



## Course Outline

### EE200 Data Science for Economic Analysis

**Semester 1/2023 (August 13 - December 2, 2024)**

**Lecture Time:** Monday 09.00-12.00 hours

**Lecture Venue:** Room 201

**Teaching Materials Platform:** [Google Classroom](#) [7op2buj](#)

#### Instructor:

**Name:** Asst.Prof. Dr.Monthien Satimanon

**Office Hours:** **By appointment**

**Email:** monthien@econ.tu.ac.th

**Phone:** 02-613-2478

**Number of Credits:** 3 Credits (3-0-6)

**Prerequisite:** 2nd-year student or above.

**Course Description:** Learning to work with data. Topics include (1) data sampling and cleaning, (2) data storage and management, (3) exploratory data analysis, (4) prediction based on statistical tools such as regression, classification, and clustering, and (5) communication of results through visualization and summary statistics. Students learn through real-world examples using programs such as advanced MS Excel, Python, or R.

**Prerequisite:** 2nd-year student or above. *(have taken at least 34 credits)*

#### Course objective:

In this course, you will not just be learning in isolation, but as part of a dynamic learning community. You will learn to conduct data analysis and projects on topics of your interest. Given the course's focus on economic issues, your topic and group project should be related to either microeconomics or macroeconomics. This course is designed to create an active learning environment, with fewer traditional lectures and more hands-on learning experiences, including group discussions and collaborative projects.

My role in this class is not just that of a lecturer, but also a coach and conductor. This course is divided into two significant parts. In the first part, I aim to provide you with a solid understanding of the fundamental concepts of data and data science.

To do this, we will discuss how to create meaningful linkages between business, economics, and data. We will then discuss different components of standard data (your team paper). This part includes an introduction, literature review, theoretical framework, data acquisition, data management, prediction, and visualization of findings. You will be assigned to read and practice data science projects and past seminar papers. The students will work on the suggested data and coding project.

In the second part, you will gradually take more roles in class. Each of you and your time will take a turn to present your topic, related literature, theoretical framework, methodology, and findings.

Other students will pay close attention to your presentation. They will then provide constructive comments and suggestions to improve your data science project. Occasionally, we will turn the class into individual meetings and team meetings. The course is where you discuss specific questions regarding your data science project and data skills with me.

## Expected Learning Outcomes

### 1. Morality and Ethics **EE200**

Applicability	Expected Learning Outcomes	Evaluation Method
●	1. Be honest	
○	2. Be fair	
○	3. Be responsible for oneself and society	
●	4. Be disciplined	
○	5. Value morality and a professional's ethical conduct	
○	6. Be service-minded	

### 2. Knowledge

Applicability	Expected Learning Outcomes	Evaluation Method
●	1. Understand principles and theories in the field of studies	
●	2. Be able to analyze the case systematically	
●	3. Be able to apply the learned knowledge appropriately	
○	4. Be able to integrate the learned knowledge appropriately	

### 3. Intellectual Development

Applicability	Expected Learning Outcomes	Evaluation Method
●	1. Be able to search for data/information systematically	
●	2. Be able to analyze the problem, assess the choices, and give an advice and results appropriately	
○	3. Be creative and have positive Thinking	
●	4. Be committed to self-development in the actual situation	

### 4. Interpersonal Skills and Responsibilities

Applicability	Expected Learning Outcomes	Evaluation Method
●	1. Be able to listen to other's opinions and be able to accept the differences	
●	2. Possess good leadership and have the courage to do the right things	
●	3. Be responsible for the assigned works	
●	4. Possess interpersonal intelligence, have patience, and be able to adjust/control oneself and emotion	
○	5. Know one's rights and duties and possess good citizenship awareness	

### 5. Quantitative Analysis, communication, and information technology

Applicability	Expected Learning Outcomes	Evaluation Method
●	1. Be able to communicate efficiently both in Thai and foreign languages.	
●	2. Be able to use and apply information technology efficiently and properly	
●	3. Possess mathematical skills	
●	4. Possess mathematical and statistical analytical skills, including the gathering of information and presentation	

Remark: ● Primary expected outcome ○ Secondary expected

## **Textbooks:**

There is no textbook for this course. However, I will post some helpful articles on Google Classroom. There is also a list of readings in Google Classroom.

## **Suggested readings by topics will be available after knowing you guys:**

**1. Data Science: business, computer, statistics, and mathematics.** Data Literacy, Fact, and Truth.

Talk to me.

Design Thinking: OKRs: Designing Your Life: The Book of Why: 3 Dimensions of Data Science, Statistics, and Econometrics. <https://clevertap.com/blog/data-science/>  
<https://www.vox.com/the-highlight/2019/5/14/18520783/harvard-economics-chetty>  
Data Science (MIT Press Essential Knowledge series) Data for the People: How to Make Our Post-Privacy Economy Work for You by Andreas Weigend Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Are by Seth Stephens-Davidowitz

<https://datarockie.com/free-data-science-books/>

Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy

<https://www.businessinsider.com/netflix-facebook-cambridge-analytica-documentary-trailer-great-hack-2019-7>

## **2. Data Sampling and Cleaning**

<https://towardsdatascience.com/sampling-techniques-a4e34111d808>

<http://siteresources.worldbank.org/INTPOVRES/Resources/477227->

[1142020443961/2311843-1142870725726/2337154-1328041661816/8405489-1342716080957/12-](http://1142020443961/2311843-1142870725726/2337154-1328041661816/8405489-1342716080957/12-)

[StataforSampling2012\(KristenHimelein\).pdf](#)

<https://www.tableau.com/learn/whitepapers/data-prep-best-practices>

## **3. Data Storage and Management**

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.661.538&rep=rep1&type=pdf>

<https://www.bmc.com/blogs/data-lake-vs-data-warehouse-vs-database-whats-the-difference/>

## **4. Exploratory Data Analysis: Describe or Explain**

- Tabulation

- Excel

- Pivot

- Power BI

- Stata®

- R

- Python Book of Why Causal Inference in Statistics Angrist, Joshua D., and Jörn-Steffen Pischke. Mastering 'Metrics: The Path from Cause to Effect.

Princeton: Princeton University Press, 2015.

- VanderPlas, J. (2016). *Python data science handbook: Essential tools for working with data.* " O'Reilly Media, Inc."

## 5. Prediction

- Machine Learning basic form.

<https://bruegel.org/2018/11/machine-learning-and-economics/>

Criteria for prediction

- Is it about causation?

[http://www.equality-of-opportunity.org/bigdatacourse\\_stanford/](http://www.equality-of-opportunity.org/bigdatacourse_stanford/)

Mullainathan, Sendhil, and Jan Spiess. 2017. "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives* 31 (2): 87-106.

Stock, James H., and Mark W. Watson. *Introduction to Econometrics*. 4th Edition. Boston: Pearson, 2018. Note: earlier editions and all international editions printed in English are acceptable.

- Shmueli, G., Bruce, P. C., Deokar, K. R., & Patel, N. R. (2023). *Machine learning for business analytics: Concepts, techniques, and applications with analytic solver data mining*. John Wiley & Sons.

## 6. Communication and Visualization

- Storytelling with data

- Show and Tell

- Draw to Win

- Dashboard White Paper from Tableau

- Ted Talk by Han Rosling

- Ted Talk by Simon Sinek

**Other teaching materials:** If there is any PowerPoint presentation, the file(s) will be on Google Classroom.

## Evaluation:

1. Group Review of Literature, AKA book review 10%
2. Group Presentations on the state of using AI and Data in Thailand 10%
3. Group presentations of project and coding requirement 10%
4. Group work on data visualization or dashboard: 20%
5. Group work on Python 10%
6. Final data science project and Communication 40%

**Data Science Projects:** A vital element of the course will be five small data projects and a final data science project, giving students hands-on experience in data. We recommend and will support using statistical software programs like Stata® and Excel® for these projects. Still, students can use other applications (e.g., Power BI, R, Python, Tableau, RapidMiner, Alteryx), provided their code and workflow are clear. The group projects are designed to be more substantial than traditional problem sets. They will include significant coding, reading, and writing elements to give students a sense of how data scientists work.

**Collaboration Policy:** Discussion and exchanging ideas and works are essential to data science work. You and your teams are encouraged to consult and discuss with your classmates on the data projects and to share resources and codes. However, you should ensure that any work you submit for evaluation is the result of your work and that it reflects your integrity. You should also understand and practice the standard citation practices, and please cite any books, articles, websites, lectures, GPTs, etc., that have helped you with your work. If you receive any help with your work (e.g., feedback on drafts, help with code, or programming), you must also acknowledge and give credit for this assistance.

**Your final Data Science Project** The project should be between a 20-25 pages long presentation, including graphs, tables, dashboards, and figures (excluding bibliography and appendices). The paper should be composed of 5 main parts: 1) introduction, 2) literature review, 3) theoretical framework and methodology, 4) findings, and 5) conclusion. We will discuss each section in more detail in class.

<b>Class</b>	<b>Topics: Schedules are subject to change based on in-class discussion and class pace.</b>
19 Aug	Introduction - Review of Course and Data Science Resource. Identification of your project and install your tools. Conduct Group Meetings to identify your topics, relevant literature, and Data Science Tools.
26 Aug	Introduction - Review of Course and Data Science Resource. Identification of your project and install your tools. Conduct Group Meetings to identify your topics, relevant literature, and Data Science Tools.
2 Sep	<i>Book Review presentation and</i>
9 Sep	<i>Workshop on Excel, Tableau, and the Use of GPTs</i>
23 Sep	<i>Group presentation on the state and development of data and AI in the Thai economy.</i>
<b>— Midterm Exam: Submission of Data Science Project Interim Report</b>	
7 Oct	<i>Workshop on Python: Data Preparation</i>
14 Oct	<i>Holidays</i>
21 Oct	<i>Workshop on Python: Data Visualization</i>
28 Oct	<i>Workshop on Python: Data Visualization</i>
4 Nov	<i>Workshop on Python: Presentation of Dashboard and Visualization</i>
11 Nov	<i>Machine Learning on Python: Group Work Presentation on Python</i>
18 Nov	<i>Orange for Machine Learning</i>
25 Nov.	<i>Orange for Machine Learning</i>
2 Dec	<i>Class Wrap-Up and terminal project presentations</i>
<b>— Final Project Presentation: TBA: Via Zoom and by Appointment</b>	

