SOSC 1100: Quantitative Data Analysis for Social Research I

Fall 2023

Tuesday, 4:30-5:50pm
Room 5619, Lift 31-32

Instructor: Dr. ZHOU, Titi
Contact Information: Room 2356; Email: ttz@ust.hk
Office Hours: Wednesday, 1:30-2:30pm (or by appointment)
Course Website: Canvas

TA: LAM, Wing Yin Lillian
Contact Information: Room 2359; Email: lillianta@ust.hk
Office Hours: By appointment

Course Description

This course provides an introduction to data analysis for the social sciences, assuming no prior knowledge of the subject matter. Using plain language, the course introduces basic statistical concepts and shows how to analyze real-world data using the statistical program R. Specifically, the course (1) covers basic practices of analyzing data for social science research, such as data management, descriptive analysis, and correlation; (2) introduces basic idea and intuition on causality; and (3) demonstrates how the aforementioned practices are actually executed with real-world data using R.

Upon completion of this course, students will be able to understand some basic concepts of quantitative reasoning, use basic function of R to summarize data numerically and visually, and make preliminary data-backed arguments.

Course Format

This course will be taught in *blended mode*. You are required to watch lecture videos outside class *and* participate in learning activities inside class. Specifically, every week:

- **On Tuesday (in the classroom)**: You will attend the face-to-face class meeting and are expected to participate in discussions and other activities. The meeting will involve practices using **R**. Please bring your laptop to class.
- **On Thursday (online learning)**: Lecture notes and videos will be released through Canvas by 12:00pm. You will watch the videos at your paces and according to your preferred time (*before* our class meeting next Tuesday). A few questions for discussion in class will be listed at the end of each lecture.

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.

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 about the lecture notes/videos for **each week**.
  - The questions/comments need to be posted on the *Discussion section* on Canvas by **10:00am, Mondays**.
  - 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, Tuesday (see "Schedule" at the end)**.
  - Submission deadline is **10:00am, Monday (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

Video (Thursday)	Meeting (Tuesday)	#	Topic	Problem Set		Quiz
				Posted*	Submitted**	
	Sep-05	0	Introduction			
			Data and Data in R			
Sep-07	Sep-12	1	R and Rstudio			
Sep-14	Sep-19	2	Data, observations, and variables			
Sep-21	Sep-26	3	Computing and interpreting means	Sep-26		
			Inferring population characteristics via survey reserach			
Sep-28	Oct-03	4	Sampling		Oct-02	Oct-03
Oct-05	Oct-10	5	Descriptive statistics			
Oct-12	Oct-17	6	Tables and graphs	Oct-17		
Oct-19	Oct-24	7	Scatter plot		Oct-23	Oct-24
Oct-26	Oct-31	8	Correlation			
			Estimating causal effects with randomized experiments: Preliminary			
Nov-02	Nov-07	9	Causal effects with randomized experiments	Nov-07		
Nov-09	Nov-14	10	Difference-in-means estimator		Nov-13	Nov-14
Nov-16	Nov-21	11	Examples of randomized experiments			
Nov-23	Nov-28	12	Review			
	Nov-30		Final Exam			

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