MODULE **DESCRIPTOR**

MODULE TITLE	EXPLORATORY DATA ANALYSIS		
MODULE CODE	CO4760 (L7)	CREDIT VALUE	20 UK CREDITS / 10 ECTS
SCHOOL	SCHOOL OF SCIENCES		

MODULE **AIMS**

The aim of this module is to provide all the necessary tools for analysing datasets and visualising their properties.

The main objectives of the module are to:
- Provide essential exploratory techniques to describe data
- Introduce computational methods for solving statistical problems
- Introduce the R programming language, its packages, statistical functions, plotting systems
- Demonstrate the principles for constructing visual representations of the data
- Evaluate the information discovered from data analysis

MODULE **CONTENT**

Indicative syllabus content:

Data objects and Attribute Types: Nominal attributes, Binary Attributes, Ordinal Attributes, Numeric Attributes, Discrete and Continuous Attributes

Data Structures: Arrays, Vectors, Lists.

Programming: Variables, Conditional Statements, Loops, Functions.

Basic Statistical Descriptors: Mean, Median Mode, Midrange, Range, Quartiles, and Interquartile Range

Mathematical Calculations in R: Numbers, Vectors, Matrices, Random Numbers.

Data Visualization: Geometric Projection Visualisation Techniques, High-level plots, Low-level plots and the layout command par, Complex Data

Statistical Inference: Descriptive Statistics, Statistical Inference for one and two samples, test of goodness of fit, Contingency Tables.

Regression: Linear Regression, Logistic Regression: Logistic Model, Probit Model, Non-Parametric Regression: Local Polynomial Regression, Smoothing Splines, Additive Nonparametric Regression.

Analysis of Variance: One-Way ANOVA, Multiple-Factor ANOVA.

Timeseries Analysis

INTENDED **LEARNING OUTCOMES**

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

1.	**Summarize** techniques for descriptive, exploratory and confirmatory data analysis
2.	**Evaluate** modelling concepts based on the characteristics of the data
3.	**Develop** programs in R for performing statistical computing and visualising data
4.	**Make judgements** about case studies through the assessment of discovered criteria from data analysis

TEACHING METHODS

Lectures deliver factual material, introduce key concepts, direct reading and relate academic aspects to practical considerations.

Tutorial sessions allow students to apply the techniques and reinforce the material delivered in the lecture.

Practical sessions enable students to discuss material and complete online or paper-based exercises.

The module will be assessed by one written coursework. The coursework requires the student to analyse datasets using the statistical methods and tools studied in class to reveal important findings and summarise, organise and communicate the generated knowledge from this data analysis through a report. The module will have an additional assessment in the format of an examination.

Distance learning

The module tutor will deliver live online lectures through MS Teams. During the live lectures the participating students will have the opportunity to engage in discussions, present their views and ask questions. The lecture sessions will be recorded and made available to the students through Blackboard. Students who cannot participate in a live lecture will have the opportunity to answer and reflect on guided questions in the subsequent live lectures or participate asynchronously on discussion boards. The module tutor will provide appropriate feedback to students' comments, as a result of the discussions. Tutor feedback will primarily be provided in an asynchronous manner through Blackboard and emails, but when the need arises, the module tutor will schedule live sessions to provide further feedback. Where appropriate, students will be also provided with relevant further reading, web links and resources for independent study. Speakers from leading organizations will be invited, where possible, to deliver invited talks and enhance the students' experience.

Students will also be provided with bi-weekly self-assessment quizzes, so that they can reflect on their progress.

Students will be provided with access to specialised software/datasets/scripts/programs, through which they will be able to complete the practical components of the module. The students will obtain the practical sheets from Blackboard and they are expected to follow the instructions included in the practical sheets to complete the lab work. If students have difficulties with a particular exercise, they are expected to contact the module tutor or post a question on the discussion forum, where the module tutor and/or their peers can provide feedback. Different means of communication will be utilized by the tutor to offer support to the students based on the reported issue, i.e. email, Skype, MS Teams, etc.

ASSESSMENT METHODS

This module is assessed through one Portfolio of coursework and one Exam.