SYLLABUS

Course: Python and Data Science for Economics

Textbooks/Resources:
A First Course in Quantitative Economics with Python (Sargent, Stachurski)
Python Data Science Handbook (Jake VanderPlas)
Data Visualisation with Python (Gilbert Tanner's blog)

Textbooks are for reference only and physical copies need not be purchased. The online versions of all of these are available for free.

Course Description:
Modern economic analysis is heavily reliant on data and computational resources. It is critical to be able to make the best use of these resources in order to be able to solve problems of practical as well as academic relevance. This course is meant to bridge the gap between theory and practice using Python. In this course, we will revisit the foundational concepts of statistics and econometrics with a focus on applications. Students will learn to utilize the vast resources of Python to real-world data analysis problems.

Modality and Logistics:
The course will be taught through lectures twice a week of 1.5 hrs each. Students may opt to attend lectures virtually or in-person. Lecture videos shall also be posted along with slides and other teaching materials each week.

There will be worksheets and assignments to be completed and submitted every week for evaluation. There will be a final project to be completed in groups of 4/5 students. The groups are to be formed autonomously but in case of mismatches, the instructor may intervene and re-arrange them.

Students may discuss anything relevant to the course outside class during either the Office Hours (Mon, Thu 10-11.30 am) or may set up a time via email.

Classes: Tue, Fri 2-3.30 pm, BRB 1.102 (map link)
Zoom link:
Course discord:

Grading:
There will be a total of 8 assignments in this course. The assignments will have both analytical and programming components. For the programming parts, the codes may be e-mailed to me or submitted to Canvas. For the analytical portion, you may submit the assignments in-class or submit screenshots on Canvas. You are encouraged to work independently but discuss with your peers and instructors if need be.

SYLLABUS

Each assignment must be submitted at the latest by the Friday 11.59 pm of the corresponding week. Assignments may be accepted up to one week after the due date but would incur 10% penalty.

The final-grade weightage for the assessments is as follows:

1. Assignments (8) - 40%
2. Project - 60%

Project Evaluation:
The project would be due the last Friday before the Finals week. A group member must submit the Github link to the project complete with the README file detailing the scope and the details of the project. There will be group presentations during the finals week.

The projects will be scored on:

1. Originality
2. Execution of coding and Analysis
3. Presentation of results
4. Clarity of concepts and code

Course Outline:

The goal of this course is to introduce Master's level students to programming and data analysis in Python, revisit some core concepts in data science and bayesian statistics and then apply them to real-world problems.

Learning outcomes:

At the end of this course, the student is expected to be proficient in:
(i) Basic programming in python
(ii) Using Python data structures effectively
(iii) Practical application of statistical inference in economic and business problems

And develop intermediate skills in:
(i) Mathematical modeling
(ii) Solving economic models
(iii) Application of linear algebra
(iv) Optimization
(v) Statistics theory

Course Timeline
Module I: Programming in Python

Week 1&2
Basics of Python: Data Structures, Control Flow, Shell scripting, Using Git and GitHub

SYLLABUS

Week 3&4
Introduction to Data Analysis with Numpy and Pandas, Exploratory Data Analysis and Visualization, Review of Statistics and Probability

Module II: Data Science concepts
Week 5&6
Linear Regression, Estimation and Optimization, Likelihood Hypothesis testing, Fisher Information, Bias-Variance tradeoff, Model selection and Shrinkage

Week 7&8
Classification: Naive Bayes, k-means clustering/regression, Decision trees, Random Forests and Gradient boosting

Week 9&10
Panel data methods and fixed effects, Simple time-series models and forecasting, Backtesting and Model Validation, Information Criteria and degrees of freedom

Week 11&12
In-class project discussion and brainstorming

Week 13&14
Final presentations