Course Outline 2018/19
Python for Economists — EC3310

Instructors:

Vitaliy Oryshchenko
vitaliy.oryshchenko@rhul.ac.uk

and

Evan Piermont
evan.piermont@rhul.ac.uk

AIMS

This course aims to acquaint students with the basics of the Python programing language and its application to data collection, manipulation, visualization, and analysis. During the course, the students will learn how to write code, gather data from web sources using modern scraping techniques, organize data into a usable format, and to analyse data collections to answer questions about the world.

We will begin with an introduction to general purpose programing: data types, conditionals, loops, functions, and classes. We will investigate how to write succinct and efficient code. We will then see how to gather data from the web by using the webscraping python library Beautiful Soup and represent it inside python as a DataFrame. This will require a brief foray into the structure of websites (i.e., HTML). We will investigate how programing can help us clean and present the data in a informative and elegant manner. Finally, we will focus on using data to answer real business questions, covering popular exploratory and statistical data analysis techniques.

LEARNING OUTCOMES

Upon completion the course students should be able to:

- Install python on a Windows of Mac computer, write and run a python file.
- Implement general purpose programing techniques.
- Understand data types, conditionals, loops, functions, and classes.
- Use the Pandas Python Library to create and manipulate DataFrames.
- Have a working knowledge of HTML, sufficient to use tools to extract data from websites.
- Use the Beautiful Soup Python Library to scrape data from websites.
- Use Python to organize and present data.
- Use Matplotlib and seaborn to visualise data.
- Perform statistical (regression) analysis and apply machine learning techniques to gain insights from the data and draw relevant conclusions.
COURSE DELIVERY
The course consists of a one-hour lecture and a two-hour seminar each week. Given the subject’s practical content, the only way to understand the material is to actually program things---therefore, the seminars will consist of directed problem solving, and will take place in the IT room.

Short problem sets will be assigned during lectures. These problem sets will help to familiarize students with the tools introduced in the lecture. The solutions to the problems sets will be discussed in the beginning of the seminar. Students are strongly encouraged to talk to the course leader about the course by coming to see him either during regular office hours or by appointment.

ASSESSMENT

Summative assessment:
- Final project (65% of the final mark).
- Midterm (25% of the final mark).

Formative assessment:
- Weekly problem sets (10% of the final mark). Students will be able to assess their own performance through discussion during seminar and by comparing their answers to the posted suggestions. In addition, discussions during seminars and the instructor's office hours will provide more personalized feedback.

Solutions will be posted on Moodle (collective feedback). Scripts/projects will be marked and numerical grades assigned (individual feedback).

WEEKLY TIMETABLE
Lectures 1–2: Introduction to programing with Python.
- How to install python and run python files.
- Data types, conditionals and loops.
- Functions and Classes.
- Examples of economics applications of programing.

Lectures 3–4: Data Manipulation in Python.
- The Pandas Package
- DataFrames
- Manipulating and Cleaning Data

Lectures 5–6: Getting Data from the Web.
- Application Program Interfaces (APIs)
• The basics of HTML
• Web scraping with the Beautiful Soup Library

Lectures 7-10: Data visualisation and analysis
• Visualisation: 2D and 3D plots; histograms and kernel smoothing
• Statistical models; regression
• Fundamentals of machine learning