2024 / 25

School of Science and Computing
+353 (0)51 302037
Eleanor.Reade@setu.ie
www.wit.ie/schools/science_computing

Ollscoil Teicneolaíochta an Oirdheiscirt

South East Technological University

Module Descriptor

Data Analysis 2
(Computing and Mathematics)

Data Analysis 2 (A35483)

Short Title: Data Analysis 2
Department: Computing and Mathematics
Credits: 5 **Level:** Advanced

Description of Module / Aims

This module will build on statistical modelling techniques introduced in Data Analysis 1 and introduce the student to some machine learning techniques for data analysis. Statistical software such as Python or R will be used in the application of the techniques studied.

Programmes

	stage/semester/status
HDip in Science in Data Analytics (WD_KDAAN_G)	2 / 3 / M

Indicative Content

- Time Series Analysis: Stationarity, Autoregressive Integrated Moving Average (ARIMA) models, Parameter estimation and forecasting.
- Dimensionality reduction: Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA).
- Introduction to supervised Machine Learning techniques such as regression and classification.

Learning Outcomes

On successful completion of this module, a student will be able to:

1. Evaluate time series models and utilise statistical software to fit and apply models to data.
2. Explore dimensionality reduction techniques such as PCA and LDA and implement them with real world datasets.
3. Utilise machine learning methodologies to facilitate pre-processing, dimensionality reduction and model selection.
4. Analyse, interpret and report findings using appropriate presentation skills.

Learning and Teaching Methods

- This module is focused on using appropriate techniques for time series analysis and dimensionality reduction and subsequent visualisation and reporting. As a result the emphasis is on practice with real data, with its inherent issues such as missing values, outliers, etc. Building on the Data Analysis 1 module, students will learn how to model and analyse real data.

Learning Modes

Learning Type	F/T Hours	P/T Hours
Lecture	24	24
Practical	24	24
Independent Learning	87	87

Assessment Methods

	Weighting	Outcomes Assessed
Continuous Assessment	100%	
Practical	50%	1,2,4
Practical	50%	1,3,4

Assessment Criteria

<40%: The learner demonstrates little or no understanding of the fundamentals of the learning outcomes. Displays a lack of competence with regards to associated data analysis skills and statistics.

40%–59%: The learner demonstrates limited understanding of the fundamentals of the learning outcomes. Displays some ability with regards to associated data analysis skills and statistics.

60%–69%: As well as the above, the learner demonstrates reasonable understanding of the fundamentals of the learning outcomes. Displays reasonable competence with regards to associated data analysis skills and statistics.

70%–100%: All of the above to excellent levels. In addition, the learner demonstrates comprehensive knowledge and understanding of all learning outcomes. Displays advanced competence and independence with regards to associated data analysis skills and statistics.

Essential Material(s)

- Wei, W.W.S. *Multivariate Time Series Analysis and Applications*. 1st ed.. New York: Wiley, 2019.

Supplementary Material(s)

- "RStudio: Integrated Development fo R. RStudio." http://www.rstudio.com
- Grolemund, G. and H. Wickham. *R for Data Science*. 1st ed.. New York: O'Reilly Media, 2017.
- Hastie, T. and R. Tibshirani. *The Elements of Statistical Learning*. 2nd ed.. New York: Springer , 2009.

Requested Resources

- Room Type: Computer Lab