

Data Analysis for Quantitative Social Research (SOSC 1110)

Spring 2025

Room 4402 (Lift 17-18); Wed, 12:00-14:50

Instructor: Yifan Shen

Email: yifanshen@ust.hk

Office Hours: by appointment

Office Address: Room 2370

Teaching Assistant:

TA: Yueran HOU

Email: yhouau@connect.ust.hk

Office Hours: Mon 10:00–11:00 am

Office: As the SOSC RPG Common Room is currently under renovation, her office location will be determined on a case-by-case basis.

TA: ZHANG Anning

Email: annizan@ust.hk

Office Hour: Thursday 10:00–11:30 am

Office: Room 2359, lift 13-15

Course Description

This course focuses on practical aspects and implementation of social data analysis by introducing basic yet hands-on techniques for presenting, analyzing, and interpreting quantitative data, many of which are rarely taught in a regular statistics course. It is deliberately designed as complementary to a formal and theoretically-oriented introductory statistics course. This course introduces basic knowledge about quantitative data analysis from a social scientific perspective, from data generating processes to causal inference. Devoted computing sessions, a signature feature of the course, demonstrate hands-on techniques, such as data extraction, data management, variable manipulation, and descriptive analysis. Students will have the chance to implement ideas and methods from the lectures through in-class exercises and by applying what they learn to real-world data. Upon completion of the course, students should have acquired useful skills for social data analysis as well as a better understanding of quantitative social scientific research.

Intended Learning Outcomes (ILOs)

1. Develop basic quantitative skills for social data analysis.
2. Develop a better understanding of quantitative social scientific research.

Computing

Stata will be used as the major computing tool. You can access *Stata* either using our classroom computers or through [Virtual Barn](#). There are probably some other computers on campus that have *Stata* installed (I remember many computers in the library have *Stata*), but you will need to find out their location by yourself.

Textbook

Great News: No required textbook. The recommended reference books are [Statistics from OpenStax](#)

[Getting Started with Stata for Windows](#)
[Data Analysis Using Stata, Third Edition](#)

Assessment

Your grade will be determined as follows:

(1) *Attendance: 15%*

The TA/instructor will record attendance on five randomly selected sessions (so that each session record is worth 3 marks). If you have to miss a class session for medical reasons, email the TA/instructor at least one hour before the class start time (leave applications submitted after that will not be accepted) with valid medical proof on which the doctor explicitly recommends leave for the day of absence.

(2) *Quiz 1: 15%*

An in-class open-book quiz held somewhere in the middle of the semester (see schedule below).

(3) *Quiz 2: 15%*

An in-class open-book quiz held somewhere in the middle of the semester (see schedule below).

(4) *Quiz 3: 15%*

An in-class open-book quiz held somewhere in the middle of the semester (see schedule below).

(5) *Lab Assignment: 30%*

There will be one lab assignment to help you hone your Stata skills. You will start working on them during a lab session under the guidance of TA and then submit your work to Canvas (see schedule below).

Grading Rubrics

All assessments will be graded on the basis of correction instead of completion. The three quizzes will consist of multiple-choice and open-ended questions. The lab assignment will consist of multiple Stata coding tasks.

Mapping of Course ILOs to Assessment Tasks

Assessed Task	Mapped ILOs	Explanation
Class Attendance	ILO1 and ILO2	
Lab Assignment	ILO1	This assignment helps students develop Stata skills
Quiz 1, 2, and 3	ILO2	These quizzes help students gain deeper understanding of the readings

Final Grade Descriptors:

Grades	Short Description	Elaboration on subject grading description
A	Excellent Performance	Demonstrates a comprehensive grasp of subject matter, expertise in problem-solving, and significant creativity in thinking. Exhibits a high capacity for scholarship and collaboration, going beyond core requirements to achieve learning goals.
B	Good Performance	Shows good knowledge and understanding of the main subject matter, competence in problem-solving, and the ability to analyze and evaluate issues. Displays high motivation to learn and the ability to work effectively with others.

C	Satisfactory Performance	Possesses adequate knowledge of core subject matter, competence in dealing with familiar problems, and some capacity for analysis and critical thinking. Shows persistence and effort to achieve broadly defined learning goals.
D	Marginal Pass	Has threshold knowledge of core subject matter, potential to achieve key professional skills, and the ability to make basic judgments. Benefits from the course and has the potential to develop in the discipline.
F	Fail	Demonstrates insufficient understanding of the subject matter and lacks the necessary problem-solving skills. Shows limited ability to think critically or analytically and exhibits minimal effort towards achieving learning goals. Does not meet the threshold requirements for professional practice or development in the discipline.

Course AI Policy

Use of AI is allowed as long as it does not lead to plagiarism. AI cannot be used to answer the in-class quizzes.

Communication and Feedback

Assessment marks will be communicated via Canvas within two weeks of submission. Feedback on assignments will include a few suggestions for further improvement. Students who have further questions about the feedback including marks should consult the instructor or the TA within five working days after the feedback is received.

Late submission Policy

Late submissions will receive a penalty of 5 marks for each day after the deadline (until reaching zero).

Required Texts and Materials

Readings and video materials will be posted on Canvas.

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to [Academic Integrity | HKUST – Academic Registry](#) for the University's definition of plagiarism and ways to avoid cheating and plagiarism.

COURSE SCHEDULE (subject to change; updates will be posted on Canvas)

<i>Week</i>	<i>Topic</i>	<i>Date</i>	<i>Note</i>
1	Course Overview	Feb 4	
2	Case Study I	Feb 11	
q	Case Study II	Feb 18	Quiz 1
4	Data Manipulation in Stata	Feb 25	Quiz 2
5	Descriptive Statistics in Stata	Mar 4	

6	Descriptive Statistics in Stata	Mar 11	
7	Descriptive Statistics in Stata	Mar 18	
8	Measurement and Sampling	Mar 25	
9	Central Limit Theorem	Apr 1	
10	Lab Assignment (TA)	Apr 8	Lab Assignment Due (11:59pm)
11	Hypothesis Testing	Apr 15	
12	Hypothesis Testing	Apr 22	
13	Hypothesis Testing	Apr 29	
14	No class	May 6	Quiz 3