

JOUR 5020 Syllabus

社會科學程式設計 Programming for Social Science

Announce

本學期除了帶學生學習基礎的程式設計技能外，另規劃如何大量融入AI Copilot於程式設計學習的目標。從期初到期末，將依序了解程式語言的基本邏輯、如何應用AI助理輔助進行文本分析、如何直接用AI助理做Vibe Coding設計探索性分析介面等。

Introduction

近年資料科學、計算機方法廣泛地在各個領域被應用，社會科學也不例外。過去我們也培育相當多人文社會科學的資訊跨域人才，有部分進入一般想像的資訊領域職場，有部分進入了跨域職場(如輿情分析師、資料新聞記者)、更有不少人社科學的畢業生告訴我們他們在人資、公關、法務、投顧等工作上運用資料科學方法，寫程式、爬取資料、彙整資料、探勘資料。

本門課主要的授課對象為對應用計算機方法於政治傳播與社會議題有興趣的學生，相較於演算法與數值運算，將著重在社會科學的文字運算、網絡結構、人口變項、地理資訊等跟資

料特性相關的運算。如果你未來想要用這類議題作為碩士論文或未來想到國外修讀計算社會科學碩士學位者，將是最合適的選修人選。或者你曾經選修過經濟系、共教中心、新聞所等R語言相關課程但卻尚未學過Python程式語言的同學，更是你接續學習第二種語言的最好機會。

本門課相較於一般程式設計課程會更重視文字的處理和實際的應用，已修過其他校內Python相關課程、電機資訊學院同學請勿選修。近年台大校內有相當多程式學習資源，如經濟系程式設計、經濟系資料科學與社會研究、資管系商管程式設計都是非常精彩、行之有年且設計完善的課程。因此，本門課選課的優先順序為新聞所 > 社科院、法學院、生傳系 > 文學院 > 其他科系。

過往本門課會比較強調傳遞完整的社會科學程式設計能力給學生，但本學期在作業與專題的設計上，會希望引導同學去發掘社會科學寫程式可以做什麼事，並大量運用ChatGPT作為教學輔助工具於課程中，也希望能引發更多同學的興趣。

Basic Information

- Instructor: 謝吉隆 (新聞所副教授)
- TA: 黃俐珏 (新聞所碩一) lisa1764737958@gmail.com
- Time: MON 2, 3, 4 時間為09:20~10:30, 10:40~12:00
- Room: 新聞所R103
- Assignment submission: NTU COOL [社會科學程式設計](#)
- Q&A: Discord Inviting link: <https://discord.gg/e4rzeB9t>





Materials




- [Programming for Social Scientists \(p4css.github.io\)](https://p4css.github.io/): Online tutorial
- 參考書籍
 - Think Python 2e [Think Python 2e – Green Tea Press](#)。看英文應該會比較容易懂，
但[《Think Python 2e》最新版中文翻译 – 《Think Python 2e》中译本1.0 文档](#)
 - [Python Data Science Handbook | Python Data Science Handbook \(jakevdp.github.io\)](#)
 - [Python for Data Analysis, 3E \(wesmckinney.com\)](https://wesmckinney.com/)

Grading

- **Absence** - 3%: Absences should be reported through the myNTU's **student leave system**. If a student is absent without leave, 3 points will be deducted from the total semester grade.
- **Assignments** 40~45%: On average, one assignment will be given per week, totaling 8 to 10 assignments throughout the semester. The duration for each assignment is one week. Late submissions within four days (until Friday, 23:59) will have the score multiplied by 0.7. All assignments must be submitted via NTU COOL [社會科學程式設計 Programming for Social Science \(ntu.edu.tw\)](https://cool.ntu.edu.tw/).
- **Quiz & Learning Sheets** 5~10%: Each quiz will consist of approximately 3 to 5 questions and will be graded on a scale of 2 points. Each learning sheet will get 1 point.
- **Mid-Exam** 15%. The exam is graded on a percentage scale (0-100).
- **Project** 35%: The project grade comprises two components:
 - Guest experts (scholars or industry professionals) will be invited to participate in the final evaluation, contributing 25% to the overall assessment.
 - Final Peer Review (10%): Assessed through peer evaluation, considering completeness, creativity, and usefulness.

Calendar

W	Date	Skills (2h)	ASs
1	09/01	Course Overview Counting, Keywords	 AS01 for Python Basic
2	09/08	Python Basic Accessing list and dictionary	Datacamp: Introduction to Python (3) Datacamp: Intermediate Python (3)
3	09/15	Flow-control	 AS02 List-dict-for-if
4	09/22	Pandas: filter, select, and timeline	 AS03 Summarize data with Pandas
5	09/29		
6	10/06		
7	10/13	Proposal (3 mins) Using Github Copilot	 AS04 Pandas and Chart
8	10/20	Scrapers	

9	10/27	Text Mining (1) Sentiment analysis Using LLM, OpenAI	 AS05 Features of text: Word frequency, collocation, POS, NER, LangExtract
10	11/03	Text Mining (2) Feature selection and embedding	 AS06 Sentiment Analysis
11	11/10	Mid-Exam UI: Streamlit & Flask	 AS07 Exploratory Data Analysis UI
12	11/17	Using LLM: Completion & RAGs	
13	11/24	Text Mining (3) Topic Modeling (LDA, BERTopics)	Hand-writing take-home review sheet for text mining
14	12/01	Early Report	Early Report
15	12/08	Working on Project	
16	12/15	Final Report [Self-efficacy survey]	Final Report