

Welcome to

Introduction to Social Data Analytics

Spring 2024

Why take this class?

Huge demand for quantitative skills related to data analysis, both in academia and industry.

Learning objectives and class goals

1. Introduce you to interesting and important social science questions
2. Show how data can be used to provide answers to these questions
3. Give you tools to perform this analysis through Excel, Stata, and R

All Information in the Syllabus is Tentative

Normally, I would stick to the syllabus.

If I must change something, I will make sure

- to inform you as early as possible and
- to explain as best as I can the rationale behind the change.

Basic Information

Professor:	Yinlin Dai
Office:	Atkinson Hall 6406
Office Hours:	Tuesdays 9:30 AM – 11:00 AM (Over Zoom)
Email:	yid031@UCSD.edu
Class Information:	Mon/Wed <u>in-class lecture</u> 2-2:50 pm (Section A) 3-3:50 pm (Section B)
Location:	MOS 0114
Lab:	Fri, various times/locations

Instructor

- Name: Yinlin Dai
- From: Chengdu, China
- Undergraduate: Southwestern University 16', Georgetown, TX
- PhD: Clemson University, Clemson 22', SC
- Hobby: Badminton, Frisbee, Hiking, Barry's, Coffee Shops/Restaurants

Instructional Team

Pai, Anjali	a1pai@ucsd.edu
Bhalothia, Aakash	abhalothia@ucsd.edu
Mittal, Ruby	rumittal@ucsd.edu
Erbabian, Maddison	merbabian@ucsd.edu
Karim, Muhammad	mukarim@ucsd.edu

- TAs' Office hours are OVER ZOOM! Office Hours Details are posted on Canvas.
- If you have any questions (content or logistics; **NO** technical issues) related to this class:
 - ***First, check the course website, the syllabus, or Piazza.***
 - Office Hours!
 - Outside Office Hours:
 - Post on Piazza!
 - Students who **first** answered questions on Piazza will get extra credits (NO anonymous!!!)
 - If you have a question related to the course that you don't think is relevant for other students, you can send an email to your TA or my email at yid031@ucsd.edu

MATERIALS FOR THE CLASS

All of the course materials will be made available through Canvas.

No textbook to buy!

We will use an online textbook that was written for the course and is available through Canvas for free.

The software we will be using (Excel, Stata, and R) is also available through Canvas. Check the first module which contains information about installation!

Are there any technology requirements?

We will be using Excel, Stata and R throughout the course. You can download these on either PCs or Macs, but **chromebooks are very difficult to download the programs on.**

There are computers in the library that have these programs installed. The Data and GIS lab: <https://library.ucsd.edu/computing-and-technology/data-and-gis-lab/index.html> has computers that have these programs.

However, if you attend lab in person, most of the lab will be using your laptop to complete a coding assignment.

If you do not have a laptop that is capable of downloading R and Stata, there is also a possibility for a long-term loaner laptop from the University. Please go here: <https://library.ucsd.edu/computing-and-technology/computers-and-laptops/index.html> to receive more information.

LECTURES (Mon/Wed)

- Lecture attendance is strongly encouraged.
 - All lectures will be **recorded and posted** to Canvas (Media Gallery).
 - I can't promise the quality of the recording!
 - I like to write on the board/iPad.
 - In addition, all the lecture material, as well as bonus material, is covered in a series of videos on Canvas.
 - In certain cases, I will ask you to watch a video or read some materials before arriving at class. Make sure you watch or read before coming to class.
 - Student questions/participation during lectures are strongly encouraged.
 - All lectures for each chapter will be posted on Canvas beforehand in PDF format.
 - Class will not be held on Monday (5/27) *in observance of Memorial Day.*

LABS (Friday)

Note:

NOT RECORDED!

Each lab will involve completing an Excel workbook, Stata Do-file, or R script.

- ✓ Work with your classmates
- ✓ With the help of a TA
- ✓ At the end of class, TA will go through the answers to the lab

Each week there is a short quiz related to the lab that you need to fill out in order to receive credit for lab.

If you cannot attend a lab in a given week, you can still receive credit by finishing the lab on your own time and then filling out the quiz related to the lab assignment and turning in the complete lab.

You should feel free to ask questions about quizzes and labs at office hours.

The last question of every lab quiz will be to submit the completed lab script. **If you do not submit a completed lab script, you will automatically receive a zero on the lab, regardless of your other answers.**

Example of a Normal Week

1. Introduce Empirical Application for the week (often highlighting work from someone at UCSD). Make sure to watch the interview videos on Canvas!
2. Learn various commands/functions in a given software
3. Use those newly learned techniques to analyze data from the empirical application of the week
4. Complete quizzes/labs to evaluate progress in understanding the content

Each week there is a “Plan of the Week” that discusses what assessments are due and when

GRADES

10%:	Lab	On Canvas
15%:	Quizzes	On Canvas
25%:	Midterm	Wednesday (5/1), in class
50%:	Final exam	Saturday (6/8), 11:30am-2:29pm, Loc: TBA

- **The final exam is cumulative! But will be weighted much more heavily towards materials (week 6-10).**
- **If you arrive late to an in-person exam, I will allow you to take the exam in the time that remains as long as no one has turned in his/ her exam and left the room. Once a classmate has turned in his/her exam, you will earn a zero on the test if you arrive late.**

How Course Grades Are Determined

If you have a score above or equal to a 90% for the course you have earned an *A of some type (A+,A, or A-)*.

If you have a score less that is greater than or equal to 80% but less than 90% you have earned no lower than a *B of some type (B+,B, or B-)*.

If you have a score less that is greater than or equal to 70% but less than 80% you have earned no lower than a *C of some type (C+,C, or C-)*.

If you have a score less that is greater than or equal to 60% but less than 70% you have earned no lower than a **D**.

I usually do not bump up your grade. If you get 89.1%, you might get a B.

➤ If you worry this might happen, please study harder throughout the semester.

Lab (10%)

- Online quizzes associated with lab need to be completed **on Canvas**.
 - Each week corresponds to one lab quiz!
- *Lab quizzes will **ALWAYS** be due by Sunday at 11:59 PM.*
- If you attend lab, the answers for the lab will be covered during the lab itself.
- If you do not attend, you can still get credit by completing the lab on your own time and completing the associated Canvas quiz.
- The last question of every lab quiz will be to submit the completed lab script. **If you do not submit a completed lab script, you will automatically receive a zero on the lab, regardless of your other answers.**
- If you submit your lab quiz after the due time, it will be marked as late submission and you will get zero for it. Please **DO NOT** redo any lab quizzes at any time after the deadlines, or these completed assignments will be counted as late submissions.

Quizzes (15%)

- Quizzes need to be completed **on Canvas.**
 - **Separate from lab quizzes.**
 - Each week corresponds to one quiz!
 - Two attempts!
 - Some questions might look different from the first attempt to the second.
- *Quizzes will be posted on Wednesday night and must be completed by Friday at 11:59 PM.*
- If you submit your quiz after the due time, it will be marked as late submission and you will get zero for it. Please **DO NOT** redo any quizzes at any time after the deadlines, or these completed assignments will be counted as late submissions.
- The **two** lowest assignment scores will be dropped before computing your homework grade to cover circumstances such as “I forgot,” “my computer died,” “I felt sick,” etc.

Midterm (25%)

- The only valid excuses for missing midterm are:
 1. severe illness.
 2. death in the immediate family (or severe illness).
 3. school-sponsored event.

- If you miss a midterm for one of these reasons, you must adhere to the following conditions:
 1. provide me with a valid and verifiable written excuse **AT LEAST 5 HOURS AHEAD**;
 2. permission must be granted to miss the midterm **PRIOR** to the deadline;
 3. the weight of the midterm will be shifted to the final exam.

- Other excuses are invalid and will warrant a zero on midterm!
- Going to doctor's appointments is not a valid excuse for missing any assignments. There are no exceptions to these policies.

ACADEMIC INTEGRITY

Academic dishonesty will not be tolerated. All suspected cases of academic dishonesty will be reported to the Academic Integrity Coordinator. Students found guilty of academic dishonesty will earn a failing grade for the course in addition to the penalties imposed by the Academic Integrity Review Board.

The following are just a few examples of academic dishonesty:

- Using unauthorized materials during an exam.
- Copying another student's answers for a homework or during an exam.
- Having someone else take your exam for you or give you the answers.
- Lying about having taken an exam or completed an assignment.

For more information, go to: <http://academicintegrity.ucsd.edu/excel-integrity/define-cheating/index.html>

Week 1: Introduction to Excel

(week 1: Excel)

Week 2: Introduction to Stata

Week 3: Data Wrangling in Stata

Week 4: Regression in Stata

Week 5: Midterm Week

(Week 2-5: Stata)

Week 6: Introduction to R

Week 7: Data Wrangling in R

Week 8: Data Visualization in R

Week 9: Linear Regression in R

Week 10: Functions in R

(Week 6-10: R)

Final Exam

Road Maps

Questions about Syllabus?