Course Leader	
Credit points	15	Teaching Period	Hilary
FHEQ Level	Level 7	Date Approved	June 2020
Compulsory/ Optional	Compulsory		
Pre-requisites	Programming with Data		
Co-requisites	None		

COURSE SUMMARY

By the end of this course, students will have a thorough advanced working knowledge of working in an environment such as IPython for full-scale coding. The module covers the scientific programming environment in Python and advanced programming techniques and plotting.

COURSE AIMS

The aims of the course are to:

- Develop familiarity with an interactive computing and development environment
- Develop a basic understanding of arrays and vectorised computation
- Develop and manifest an elementary understanding of data structures and their functionality, and of methods of data transformation
- Be able to load and clean data sets, summarise and compute descriptive statistics, and plot and visualize data

LEARNING OUTCOMES

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

KNOWLEDGE AND UNDERSTANDING

K1d demonstrate critical awareness and knowledge of basic data science concepts

K2d demonstrate comprehensive knowledge of feasible operations on data and transformation on data

K3d show knowledge and understanding of plotting and visualising data

K4d demonstrate sound knowledge of applying techniques to a data case-study

SUBJECT SPECIFIC SKILLS

S1d engage critically in the theory behind the concepts taught in the class

S2d demonstrate originality in applying the data transformation and techniques in an appropriate manner to the chosen dataset

S3d identify and apply the correct choice of appropriate data transformation techniques

TRANSFERABLE AND PROFESSIONAL SKILLS

T1d demonstrate initiative in working independently, effectively, and to deadlines

T2d identify, transform, critically evaluate and plot accordingly from the dataset

T3d produce clear and concise and well documented code

TEACHING AND LEARNING

Teaching and learning strategies for this course will include:

- 10 X 1 hours of full-cohort lectures
- 10 x 2 hours of lab-based tutorials
- 2 hours of office hour per teaching week

Course information and supplementary materials are available on the College's Virtual Learning Environment (VLE).

At the end of Michaelmas and Hilary, students will attend Collections (formal meetings) in which they receive comprehensive and collated feedback about their performance over the term.

Students are required to attend and participate in all the formal and timetabled sessions for this course. Students are also expected to manage their directed learning and independent study in support of the course.

EMPLOYABILITY SKILLS

- Skills in writing and analysing complex code.
- Presentation skills in presenting code accordingly.
- Skills in organisation of written and coding discourse
- Skills in being able to read, understand and comprehend the code

ASSESSMENT

FORMATIVE

Formative assessment will build on the material taught in the classroom notebooks. The material will be in the form of end of session exercises and in some cases questions and answers. Oral explanations are also part of summative assessment.

SUMMATIVE

Students will be assessed during the course by means of set assignments. Assessment will be in two forms:

AE:	Assessment Activity	Weighting (%)	Online sub-mission	Coding	Notebook Submission
1	Coding Assignment	50%	No	Yes	Code and 2500 word explanation
2	Coding Assignment	50%	No	Yes	Code and 2500 word explanation

The examination will consist of two written coding assignments which the student will have to do to the set guidelines for coding. The written assignment will be assessed in accordance with the assessment aims set out in the Programme Specification.

FEEDBACK

Students will receive formal feedback in a variety of ways: written (including via email correspondence); oral (within one-to-one tutorials or on an *ad hoc* basis) and indirectly through discussion during group tutorials. Student's will also attend the formal meeting, at the end of Michaelmas and Hilary terms in which they will receive constructive and developmental feedback on their term's performance.

Feedback is provided on written assignments (including essays, briefings and reports) and through generic internal examiners' reports, both of which are posted on the College's VLE.

INDICATIVE READING

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

BOOKS

Joel Grus (2019), *Data Science from Scratch*, 2nd ed., O'Reilly: Boston.

Wes McKinney (2017), *Python for Data Analysis*, 2nd ed., O'Reilly: Boston.

Hadrien Jean (2020), *Essential Math for Data Science*, O'Reilly: Boston.

ELECTRONIC RESOURCES

Students can visit courses on Datacamp, Coursera and Udemy to watch videos on Python Programming.

INDICATIVE TOPICS

- IPython: An Interactive Computing and Development Environment
- NumPy Basics: Arrays and Vectorised Computation
- Pandas
- Data Transformation
- Summarising and Computing Descriptive Statistics
- Plotting and Visualisation
- Advanced Pandas - Data Aggregation and Group Operations.

Title: NCHAI759 Foundations of Data Science Course Descriptor **Approved by: Academic Board**					
Version number	Date approved	Date published	Owner	Location	Proposed next review date
1.0	June 2020	June 2020	Brian Ball	1. Academic Handbook > Course Descriptors 2. VLE	April 2025
Modifications (As per AQF4)					
Version number	Date approved	Date published	Modification (including category number)		