Lancaster University Management School

timberlake
statistics | econometrics | forecasting

Econ626 : Python Programming: Best Practices for Data Science

COURSE AIMS & OBJECTIVES, KEY SKILLS AND LEARNING OUTCOMES

Course Aims & Objectives: This course covers best practices in Python programming. It focuses on structuring code, commenting code, using versioning tools like Github, and setting up data science projects in Python which allow for effective collaboration and maintenance of code while working with other coders in large projects.

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

- Cloud Computing Proficiency: Ability to leverage cloud computing services for data science projects.
- Collaborative Coding with GitHub: Skills in using collaborative coding environments, particularly GitHub.
- Machine Learning Deployment: Proficiency in deploying, testing, and maintaining machine learning models.
- Tools Utilisation: Effective use of tools such as GitHub for collaborative coding.
- Practical Application: Application of cloud computing and collaborative coding skills to real-world projects.
- Testing and Maintenance of Models: Rigorous testing and ongoing maintenance of machine learning models.
- Cloud Service Integration: Integrating cloud computing services into data science workflows.
- Data Science Project Implementation: Applying skills to deploy and maintain machine learning models in data science projects.

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

- Write high quality and easy-to-maintain Python scripts;
- Adhere to best coding practices;
- Be confident in sharing code with others;
- Be familiar with using versioning systems such as Github.
- Be confident in using Python in data science projects;
- Understand how to best structure code for clarity, readability, and easy maintenance;
- Know how to use Github for collaborative coding projects;

- Know how to make use of the Python programming language in the best possible way for data science projects.

COURSE STRUCTURE

Econ 626 is a 10 credits course and therefore students are expected to input approximately 100 hours of study into the course. The total number of contact hours on Econ 626 is 15 hours. This leaves 85 hours for private study. Course Delivery comes in the form of Lectures with 15 hours delivered over the first 3 weeks of the term (10 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 in to 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
Econ620
Econ621

COURSE CONVENOR

Régis Amichia

LECTURERS CONTACT INFORMATION (Including Office Hours)

Régis Amichia

email: regis@foxintelligence.io

Thomas Pical

email: tpical@equancy.com

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.

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

Deep Learning with Python. Francois Chollet. Simon and Schuster.

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

- The basics of writing clean code in Python.
- The Python style guide.
- Naming variables, functions, classes, methods, and sequences.
- Documenting the code.
- Block comments, in-line comments, and docstrings.

- How to write clean Python loops with enumerate, zip, break, and the else clause.
- Best practices for indentation, line breaks, blank lines, and whitespaces in Python.
- Best practices for object-oriented programming.
- What is Github?
- Tracking code changes.
- Collaborative coding.
- Managing projects with Repositories.
- Project cloning and working on local copies.
- Staging and Committing.
- Branching and Merging.
- Pull the latest version of the project to a local copy.
- Pushing local updates to the main project.
- Files that are changed, added or deleted.
- You select the modified files you want to Stage.
- Seeing the full history of every commit.
- Reverting back to a previous commit.