COURSE OUTLINE

Course Code / Title	**:**	HE4045 Quantitative Economics with Data Science
Pre-requisites	**:**	HE3001 Microeconomics III, HE3002 Macroeconomics III, HE3003 Econometrics III or HE2001 Microeconomics II, HE2002 Macroeconomics II with HE2005 Principles of Econometrics / HE2004 Introductory Econometrics
No. of AUs.	**:**	4
Contact Hours	**:**	52 hours (2 hours lecture + 2 hour tutorial)

Course Aims

This course is on data science, economics, programming and how they can be used to understand the world around us. It is a complement to econometrics. The focus of this course is to learn practical programming skills for the workplace and future studies in economics and finance. Unlike courses in computer science, data science, or statistics, the emphasis of this course includes both the programming and statistics necessary to analyze data and subsequently interpret results through the lens of economics. This course covers programming and basic scientific computing in Python by re-examining basic economic concepts and models. Students who complete this course will be prepared to work in a data analyst or a data science role with the ability to frame problems in a broader economic context and provide unique insights associated with that context. No prior programming experience is required. Students who complete this module will gain valuable insights into how programming can be utilized in an economic context, providing them with a foundational understanding. It provides a solid starting point for students who want to specialize in data analysis with the ability to frame problems in a broader economic context and provide unique insights associated with that context. No prior programming experience is required.

Intended Learning Outcomes (ILO)

- Recognize and understand the connections between economic theory and the practice of data science.

- Develop ability to frame problems in a broader economic context and provide unique insights.
- Master the fundamentals of programming in Python for data science and economic analysis.
- Communicate the findings of analyses both verbally and as part of a presentation.

Course Content

1. **Python Fundamentals**: Collections, Control Flow, Functions
2. **Scientific Computing**: Introduction to Numpy, Plotting, Applied Linear Algebra, Randomness, Optimization
3. **Pandas**: Introduction to Pandas, Basic Functionality, Storage Formats, Cleaning Data, Reshape, Merge, GroupBy, Time Series, Intermediate Plotting
4. **Applications**: Data Visualization, Regression, Mapping in Python, Classification, Simulate Linear/Nonlinear Difference Equations, Introduction to Business Cycle Data, Working with Text in Economics and Finance, Machine Learning in Economics, Heterogeneous Effects

Course Assessment (To be specific)

Class Participation	: 10%
Programming Assignments	: 20%
Quizzes	: 30%
Final Group Project and Presentation	: 40%

Total	**100%**

Reading and References

We will follow the QuantEcon DataScience (https://datascience.quantecon.org) textbook Introduction to Economic Modeling and Data Science, Chase Coleman, Spencer Lyon, and Jesse Perla (2021)

Recommended: Intro to Python for Computer Science and Data Science, Deitel & Deitel. Pearson Education (2020)

Thomas J. Sargent and John Stachurski's Python lectures (http://lectures.quantecon.org/py/index.html). In particular, the content in the section: Introduction to Python.

Main Readings:
Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. "Text as Data." Journal of Economic Literature 57, no. 3 (2019): 535-574.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis. "Measuring Economic Policy Uncertainty." The Quarterly Journal of Economics 131, no. 4 (2016): 1593-1636.

Hassan, Tarek A., Stephan Hollander, Laurence Van Lent, and Ahmed Tahoun. "Firm-level Political Risk: Measurement and Effects." The Quarterly Journal of Economics 134, no. 4 (2019): 2135-2202.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy. "Measuring Technological Innovation over the Long Run." American Economic Review: Insights 3, no. 3 (2021): 303-320.

Jegadeesh, Narasimhan, and Di Wu. "Word Power: A New Approach for Content Analysis." Journal of financial economics 110, no. 3 (2013): 712-729.

Loughran, Tim, and Bill McDonald. "When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks." The Journal of Finance 66, no. 1 (2011): 35-65.

Loughran, Tim, and Bill McDonald. "Textual Analysis in Finance." Annual Review of Financial Economics 12 (2020): 357-375.

Loughran, Tim, and Bill McDonald. "Textual Analysis in Accounting and Finance: A Survey." Journal of Accounting Research 54, no. 4 (2016): 1187-1230.

Sautner, Zacharias, Laurence Van Lent, Grigory Vilkov, and Ruishen Zhang. "Firm-level Climate Change Exposure." The Journal of Finance 78, no. 3 (2023): 1449-1498.

Course Instructors

Instructor	**Office Location**	**Phone**	**Email**
Dr Ye Guangzhi	SHHK-04-79	84323282	guangzhi.ye@ntu.edu.sg

Planned Weekly Schedule

Week	Topic	ILO	Readings/ Activities
Week 1	Lecture 1: Overview and Python Basics 1.1 Variable Assignments 1.2 Objects and Types 1.3 Modules 1.4 Numbers 1.5 Strings 1.6 Booleans	1,2,3	Lecture notes, Jupyter notebooks
Week 2	Lecture 2: Collections, Control Flow and Functions 2.1 Ordered Collections 2.2 Associative Collections 2.3 Asset Pricng and NPV 2.4 Conditionals 2.5 Iteration 2.6 Economic Production Functions	1,2,3	Lecture notes, Jupyter notebooks
Week 3	Lecture 3: Numpy and Plotting Intro 3.1 Basics about Numpy Arrays 3.2 Multi-dimensional Arrays 3.3 Operations on Arrays 3.4 Matplotlib Plots	1,2,3	Lecture notes, Jupyter notebooks
Week 4	Lecture 4: Linear algebra and Randomness 4.1 Linear Algebra Concepts 4.2 Unemployment and Pricing Portfolios 4.3 Use Numpy to do Linear Algebra Operations 4.4 Basic Probability 4.5 Draw Random Numbers 4.6 Simulate Discrete and Continuous Random Variables and Processes	1,2,3	Lecture notes, Jupyter notebooks
Week 5	Lecture 5: Optimization and Pandas 5.1 Optimization using Derivatives 5.2 Gradient Descent 5.3 Series and DataFrame 5.4 Make Basic Visualizations	1,2,3	Lecture notes, Jupyter notebooks
Week 6	Lecture 6: Storage Formats, Cleaning Data, Reshape, and Merge 6.1 Storage Formats 6.2 Cleaning Methods to Prepare and Analyze Data 6.3 Reshape 6.4 Combine Datasets	1,2,3	Lecture notes, Jupyter notebooks
Week 7	Lecture 7: Groupby, Timeseries, Matplotlib, and Visualization 7.1 Groupby 7.2 Handle Datatime Objects 7.3 Intermediate Plotting with Matplotlib 7.4 Steps of Creating a Visualization	1,2,3	Lecture notes, Jupyter notebooks

	7.5 Effective Visualizations Midterm Quiz		
Week 8	Lecture 8: Textual Analysis in Economics 8.1 Recent Studies in Economics 8.2 Text Processing 8.3 Natural Language Processing	1,2,3	Lecture notes, Jupyter notebooks, Required Readings
Week 9	Lecture 9: Textual Analysis in Finance 9.1 Recent Studies in Finance 9.2 Scraping the Web	1,2,3	Lecture notes, Jupyter notebooks, Required Readings
Week 10	Lecture 10: Simulate Systems of Difference Equations and Introduction to Business Cycle Data 10.1 Simulate Linear First-order and Linear Higher-order Difference Equations 10.2 Simulate Nonlinear First-order Difference Equations and Systems of Difference Equations - Example: Solow Growth Model 10.3 Introduction to Business Cycle Data	1,2,3	Lecture notes, Jupyter notebooks
Week 11	Lecture 11: Classification, Mapping, and Machine Learning 11.1 Classification 11.2 Evaluate Classification Models 11.3 Mapping in Python 11.4 Machine Learning in Economics	1,2,3	Lecture notes, Jupyter notebooks
Week 12	Lecture 12: 12.1 Heterogeneous Effect 12.2 Basic SQL 12.3 Accessing Databases using Python	1,2,3	Lecture notes, Jupyter notebooks
Week 13	Final Project Group Presentation	1,2,3,4	Nil