SOSC 2400: Quantitative Data Analysis for Social Research II
Fall 2023

Wednesdays and Fridays, 3:00-4:20pm
Room 2404, Lift 17-18

Instructor:	Dr. ZHOU, Titi
Contact Information:	Room 2356; Email: ttz@ust.hk
Office Hours:	Friday, 1:30-2:30pm (or by appointment)
Course Website:	Canvas
TA:	LIU, Ying
Contact Information:	Room 2359; Email: liuyingly@ust.hk
Office Hours:	By appointment
TA:	HO, Pat Shu Roy
Contact Information:	Room 2359; Email: psrho@ust.hk
Office Hours:	By appointment

Course Description

Data analysis has significantly transformed the way we understand social issues and decision-making processes. This course aims to equip students with the essential knowledge and skills needed to perform data analysis in their own research and critically evaluate statistical claims made in news media, policy reports, and academic research.

Students will be introduced to the core principles of statistical inference and acquire the necessary programming skills to address various social questions through data analysis. We will demonstrate how to carry out the analyses, interpret the results, and identify the strengths and potential limitations of each analysis.

Upon completing this course, students will be able to recognize fundamental statistical concepts and methods, use R to summarize data numerically and visually, employ linear regression to analyze data and interpret the results, and understand uncertainty in data analysis and how to quantify it.

Prerequisite

SOSC 1100.

Canvas

Canvas is the primary learning platform by which the course is delivered. If you encounter any problems with Canvas, please visit https://cei.hkust.edu.hk/canvas/faqs-students first for help.

You are responsible for all the information posted on Canvas for this course. Please check Canvas frequently since this is where we will post announcements, class assignments, and any schedule changes.

Textbook

- Llaudet, Elena and Kosuke Imai. 2022. *Data Analysis for Social Science: A Friendly and Practical Introduction*. Princeton: Princeton University Press.

- Agresti, Alan, Christine Franklin and Bernhard Klingenberg. 2023. *Statistics : the Art and Science of Learning from Data*. Harlow ; Hoboken, NJ : Pearson.

Software

We will be doing data analysis with R, an open-source statistical software. Given its power and flexibility, R has been widely used by data analysts in both corporations and academia. You can download it and find useful documentation at http://www.r-project.org/.

To help make using R easier, you are strongly encouraged to use RStudio (https://posit.co/downloads/) – a freely available user interface that simplifies many common operations.

Grading Policy

Your final grade will be based on the following:

- Attendance (10%): You are required to attend each class meeting. Attendance check may be randomly conducted.

- Weekly question/comment submission (10%):

  - To promote discussion in class meeting, you need to submit a question or comment related to the weekly topic and lecture.
  - The questions/comments need to be posted on the *Discussion section* on Canvas by **10:00am, Tuesdays**. The first will be due at **10:00am, Sep 12**.
  - Late submission is not accepted, but you are allowed to miss one submission.

- Problem Sets (20%):

  - There will be **three take-home** problem sets. The purpose is to practice data analysis skills using R and learn key statistical concepts.
  - Problem set will be posted on Canvas by **6:00pm, Wednesday (see "Schedule" at the end)**.
  - Submission deadline is **10:00am, Tuesday (see "Schedule" at the end)**.
  - Late submission will cause a grade deduction by half. Six hours late submission will not be accepted.

- In-class Quizzes (30%):

  - There will be **three closed-book** quizzes in our class meetings. The purpose is to assess how well you are understanding the key concepts covered in the lecture videos.
  - Only the **two highest** grades will count towards your final grade.

- Final Exam (30%): The closed-book final exam will cover the content learned throughout the semester. Specific arrangements will be announced later.

Honesty in Academic Work

Honesty and integrity are integral components of the academic process. Disciplinary actions for dishonesty cases have been strengthened. Students are advised to read the guideline at: https://registry.hkust.edu.hk/resource-library/regulations-student-conduct-and-academic-integrity

Course Outline and Tentative Schedule

Friday	Wednesday	#	Topic	Problem Set Posted*	Problem Set Submitted**	Quiz
Sep-01		0	Introduction			
	Sep-06	1	Data and descriptive analysis			
Sep-08	Sep-13	2	Relationship between two variables			
Sep-15	Sep-20	3	Predicting outcomes using linear regression			
Sep-22	Sep-27	4	Multiple linear regression	Sep-27	Oct-03	
Sep-29	Oct-04	5	Probability distributions			Oct-04
Oct-06	Oct-11	6	Statistical inference: confidence intervals			
Oct-13	Oct-18	7	Statistical inference: significance tests			
Oct-20	Oct-25	8	Statistical inference: comparing two groups	Oct-25	Oct-31	
Oct-27	Nov-01	9	Linear regression: inferences			Nov-01
Nov-03	Nov-08	10	Linear regression: practical issues			
Nov-10	Nov-15	11	Linear regression: with continuous and categorical predictors	Nov-15	Nov-21	
Nov-17	Nov-22	12	Logistic regression			Nov-22
Nov-24		13	Review			
	Nov-29		Final exam			

* Problem set will be posted by 6:00pm on that day. ** The submission deadline is 10:00am on that day.