

Module Details	
Module Code:	DATA H4001
Module Long Title:	Data Analysis and Programming DRAFT
Banner Title:	
Version:	3
Indicative NFQ level:	Level 8
Available From:	Jan 2022 (January 2022)
Language of Instruction:	English
ECTS Credits::	5
ISCED Code:	0611 - Computer use
Module Type	
No Module study modes listed	
Current Coordinator::	JOHN BURNS
Module Coordinators:	JOHN BURNS (21 December 2021 to ---)
Please enter the School responsible for the module.:	Tallaght Campus (TA)
Campus:	Tallaght
Module Overview	This module provides the student with an understanding of state-of-the-art data analysis concepts and techniques, and the ability to develop solutions to big data problems using suitable algorithms and software.
Indicative Syllabus	
1. Data Analysis, Theory and Methods 1.1) Data Collection, Sampling, Descriptive statistics, Inferential statistics	
2. Data Theory, Concepts and Methods 2.1) Data bias, handling missing data, investigating outliers and erroneous data, data generation techniques. 2.2) Effects of data bias.	
3. Data Analysis & Scripting 3.1) Reading Data; Cleaning and Filtering Data; Data Pre-processing and attribute techniques for ML algorithms. 3.2) Visualisation of Data	
4. Public Cloud Solutions 4.1) Developing solutions using Public Cloud Providers, using analysis tools (eg Jupyter notebooks) with access to HPC architectures.	

Learning Outcomes		
Upon successful completion of this module the learner will be able to		
#		
MLO1	Be able to apply a range of appropriate statistical techniques to analyze a variety of data sets	
MLO2	Demonstrate an understanding of the theory, concepts and techniques for data processing	
MLO3	Demonstrate a detailed knowledge and understanding of data analysis scripting tools and techniques	
MLO4	Apply best practice when using solutions from providers such as public cloud providers	
Requisites		
Requisite Type	Module Title	Type
No requisites exist.		

Module Content & Assessment			
Assessment Breakdown		%	
Other Assessment(s)		100.00%	
Assessments			
No Formal Examination			
Other Assessment(s)			
Assessment Type	Practical/Skills Evaluation	% of Total Mark for Module	25
Indicative Week	Week 4	Learning Outcomes	1,2
Semester	Semester 1	Assessment Threshold:	None
Assessment Role	Not yet determined	Assessment Authenticity	Not Online
Pass/Fail	No		
Assessment Description			
Students will solve a specific problem(s) and describe how to implement and test the solution from a specific data set/problem context.			
Assessment Type	Project	% of Total Mark for Module	25
Indicative Week	Week 8	Learning Outcomes	1,4
Semester	Semester 1	Assessment Threshold:	None
Assessment Role	Not yet determined	Assessment Authenticity	Not Online
Pass/Fail	No		
Assessment Description			
A group project in which students will programme and implement an example of a data analysis problem with pre-processing and bias investigation.			
Assessment Type	Project	% of Total Mark for Module	50
Indicative Week	Week 12	Learning Outcomes	1,2,3,4
Semester	Semester 1	Assessment Threshold:	None
Assessment Role	Not yet determined	Assessment Authenticity	Not Online
Pass/Fail	No		
Assessment Description			
An individual project in which students programme and implement an example of a data analysis project for the complete life cycle, data, statistical analysis ideally using cloud infrastructure.			
Reassessment Requirement			
A repeat examination			
Reassessment of this module will consist of a repeat examination. It is possible that there will also be a requirement to be reassessed in a coursework element.			

Module Activity

Full Time hours per semester	
Activity Type	Duration (Hours)
Lecture	2
Lab	2
Self Directed	3
Hours (up to 100 for 5 ECTS credits)	7.00

Part Time hours per semester	
Activity Type	Duration (Hours)
Lecturer/Lab (Inactive)	4
Self Directed	5
Hours (up to 100 for 5 ECTS credits)	9.00

Recommended Reading List

Supplementary Book Resources

EMC Education Services. (2015), Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, Wiley.

D. Miner, A. Shook. MapReduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop and Other Systems, O'Reilly.

Ian H. Witten, Eibe Frank, Mark A. Hall, Christopher J. Pal. (2016), Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques (Morgan Kaufmann Series in Data Management Systems), 4. Morgan Kaufmann.

Jason Brownlee. (2017), Machine Learning Mastery With Python: Understand Your Data, Create Accurate Models, and Work Projects End-to-End, Machine Learning Mastery.

Sebastian Raschka, Vahid Mirjalili. (2019), Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2., 3rd. [ISBN: 978-1789955750].

This module does not have any journal article/paper resources

This module does not have any other resources

Essential Reading List

Review

Module Extra Information

Editor(s)

Staff Member

FINBARR FEENEY

JOHN CARDIFF

JOHN BURNS

Siobhan Jacob

Keith Quille

Rajesh Jaiswal

Affiliated Programmes

Programme Code	Programme Title	Programme Version
TA_KACOD_B	Bachelor of Science (Hons) in Computing with Data Analytics	2
TA_KACOD_B	Bachelor of Science (Hons) in Computing with Machine Learning/ Artificial Intelligence	3
TA_KCODA_B	Bachelor of Science (Hons) in Computing with Data Analytics (Add-On)	1
TA_KCODA_B	Bachelor of Science (Hons) in Computing with Machine Learning/ Artificial Intelligence	2
TU_HCAIM_M	Master of Science in Computing - Human Centered Artificial Intelligence (90)	1

Other Resources

No Links to Documents

statuslog

Initial Status	End Status	User	Date	Comment
	Draft	FINBARR FEENEY	21/Dec/2021 17:14	Update content for 2021 and MSc precursor