Lancaster University Management School

timberlake
statistics | econometrics | forecasting

Econ619 : Foundations of Data Science

COURSE AIMS & OBJECTIVES, KEY SKILLS AND LEARNING OUTCOMES

Course Aims & Objectives: This course offers a first overview of data science techniques at the graduate level. It takes participants through basic data types and properties, supervised learning techniques, and unsupervised learning methods. The course also addresses prediction and classification techniques using parametric, non-parametric and ensemble methods. Participants understand how to approach data analysis problems, which tools are available to them, and how to address common problems faced in data science applications.

Key Skills: By the end of this course, students should have some knowledge and understanding of:

- Data Fundamentals: Handling diverse data types and understanding their properties.
- Machine Learning: Applying supervised and unsupervised learning algorithms.
- Prediction and Classification: Using various methods for accurate predictions and classifications.
- Analytical Approach: Systematic problem-solving for data analysis.
- Tool Proficiency: Competence in common data science tools for analysis and visualisation.
- Problem-solving: Addressing challenges in real-world data science applications.
- Ensemble Methods: Application of ensemble methods for improved predictions.
- Communication Skills: Effectively conveying insights from data analysis.

Desired Outcomes: By the end of this course, students should be able to:

- Demonstrate a sound knowledge of applied econometric principles and basic quantitative techniques;
- Demonstrate a sound knowledge of supervised and unsupervised learning techniques;
- present, interpret and analyse information in numerical form and use econometric and other packages effectively;
- Understand the relevance of different econometric approaches to specific applications in economics;
- Select relevant information from large amounts of data;

COURSE STRUCTURE

Econ 619 is a 20 credits course and therefore students are expected to input approximately 200 hours of study into the course. The total number of contact hours on Econ 619 is 25 hours. This leaves 175 hours for private study. Course Delivery comes in the form of Lectures with 25 hours delivered over the first 5 weeks of the term (20 hours of lectures and 5 hours of tutorials). There will be optional clinics on the last day of the course.

During your private study you should strike a balance between reading the course material (which is the primary source of information) and the recommended textbooks, thinking critically about how these fit into the body of knowledge on the subject and about how our level of knowledge can be improved, performing exercises, completing coursework and revising for examinations. You can expect to perform well on this course only if you work consistently through the year.

Prior to enrolment on the module, the student must have successfully completed ECON601

COURSE CONVENOR

Dr Arthur Charpentier

LECTURERS CONTACT INFORMATION (Including Office Hours)

Dr Arthur Charpentier

email: arthur.charpentier@gmail.com

Prof. Russel Davidson

email: russell.davidson@mcgill.ca

COURSEWORK ASSESSMENT

The CWA mark will be calculated as 100% coursework. The coursework will be assigned at the end of the course

The coursework will be delivered to students at the end of week 6 of each term and is due for submission at the end of week 10 of the term, allowing students 4 weeks for completion.

Coursework must be submitted electronically through the Moodle site for this course:

FEEDBACK ON COURSEWORK:

The coursework will be marked and returned to students within 4 weeks of the submission deadline. Feedback will consist of marker's notes appended to the pdf of your coursework.

MARKING CRITERIA AND PENALTIES

Marking criteria can be found in the Economics Undergraduate Handbook and the general course information paper. An electronic copy of this can be found via the Current Student page of the

university website then follow the Academic Regulations link
https://gap.lancs.ac.uk/ASQ/QAE/MARP/Documents/UG-Assess-Regs.pdf

FINAL MARK INFORMATION

This course is assessed 100% by means of coursework. The final mark is the average of the marks obtained in the two pieces of coursework.

COURSE TEXT AND RECOMMENDED READING

Lecture notes and Lecture slides.

Statistical Foundations, Reasoning and Inference: For Science and Data Science, Göran Kauermann , Helmut Küchenhoff, Christian Heumann, Springer Nature

Mathematical Foundations for Data Analysis. Jeff M. Phillips, Springer Nature

Python Data Science Handbook: Essential Tools for Working with Data. Oreilly.

MARKING CRITERIA AND PENALTIES

Marking criteria can be found in the Economics Undergraduate Handbook and the general course information paper. An electronic copy of this can be found via the Current Student page of the university website then follow the Academic Regulations link
https://gap.lancs.ac.uk/ASQ/QAE/MARP/Documents/UG-Assess-Regs.pdf

COURSE OUTLINE/LECTURE SCHEDULE

Introduction to Data Science

- What is data science and why is it important?

- The role of data in decision-making

- The principles of statistical thinking

Data types

- Importing and cleaning data using popular tools and techniques

- Basic data visualisation using graphs and plots

- Introduction to the Pandas library for data manipulation

Statistics and Machine Learning Basics

- Supervised vs. unsupervised learning

- Important data science methods for prediction and classification

- Basic machine learning algorithms and their applications

Model Evaluation and Selection

- Measuring the performance of machine learning models

- Choosing the appropriate model for a given problem

- Avoiding common pitfalls in model evaluation and selection