

Module Title Data Science Development Keywords Data science, data preparation, data exploration, data visualisation, data science evaluation, machine learning	Reference CMM535 SCQF Level SCQF 11 SCQF Points 15 ECTS Points 7.5 Created October 2014 Approved April 2015 Amended November 2016 Version No. 3
--	--

This Version is No Longer Current

The latest version of this module is available [here](#)

Prerequisites for Module

None.

Corequisite Modules

None.

Precluded Modules

None.

Aims of Module

To provide students with necessary skills for developing complete data science products using a state-of-the-art high level programming language.

Learning Outcomes for Module

On completion of this module, students are expected to be able to:

1. Discuss the main concepts and tools for a data science project.
2. Load, explore, model and visualise data using off-the-shelf tools and packages.
3. Report data science results to a wider audience by tailoring them at different levels of detail.

Mode of Delivery

This is a lecture based module, supplemented with practical sessions, where a data science programming language will be used to teach students how to develop a complete data science project from data preparation to advanced analytics. Online materials and online sessions will be used to support DL students.

Assessment Plan

	Learning Outcomes Assessed
Component 1	1,2,3,4

Component 1 - Data science project, 15% in class presentation, 85% Report including results and R scripts.

Indicative Bibliography

1. JAMES, G., WITTEN, D., HASTIE, T., & TIBSHIRANI, R. (2021) An introduction to statistical learning with applications in R. New York: Springer.
2. An Introduction to R, Version 4.1.2 (2021), <https://cran.r-project.org/doc/manuals/R-intro.pdf>
3. ZHAO, Y. (2012-2015). R and Data Mining: Examples and Case Studies, Elsevier. <http://www2.rdatamining.com/uploads/5/7/1/3/57136767/rdatamining-book.pdf>
4. KORDON, K. (2020) Applying Data Science: How to Create Value with Artificial Intelligence. Springer.
5. MAILUND, T. (2017) Beginning Data Science in R Data Analysis, Visualization, and Modelling for the Data Scientist. APress.

4.Design, implement and evaluate a data science product that addresses a given data problem.

Indicative Module Content

1. Data science programming concepts
2. Data preparation methods
3. Data exploration, summarisation, transformation and visualisation techniques
4. Descriptive analytics (Cluster and link analysis)
5. Predictive analytics (Classification and regression analysis)
6. Advanced analytics (Text mining and social network analysis)

Indicative Student Workload

<i>Contact Hours</i>	Full Time	Part Time
Laboratories	24	24
Lectures	12	12

Directed Study

Coursework Preparation	25	25
Directed Reading	47	47

Private Study

Private Study	42	42
---------------	----	----