

SOSC 2400 **Quantitative Data Analysis for Social Research II**

Fall 2021

Monday & Wednesday, 10:30-11:50am

G009A, CYT Bldg

Instructor: Dr. WANG Hongbo (hbwang@ust.hk)
Office: Academic Building, Room 2372 (Ext. 7804)
Office Hours: By appointment

TA: CHEN Ziyang (zchenes@connect.ust.hk)
Office: Academic Building, Room 3001
Office Hours: By appointment

Course Description and Objectives:

This course mainly covers implementation of linear regression from a social scientific perspective. The focus is on the specification of models including choices of variables, incorporation of different types of effect, and interpretation of coefficients.

Note that the course does not treat statistical inference explicitly, leaving the latter almost entirely to a formal statistics course.

Organization:

The class meets twice a week on Monday and Wednesday, respectively, each lasting for 80 minutes. The lecture will be given on Monday each week, while Wednesday usually is reserved for an accompanying computing session (See “Schedule” below).

All course materials will be distributed through [Canvas](#). Note that they should be used *exclusively* for the purpose of this course.

Students are encouraged to form groups of 5-6 individuals to work on the group project. Although division of work is allowed among group members, each member must make adequate contribution to the project.

Computing:

We will mainly use R as computing tool.

Prerequisite:

SOSC 1100

References:

Baumer, Benjamin S., Daniel T. Kaplan, and Nicholas J. Horton. 2017. *Modern Data Science with R*. Chapman and Hall/CRC. [BKH]

Imi, Kosuke. 2018. *Quantitative Social Science: An Introduction*. Princeton University Press. [I]

Navarro, Danielle. [Learning statistics with R: A tutorial for psychology students and other beginners](#) (Version 0.6). [N]

Treiman, Donald J., 2009. *Quantitative Data Analysis: Doing Social Research to Test Ideas*. Jossey-Bass. [T]

Assessment:

Your grade will be determined as follows:

(1) Attendance: 20%

Attendance is required for all classes. We will take attendance via iPRS in class. One point will be deducted for each missed class *without legitimate justification*.

(2) Class participation: 10%

Your class participation will be evaluated using a three-level scheme, i.e. excellent (=10), satisfactory (=8), and unsatisfactory (=4).

(3) Group project: 70% (presentation, 20%; written report, 50%)

Under the supervision of the instructor, each group will choose a topic of their own, locate appropriate data sources, carry out data analysis, present the findings, and, finally, submit a written report. Detailed guidelines will be provided in a separate document.

Each group should keep a diary of your work on the project, which describes all related activities, including exploration of literature and data, downloading and processing of data, and analyzing data. This should document that you have spent reasonable efforts on the project, including on data processing and exploration that are not reflected in the final report.

Each group is required to submit a brief report on the progress of the project via Canvas **in Week 8**. A special assignments will be created for you to upload the report.

Schedule (*Subject to adjustment*)

Calendar Week	Topic	Important Dates
Week 1: Monday Wednesday	NO CLASS	9/1
Week 2: Monday Wednesday	[R] Course Introduction [L] Causality	
Week 3: Monday Wednesday	[R] Dataframe: A Review [L] Simple Linear Regression (SLR)	
Week 4: Monday Wednesday	[R] <i>Practice Session</i> [L] NO CLASS	9/22
Week 5: Monday Wednesday	[R] <i>Guidelines for Project</i> [L] Choice of X	9/27: Project Group Finalized
Week 6: Monday Wednesday	[R] Check for Correlation between Variables [L] Interpreting SLR	
Week 7: Monday Wednesday	[R] $\ln(\)$ Function [L] Log transformation	
Week 8: Monday Wednesday	[R] Log Transformation for SLR [L] Curvilinear Effect	10/20: Progress Report Due
Week 9: Monday Wednesday	[R] Incorporating Curvilinear Effect [L] Categorical X	
Week 10: Monday Wednesday	[R] Factor Variable [L] Interaction Effect	
Week 11: Monday Wednesday	[R] SLR by Group [L] Motivation for MLR	11/3: Proposal Due
Week 12: Monday Wednesday	[R] <i>Practice Session</i> [L] Visualization	
Week 13: Monday Wednesday	*Project Presentation *Project Presentation	11/22 11/24
Week 14: Monday Wednesday	[R] Q & A for Group Project	11/29
<i>TBD</i>		Final Exam

[L] Lecture

[R] R and computing