The Hong Kong University of Science and Technology

UG Course Syllabus

Data Analysis for Quantitative Social Research (Spring 2026)

SOSC 1110

3 Credits

Pre-requisites: None

Instructor Name: Jiaxin Shi
Instructor Email: jxshi@ust.hk
Office Hours: Wednesdays 16:00–17:00, Room 3350
TA Name: Anning Zhang
TA Email: annizan@ust.hk
Office Hours: By appointment, Room 2359

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 sources to probability fundamentals. Devoted R computing sessions, a signature feature of the course, demonstrate hands-on techniques, such as data extraction, data management, variable manipulation, descriptive analysis, and data visualization. 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 and a better understanding of quantitative social scientific research.

Intended Learning Outcomes (ILOs)

By the end of this course, students should be able to:

1. Obtain basic knowledge about quantitative data analysis from a social scientific perspective.
2. Conduct computer-based data extraction, data management, variable manipulation, and descriptive analysis.
3. Acquire statistical skills that can be appliable to real-world data.

Computing

Students must bring a laptop to all R computing sessions. Required software (free):
 - R (version 4.0 or higher): https://cran.r-project.org/
 - RStudio Desktop: https://posit.co/download/rstudio-desktop/

Installation instructions and support will be provided in the first computing session. If you do not have access to a personal laptop, please contact the TA or instructor by the end of Week 1 to discuss alternative arrangements.

Schedule

Sequence	Topic
1	Introduction to Data Literacy
2	Data Type and Measurement
3	Data Sources in Social Research
4	Introduction to R Programming
5	Descriptive Statistics
6	Comparing groups
7	Data Visualization
8	Probability Fundamentals
9	Measuring Relationships
10	Randomness and Patterns in Data
11	Special Topics
12	Review

Assessment and Grading

This course will be assessed using criterion-referencing and grades will not be assigned using a curve. Detailed rubrics for each assignment are provided below, outlining the criteria used for evaluation.

Assessments:

In addition to the final examination, students will do two group projects (one on summary statistics and one on visualizations). The group projects will be evaluated in terms of presentations and written reports. Groups of two will be self-selected during Week 2 and remain the same for both projects. In cases of significantly unequal contribution documented through peer evaluation, individual grades may be adjusted. Both group members must participate in presentation delivery.

Presentation 1 and Report 1 must include at least three different summary statistics and comparisons across groups or categories, presented in clear tables. Presentation 2 and Report 2 must include at least four graphs following design principles covered in class, integrated into a coherent data story. The group will complete both projects using an instructor-provided dataset.

Assessment Task	Contribution to Overall Course grade (%)	Due date
Attendance	15%	/
Group Presentation	15% x 2	12 March & 2 April *
Project Report	15% x 2	5 PM, 13 March & 3 April
Final Examination	25%	7 May *

* Assessment marks for individual assessed tasks will be released within two weeks of the due date.

Mapping of Course ILOs to Assessment Tasks

Assessed Task	Mapped ILOs	Explanation
Attendance	ILO1−3	Attendance and active participation in class allows students to engage with course concepts in real-time, practice R coding techniques during computing sessions, and develop hands-on data analysis skills through in-class exercises.
Group Presentation	ILO1−3	Group presentations require students to demonstrate their understanding of quantitative data analysis concepts, apply computer-based techniques for data manipulation and analysis, and present statistical findings from real-world data in a clear and compelling manner. Students work collaboratively to analyze datasets and communicate their insights effectively.
Project Report	ILO1−3	Group project reports require students to demonstrate their understanding of quantitative data analysis concepts, conduct computer-based data extraction, management, and analysis using R, and apply statistical skills to real-world data with clear written communication.
Final Exam	ILO1 & 3	The exams assess students' conceptual understanding of quantitative data analysis from a social scientific perspective as well as statistical skills that can be appliable to real-world data.

Grading Rubrics

Group Presentation: Each group presentation is 6 minutes. Assessment is based on four criteria: Analysis Quality (35%) evaluates whether appropriate methods are applied correctly, the exploration is thorough, and limitations are acknowledged. Visuals/Tables (25%) examines clarity, formatting, and adherence to design principles covered in class. Interpretation (25%) evaluates whether the analysis is insightful and addresses the research question meaningfully. Delivery (15%) considers clear communication and balanced participation among group members.

Project Report: Each report should be 3−5 pages with R code in an appendix. Assessment is based on six criteria: Research Question (5%) evaluates whether the question is clear, well-motivated, and appropriate for the data and methods. Methods (5%) assesses the description of data and methods. R Code (20%) examines whether code is efficient, well-documented, and fully reproducible. Results (20%) evaluates the clarity, formatting, and labeling of tables or figures. Interpretation (25%) assesses whether findings address the research question and acknowledge uncertainty. Writing Quality (10%) considers organization, clarity, and professionalism. Reflection (15%) assesses students' critical thinking about their analytical process, including what worked well, what challenges they encountered, limitations of their approach, and what they would do differently.

Attendance: Students are allowed 2 absences without penalty or documentation required. Additional absences result in a 2% deduction each (maximum 15%) unless excused with official

documentation (medical certificate, university-approved leave). Documentation should be submitted within two weeks of the absence when possible. In cases of extended illness or emergency, students should contact the instructor to discuss alternative arrangements.

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

The final exam will be performed under strict exam conditions. However, students may use AI (in tandem with other tools) to help prepare them for the exams. For the in-class activities and group projects, students are welcome to use generative AI to support their learning in completing the project. However, students should understand that generative AI tools should only be used as *tools* and should NOT be a substitute for students' own work.

Communication and Feedback

Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. Feedback on assignments will include comments and suggestions for further improvement. Students who have further questions about the feedback, including scores, should consult the Instructor OR Teaching Assistant within five working days after the feedback is received.

Resubmission Policy

Students who are unable to submit any of the assessed tasks should contact the Instructor or Teaching Assistant **within five working days** after the respective deadlines to discuss arrangements for resubmission.

Required Texts and Materials

There are no textbooks required for this course. Supplementary materials and handouts will be available on Canvas throughout the course.

Students may find the following resources useful to learn basic as well as more advanced R programming techniques:
- https://r4ds.hadley.nz/
- https://swirlstats.com
- https://www.coursera.org/course/rprog

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.