

DATA9009: Data Science Principles

Module Details

Module Code:	DATA9009
Title:	Data Science Principles APPROVED
Long Title:	Data Science Principles
NFQ Level:	Expert
Valid From:	Semester 1 - 2025/26 (September 2025)
Duration:	1 Semester
Credits:	5
Field of Study:	4816 - Data Format
Module Delivered in:	2 programme(s)
Module Description:	This module will provide the learner with an overview of the key themes, methods and technologies in the field of data science and analytics. An emphasis will be placed both on the real-world practical and the technical aspects of data science. The business, process and technical contexts in which data analytics is used in real world scenarios will be investigated with topics such as business use case scenarios, analytic team formation and delivery methodologies. The learner will gain an insight into statistical theory, mathematical algorithmic design and modelling concepts as well data architectures and data quality issues. Data analytics/mining software, e.g. R, Python, PowerBI, will be used in both the lectures and labs. GenAI methodologies and software, e.g. Microsoft CoPilot, Agentic AI, will also be investigated.

Learning Outcomes

On successful completion of this module the learner will be able to:

#	Learning Outcome Description
LO1	Appraise the field of data science and analytics, its concepts, technologies and its historical roots.
LO2	Evaluate the main approaches to creating and managing a data analytics/mining project. Explain the importance of ethical considerations, data protection and security.
LO3	Evaluate Exploratory Data Analysis (EDA) results and appraise which modelling method is most suitable.
LO4	Determine how to test and evaluate a model. Interpret the results of a variety of testing and evaluation techniques.
LO5	Explain how a variety of statistical models work and how they are optimised.
LO6	Appraise business situations using enhanced data literacy skills - students will gain the competency to confidently understand, interpret, and utilise data, empowering them to make smarter, more informed decisions in real-world settings.

Dependencies

Module Recommendations

Incompatible Modules

No incompatible modules listed

Co-requisite Modules

No Co-requisite modules listed

Requirements

No requirements listed

Indicative Content

Introduction

Investigate the data science and analytics landscape; its historical development, terminology and technologies. Identify how businesses can gather and store data and then harness the knowledge within the data using analytical techniques. Appraise the key software products within the market. Understand the key skills, roles and resources that make up a modern data analytics team.

Data analytics project life cycle

Examine how to manage a data analytical project using a framework, e.g. CRISP-DM. Investigate case studies in the field, looking at a variety of approaches, technologies with successes, failures, new developments and unusual applications of analytics. Determine how to present data analytics to different audiences within a business environment, namely a business, non-technical and a technical person.

Data types, quality and EDA

Analyse how to use exploratory data analysis (EDA) using summary statistics, plots and visualisations to discover patterns and trends within data to solve a business problem. Assess data quality, examine different data types, cleaning/scrubbing data techniques and how to evaluate which type of mathematical model is most appropriate for a given project. Use appropriate software, e.g. R, Python, PowerBI, to perform these tasks.

Data science and statistical learning theory

Investigate sampling theory (e.g. cross fold validation, bootstrapping), discuss testing and evaluation of models for both classification and regression model. Examine the concept of tuning a model's parameters and the overfitting challenge. Perform experiments using appropriate software, e.g. R, Python, PowerBI to investigate these concepts.

Model and algorithm design

Investigate the different types of statistical learning models, e.g. decision trees, ensemble techniques (random forests), support vector machines, neural networks; their characteristics, strengths and weaknesses. Examine how these models are optimised subject to constraints using such techniques such as MLE (Maximum Likelihood Estimation), Stochastic Gradient Descent (SGD) and multivariate/matrix calculus.

Ethics, data privacy and security

Investigate the ethical consequences of data analytical techniques. Evaluate data privacy, security, data protection legislation, including GDPR, the EU AI Act and related topics in data governance.

Module Content & Assessment

Assessment Breakdown	%
Coursework	40.00%
End of Module Formal Examination	60.00%

Assessments

Coursework			
Assessment Type	Multiple Choice Questions	% of Total Mark	20
Timing	Week 6	Learning Outcomes	1,2,3,4,6
Assessment Description Multiple Choice Questions on such topics as Exploratory Data Analysis (EDA), fundamental statistical learning techniques and concepts, performance evaluation methods and approaches to managing a data analytics project.			
Assessment Type	Short Answer Questions	% of Total Mark	20
Timing	Week 10	Learning Outcomes	2,3,4,5,6
Assessment Description Short answer questions requiring the students to calculate a variety of metrics, to analyse scenarios and evaluate results using statistical and data analytical, modelling techniques using			

software, e.g. R, Python, Excel, where appropriate.			
End of Module Formal Examination			
Assessment Type	Formal Exam	% of Total Mark	60
Timing	End-of-Semester	Learning Outcomes	1,2,4,5,6
Assessment Description End-of-Semester Final Examination			
Reassessment Requirement			
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.			

The University reserves the right to alter the nature and timings of assessment

Module Workload					
Workload: Full Time					
Workload Type	Contact Type	Workload Description	Frequency	Average Weekly Learner Workload	Hours
Lecture	Contact	Theory and Case Studies	Every Week	2.00	2
Lab	Contact	Computer-based lab	Every Week	2.00	2
Independent & Directed Learning (Non-contact)	Non Contact	Independent Study	Every Week	3.00	3
Total Hours					7.00
Total Weekly Learner Workload					7.00
Total Weekly Contact Hours					4.00
Workload: Part Time					
Workload Type	Contact Type	Workload Description	Frequency	Average Weekly Learner Workload	Hours
Lecture	Contact	Theory and case Studies	Every Week	2.00	2
Lab	Contact	Computer-based lab	Every Second Week	1.00	2
Independent Learning	Non Contact	Independent Study	Every Week	4.00	4
Total Hours					8.00
Total Weekly Learner Workload					7.00
Total Weekly Contact Hours					3.00

Module Resources	
Recommended Book Resources	
Peter Bruce, Andrew Bruce. (2017), Practical Statistics for Data Scientists, 1st. O'Reilly Media, California, USA, [ISBN: 9781491952962]. Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. (2021), An Introduction to Statistical Learning, 2nd. Springer, p.603, [ISBN: 978-1-0716-1417-4].	
Supplementary Book Resources	
Bill Schmarzo. (2023), AI & Data Literacy, Packt Publishing, p.0, [ISBN: 9781835083505]. Matthew North. Data Mining for the Masses, [ISBN: 0615684378]. Thomas H. Davenport. (2018), The AI Advantage, MIT Press, p.243, [ISBN: 9780262039178]. Kabacoff, Robert. (2015), R in Action, 2nd. Manning, New York, [ISBN: 1617291382]. Norman Matloff. (2011), The Art of R Programming, 1st. No Starch Press, San Francisco, [ISBN: 9781593273842]. Andy Field, Jeremy Miles. (2012), Discovering Statistics Using SAS, 1st. [ISBN: 1849200920].	
Supplementary Article/Paper Resources	
Watson, Hugh. (2011), Business Analytics Insight: Hype or Here to Stay?, Business Intelligence Journal, vol. 16, No. 1, p.1-8. Vijay Khatri, Carol V. Brown. (2010), Designing data governance, Communications of the ACM, Volume 53 Issue 1, https://dl.acm.org/citation.cfm?id=16292_10	
Other Resources	
Website, RStudio IDE, https://www.rstudio.com/ Website, Data Protection Commissioner Ireland, https://www.dataprotection.ie/docs/Home/4.htm Website, GDPR - EU, https://www.eugdpr.org/ Website, EU. EU AI Act, https://artificialintelligenceact.eu/the-act/	