Statistical Analysis for Social Sciences
GOV 350K 38930, Fall 2021
MWF 12-1pm
GAR 2.112

Kyle Shen
Department of Government
Office hours: Completely Online
- W 1-4pm
  - Access via Zoom tab on Canvas Book.
  - Select a time slot on https://gov350k2021.youcanbook.me/
- Schedule an appointment (email me directly)

Email: Kyle.shen@utexas.edu

Course Description:

"Correlation is not causation"

Datasets have become ubiquitous, and so are the attempts to explain the world using quantitative analysis. Data can help us understand our world in profound ways, systematically exploring and illuminating relationships among seemingly unrelated factors. Yet, the story that data tells us is never objective, instead emerging from a cumulation of choices made by whomever is telling the story.

Faced with an overabundance of data and different interpretations of that data, how do we parse the signal from the noise? What tools can we use to understand data more effectively?

This course offers an introduction to basic methods of describing and analyzing data using statistical methods. It will begin with methods of description and data visualization before proceeding to methods of statistical analysis and linear regression. In brief, these are ways to show and test arguments using data, respectively. After establishing a working knowledge of using and manipulating data, we will explore some limitations of data analysis and how to address them.

This course has a Quantitative Reasoning Flag.

I. Required Readings and Software:

Required:

1) The textbook for this class is ***OpenIntro Statistics (4th Edition)***, by David Diez, Mine Cetinkaya-Rundel and Christopher Barr, available here: https://www.openintro.org/book/os/

2) ***Naked Statistics*** by Charles Wheelan. Available at the co-op or at an e-copy through the library: An e-copy is available through the library: https://search.lib.utexas.edu/permalink/01UTAU_INST/be14ds/alma991035799729706011
3) All additional readings will be posted to canvas.

4) Software for Statistical Analysis:

This course will use R statistical software and RStudio, a programming environment that eases accessibility to the R coding language. Both can be downloaded for free:

R: https://www.r-project.org/
Rstudio: https://www.rstudio.com/products/rstudio/download/#download

Nota Bene: Learning to code often can be a frustrating process, especially if you are new to the process. I will provide guidance in-class and I encourage you to reach out if you have any questions. I also encourage you to budget a significant amount of time for the problem sets and coding exercises. Below are a few resources for coding.

1. Stack Exchange (https://stats.stackexchange.com/questions/tagged/r)
2. Quick R: https://www.statmethods.net/
3. https://www.r-bloggers.com/
4. GG Plot Cheat Sheet (Uploaded to Canvas Files)
5. Chatter (Chat room on Canvas site)

Additionally, the open secret among anyone who codes is typing an error message into Google often will lead you to a solution. If you are running into a problem, the odds are that you aren't the first, and someone else has sought for help online. Stack Exchange (linked above) is an especially useful website. Of course, I always will be available to help with any technical difficulties.

II. Assignments and Grading:

- Problem Sets: 30%
- Midterm: 30%
- Final Paper: 25% (40% total, including short assignments)
    - Research Topic 5%
    - Data Visualization: 5%
    - Proposed Model: 5%

A. Assignments:

- Your grade will consist of one midterm, one final paper and problem sets.

- Problem Sets will be assigned approximately every two weeks and will include assigned problems as well as extensions of coding exercises begun in class.

- The mid-term examination will be during class time on October 22. It will be closed book.

- The final grade will be a research paper asking you to use a dataset to answer a research question. To help you prepare for the final paper, I also will assign short assignments throughout the course to develop your topic and think through your analysis. The final paper will be due December 10.

B. Grading Scale:

$93 \leq$ A	$80 \leq$ B- < 83	$67 \leq$ D+ < 70
$90 \leq$ A- < 93	$77 \leq$ C+ < 80	$63 \leq$ D < 67
$87 \leq$ B+ < 90	$73 \leq$ C < 77	$60 \leq$ D- < 63
$83 \leq$ B < 87	$70 \leq$ C- < 73	F < 63

III. Policies and Procedures:

a. Office Hours over Zoom, by Appointment:

My scheduled office hours are immediately after class on Mondays from 1-4pm. Office hours will be held over Zoom. Please email me to schedule an appointment.

My office hours are your time. I am more than happy to continue conversations from class in office hours. Furthermore, the topics and skills developed in this class are relevant to a wide variety of academic and professional settings, so you should feel free to chat about anything on your mind.

b. Late Assignment and Make-up Exam Policy:

For all assignments, a late penalty of one-third of a letter grade (3.3%) will be deducted for each 24-hour period that passes after the end of class on the due date. Turning in an assignment during the 24 hours after the due date turns an A paper becomes an A-, then a B+ the next day, then a B, etc.

If you miss the midterm examination, I will schedule a single time during the finals study period at the end of the semester for all students who need to take the midterm exam at a later date. I will not schedule a make-up during the semester.

c. Grading Feedback and Email Policy:

If you have a question over how something was graded, you have two weeks from the day that the assignment was handed back to challenge the grade. Grade challenges must be made in person.

Substantive discussion must take place in person. I will not answer substantive questions about the course or about grading through email. If you would like to discuss feedback, you may come to office hours.

d. Attendance Policy:

This class will not have an attendance policy. If you miss class, you do not need to explain your absence. However, you will be responsible for keeping up with assignments and contents of the course. Keeping up is especially important in a statistics course, whose material be cumulative over the semester.

Lectures, slides, and assignments will be posted online. I encourage you to schedule an office hour with me to clarify any details if you miss class.

Given the continued status of the pandemic, I expect there to be periodic absences due to illness or exposure. If you feel the slightest bit ill or under the weather for any reasons, or if you suspect you may have been exposed to Covid, please stay at home and schedule a covid test with Health Horns: (https://www.healthyhorns.utexas.edu/coronavirus_proactive_testing.html).

e. Lectures Online:

This class is using the Lectures Online recording system. This system records the audio and video material presented in class for you to review after class. Links for the recordings will appear in the Lectures Online tab on the Canvas page for this class. You will find this tab along the left side navigation in Canvas.

To review a recording, simply click on the Lectures Online navigation tab and follow the instructions presented to you on the page. You can learn more about how to use the Lectures Online system at http://sites.la.utexas.edu/lecturesonline/students/how-to-access-recordings/.

You can find additional information about Lectures Online at: https://sites.la.utexas.edu/lecturesonline/.

f. Academic Dishonesty:

This class has a no tolerance policy for academic dishonesty. Instances of academic dishonesty may be met with, among other measures, grade penalties, failure in the class for the semester, and reporting to the Dean of Students.

IV. University Resources for Students

a. Services for Students with Disabilities

Please let me know if you experience any barriers to learning so I can work with you to ensure you have equal opportunity to participate fully in this course. If you are a student with a disability, or think you may have a disability, and need accommodations please contact Services for Students with Disabilities (SSD). Please refer to SSD's website for contact and more information: http://diversity.utexas.edu/disability/. If you are already registered with SSD, please deliver your Accommodation Letter to me as early as possible in the semester so we can discuss your approved accommodations and needs in this course.

b. Counseling and Mental Health Center

The Counseling and Mental Health Center serves UT's diverse campus community by providing high quality, innovative and culturally informed mental health programs and services that enhance and support students' well-being, academic and life goals. To learn more about your counseling and mental health options, call CMHC at (512) 471-3515.

If you are experiencing a mental health crisis, call the CMHC Crisis Line 24/7 at (512) 471-2255.

c. The Sanger Learning Center

All students are welcome to take advantage of Sanger Center's classes and workshops, private learning specialist appointments, peer academic coaching, and tutoring for more than 70 courses in 15 different subject areas. For more information, please visit http://www.utexas.edu/ugs/slc or call 512-471-3614 (JES A332).

i. **Undergraduate Writing Center**: http://uwc.utexas.edu/
ii. **Libraries**: http://www.lib.utexas.edu/
iii. **Information Technology Services**: https://its.utexas.edu/
iv. **Student Emergency Services**: http://deanofstudents.utexas.edu/emergency/

d. Behavioral Concerns and Safety

If you have concerns about the safety or behavior of fellow students, TAs or Professors, call BCAL (the Behavior Concerns Advice Line): 512-232-5050. Your call can be anonymous. If something doesn't feel right – it probably isn't. Trust your instincts and share your concerns.

e. Title IX Reporting

Title IX is a federal law that protects against sex and gender-based discrimination, sexual harassment, sexual assault, unprofessional or inappropriate conduct of a sexual nature, dating/domestic violence and stalking at federally funded educational institutions. When unprofessional or inappropriate conduct of a sexual nature occurs in our community, the university can:

1. Intervene to prevent harmful behavior from continuing or escalating.
2. Provide support and remedies to students and employees who have experienced harm or have become involved in a Title IX investigation.
3. Investigate and discipline violations of the university's relevant policies.

Beginning January 1, 2020, Texas Senate Bill 212 requires all employees of Texas universities, in cluding faculty, report any information to the Title IX Office regarding sexual harassment, sexual assault, dating violence and stalking that is disclosed to them. Texas law requires that all employees who witness or receive any information of this type (including, but not limited to, writing assignments, class discussions, or one-on-one conversations) must be reported. I am a Responsible Employee and must report any Title IX related incidents that are disclosed in writing, discussion, or one-on-one. Before talking with me, or with any faculty or staff member about a Title IX related incident, be sure to ask whether they are a responsible employee.

If you would like to speak with someone who can provide support or remedies without making an official report to the university, you may make an appointment to speak with an advocate at https://deanofstudents.utexas.edu/emergency/advocate_appointment.php. For more information about reporting options and resources, visit http://www.titleix.utexas.edu/, contact the Title IX Of fice via email at titleix@austin.utexas.edu, or call 512-471-0419.

f. **Academic Integrity**

Each student in the course is expected to abide by the University of Texas Honor Code: "As a student of The University of Texas at Austin, I shall abide by the core values of the University and uphold academic integrity." Plagiarism is taken very seriously at UT. Therefore, if you use words or ideas that are not your own (or that you have used in previous class), you must cite your sources. Other wise you will be guilty of plagiarism and subject to academic disciplinary action, including failure of the course. You are responsible for understanding UT's Academic Honesty and the University Honor Code which can be found at the following web address: https://deanofstudents.utexas.edu/conduct/standardsofconduct.php

g. **Q Drop Policy**

If you want to drop a class after the 12th class day, you'll need to execute a Q drop before the Q-drop deadline, which typically occurs near the middle of the semester. Under Texas law, you are only allowed six Q drops while you are in college at any public Texas institution. For more information, see: http://www.utexas.edu/ugs/csacc/academic/adddrop/qdrop

h. **Sharing of Course Materials is Prohibited**

No materials used in this class, including, but not limited to, lecture hand-outs, videos, (quizzes, exams, papers, projects, homework assignments), in-class materials, review sheets, and additional problem sets, may be shared online or with anyone outside of the class unless you have my explicit, written permission. Unauthorized sharing of materials promotes cheating. It is a violation of the University's Student Honor Code and an act of academic dishonesty. I am well aware of the sites used for sharing materials, and any materials found online that are associated with you, or any suspected unauthorized sharing of materials, will be reported to Student Conduct and Academic Integrity in the Office of the Dean of Students. These reports can result in sanctions, including failure in the course.

i. **Class Recording**

Class recordings are reserved only for students in this class for educational purposes and are protected under FERPA. The recordings should not be shared outside the class in any form. Violation of this restriction by a student could lead to Student Misconduct proceedings.

V. Course Schedule

Part I: Introduction to Probability and Statistics

1) Week 1: What is Data? (8/25, 8/27)
   - Read: Wheelan, Conclusion

2) Week 2 Introduction to Probability and Statistics (8/30, 9/1, 9/3)
   - Read: Diez et al., 12-27, Wheelan Ch. 10

3) Week 3: Description (9/8, 9/10)
   - Read: Diez et al., 41-55, Wheelan Ch. 2

4) Week 4: Probability Redux (9/13, 9/15, 9/17)
    - Read: Diez et al., 81-91, 95-108

5) Week 5: Estimation (9/20, 9/22, 9/24)
    - Read: Diez et al., 115-127

6) Week 6: Estimation, Part 2 (9/27, 9/29, 10/1)
    - Read: Diez et al., 133-144, 149-155
    - ***Research Topic due Oct 1****

Part II: Hypothesis Testing

7) Week 7: Hypothesis Testing (10/4, 10/6, 10/8)
    - Read: Diez et al., 170-178, 181-186

8) Week 8: Hypothesis Testing, Part 2 (10/11, 10/13, 10/15)
    - Read: Diez et al., 189-201, 208-213

9) Week 9: Difference of Means and ANOVA (10/18, 10/20, 10/22)
    - Read: Diez et al., 217-224, 251-258
    - ***Midterm Oct 22****

Part III: Regression

10) Week 10: Univariate Regression: (10/25, 10/27, 10/29)
    - Read: Diez et al. 305-311, 322-324, 331-334
    - ***Data Visualization due Oct 29****

11) Week 11: Multivariate Regression, Part 1 (11/1, 11/3, 11/5)
    - Read: 346-349, 353-356

12) Week 12: Multivariate Regression, Part 2 (11/8, 11/10, 11/12)
    - Read: Wheelan, Ch. 12

13) Week 13: Conceptualization and Measurement (11/15, 11/17, 11/19)
    - ***Proposed Models Due Nov 15****
    - Read: Wheelan, Ch. 3 (on Canvas); Elkins, Collier & Levitsky

14) Week 14: Discuss papers (11/22)

15) Week 15: Causal Inference and Experiments (11/29, 12/1, 12/3)
    - Read: Angrist & Pischke, DuFlo

16) Week 16: How can I use statistics in the future? Q&A with three experts (12/6)

****Final paper due Dec 10****