SAMPLE course outline only
Course contents will vary if taught by different teachers.
Please refer to the course teacher for the actual course content offered in each term.

As of December 27, 2020

SOCI 3229: Quantitative Data Analysis

Department of Sociology
The Chinese University of Hong Kong

Spring, 2021
Thursdays 09:30-11:15 (Lectures fully online)

Contact Information

Professor: LEE Jaemin, Ph.D.
E-mail: jaeminlee@cuhk.edu.hk
Office hours: TBA
Office: Sino Building Room 418

Teaching Assistant: CAO Ji, M.A.
E-mail: 1155152057@link.cuhk.edu.hk
Office & office hours: TBA

Course Description

Data analysis is a powerful tool for answering many interesting questions about societies and human behavior. Data-driven approaches are useful for solving the multitude of the challenging problems that today's societies confront. The focus of this course is on hands-on data analysis and the practical application of basic statistical methods to real-world social problems. Topics include causality, measurement, probability and statistical theory, linear and multivariate regression models, analyses of network, textual, and spatial data.

At the end of this course, students should be able to analyze data, interpret the results, and effectively communicate their empirical findings. The best way to learn about data analysis and new statistical procedures is by doing—not by reading and paper-and-pencil statistics per se. Hence, particular attention throughout the course will be paid to learning and implementing the **R** statistical program.

This course will also expose students to recent data and computational revolutions leveraged by the emerging field of computational social science. This field collectively addresses longstanding sociological questions—many of which have been difficult to be unearthed by conventional survey methods. We will cover, for example, automated methods and visualization tools to identify patterns in network, textual, and spatial forms of data sets or "big data."

There is no prerequisite to attend this course. Having taken Social Statistics (SOCI 2004 or equivalent elsewhere) or prior computing experiences might be helpful but not required.

Assessment and Grading

The grade for the course will be calculated as a weighted average of the following components:

Participation (Lecture & Tutorial)	10%
Assignments	30%
Examination	20%
Final Paper	40%

Participation (10%)

- Attendance will be a crucial part for your success, as the material builds on itself cumulatively throughout the course. Do not fall behind. If you start falling behind, see your tutor immediately, in order to catch up.
- To attend or skip is an adult's decision, so I do not want to punish your few times of absence/lateness with a harsh grade. However, I will highly compensate diligent and active students.
- It starts promptly on time. Arrivals after 5 minutes of the class beginning will be counted as lateness. I will check entry logs in Zoom.
- Equal weights will be given for participation in tutorials.

Assignments (30%)

- Assignments consist of problem sets, quizzes, programming exercises, and short memos, on a (roughly) biweekly basis. The goal is to maximize your understanding of the content of the week.
- You are allowed to help each other to understand materials or programming techniques. But do not, under any circumstances, copy another person's code or answers.
- All submissions should be done on Blackboard.
- For all assignments, late submission is not allowed without at least 24 hours prior notice.

Examination (20%)

- There will be one open-book in-class exam held during the lecture time on March 25.
- No other date can be scheduled to accommodate individual needs.
- You can use your books, things on the Internet etc. However, no collaboration at all is allowed on—don't talk to other humans.
- Whether the exam is arranged offline or online will be announced later.

Final Paper (40%)

- Final paper is a team project of two. You will be paired with another student to work together.
- Your team is required to submit it in a form of social science research—the product of your own, original quantitative data analysis.
- Two dues: a proposal (10%, by April 8) and a final paper (30%, by May 10). Outstanding proposals will be invited for presentation at the last week.
- Specific instructions and evaluation criteria will be provided.

Grading

Grade	Descriptors
A	Excellent: Outstanding performance on all learning outcomes.
A-	Very Good: Generally outstanding performance on all (or almost all) learning outcomes.
B	Good: Substantial performance on all learning outcomes, OR high performance on some learning outcomes which compensates for less satisfactory performance on others, resulting in overall substantial performance.
C	Fair: Satisfactory performance on the majority of learning outcomes, possibly with a few weaknesses.
D	Pass: Barely satisfactory performance on a number of learning outcomes.
F	Failure: Unsatisfactory performance on a number of learning outcomes, OR failure to meet specified assessment requirements.

Tutorials

You are required to attend tutorials. Tutorials will offer solutions to problem sets given in assignments, demonstrate R programming tutorials, and catch up/further the content of the lectures. Tutorial times: TBA. Some hands-on tutorials wil be scheduled to perform offline but are subject to change to the online mode, depending on the pandemic situation.

Textbook

It is your responsibility to do the reading *before* class. Required readings will be based on the combination of the following books (in the order of importance):

Imai, Kosuke. 2018. *Quantitative Social Science: An Introduction*. Princeton University Press.

- Purchasable at Google Play (US$19.8 to rent for 180 days, US$39.1 for ebook). I have processed a request for acquisition at the CUHK Bookstore. It will be HK$385.4 for a paperback copy when it's here.

Diez, David M., Christopher D. Barr, and Mine Cetinkaya-Rundel. 2012. *OpenIntro Statistics*. 3rd edition. Online accessible: https://drive.google.com/file/d/0B-DHaDEbiOGkc1RycUtIcUtIelE/view

Treiman, Donald J. 2014. *Quantitative Data Analysis: Doing Social Research to Test Ideas*. John Wiley & Sons. Online accessible: https://ebookcentral.proquest.com/lib/cuhk-ebooks/detail.action?docID=706553&pq-origsite=primo

Academic Honesty

Please keep in mind the university's policy on academic honesty. Plagiarism will not be tolerated in the term paper and assignments. The ideas and language should be your own, and any outside sources must be clearly and properly cited. There are severe consequences if you commit any acts of academic dishonesty. In addition to the department's policy and guidelines for citations, please refer to the university-level disciplinary guidelines and procedures. The Faculty of Social Science has also compiled a handout to alert students of the importance of academic honesty and the consequences of violating the University's Rules. To this end, the final term paper should be submitted to VeriGuide.

Online Class Logistics

- You are required to do CUHK authentication to enter Zoom meetings.
- Class recordings: Every lecture will be recorded and stored. The video recordings will be available upon individual requests and for personal use with a limited timeframe. Feel free to email me if there was a technical problem related to your connection to a live Zoom meeting or if you need a deeper review on subject matters covered in classes. However, not attending—or missing a significant portion of—the class cannot be a reason for your request, unless otherwise noticed beforehand.
- Lecture slides will be uploaded to the course Blackboard.

Schedule

Lectures

Week	Date	Topic	Reading
1	Jan-14	Introduction	
2	Jan-21	Causality	QSS 1,2
3	Jan-28	Measurement	QSS 3
4	Feb-04	Prediction	QSS 4
5	Feb-11	NO CLASS: Lunar New Year	
6	Feb-18	Inference I	OpenIntro 4; QSS 7.1
7	Feb-25	Inference II	OpenIntro 5.1-3, 6.1-2; QSS 7.2
8	Mar-04	Regression I	OpenIntro 7; QSS 7.3
9	Mar-11	Regression II	OpenIntro 8
10	Mar-18	Regression III	Treiman (TBA)
11	Mar-25	Examination	
12	Apr-01	NO CLASS: Reading Week	
13	Apr-08	Network Data	QSS 5
14	Apr-15	Text and Spatial Data	QSS 5
15	Apr-22	Wrap-up/Presentation	TBA

Tutorials and Assignment Dues (TBA)

Week	Date	Tutorial	Dues
1	Jan-14		
2	Jan-21		
3	Jan-28		
4	Feb-04		
5	Feb-11		
6	Feb-18		
7	Feb-25		
8	Mar-04		
9	Mar-11		
10	Mar-18		
11	Mar-25		
12	Apr-01		
13	Apr-08		Term Paper Proposal
14	Apr-15		
15	Apr-22		
	May-10		Final Paper